



**HAL**  
open science

# Répression Polycomb et modèle poisson zèbre en cancérologie

Ludivine Raby

► **To cite this version:**

Ludivine Raby. Répression Polycomb et modèle poisson zèbre en cancérologie. Cancer. Université de Lille, 2023. Français. NNT : 2023ULILS101 . tel-04228877

**HAL Id: tel-04228877**

**<https://theses.hal.science/tel-04228877>**

Submitted on 4 Oct 2023

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

UNIVERSITE DE LILLE

LABORATOIRE CANTHER UMR9020-U1277 INSERM

## THESE DE DOCTORAT

Pour l'obtention du titre de

### DOCTEUR DE L'UNIVERSITE DE LILLE

École Doctorale Biologie-Santé de Lille (ED446)

Aspect moléculaires et cellulaires de la biologie

---

## RÉPRESSION POLYCOMB ET MODÈLE POISSON ZÈBRE EN CANCÉROLOGIE

---

Présentée et soutenue publiquement par

**Ludivine RABY**

Le 27 Janvier 2023

### Membres du Jury :

Président	<b>Pr. Xuefen Le Bourhis</b> Laboratoire Hétérogénéité, Plasticité et Résistance aux Thérapies des Cancers (CANTHER), Université de Lille
Rapporteurs	<b>Dr. Christelle Golzio</b> Institut de Génétique et de Biologie Moléculaire et Cellulaire (IGBMC), Université de Strasbourg <b>Dr. Jean-Pierre Levraud</b> Institut des Neurosciences Paris-Saclay (NeuroPSI), Université de Paris-Saclay
Examineurs	<b>Dr. Patrick Blader</b> Centre de Biologie Intégrative (CBI), Université de Toulouse <b>Dr. Karima Kissa</b> Laboratoire des Interactions Hôte Pathogène (LPHI), Université de Montpellier <b>Pr. Florence Petit</b> Clinique de Génétique Guy Fontaine, Unité des Maladies rares du Développement et du métabolisme (RADEME), Université de Lille, CHRU Lille
Directeur de thèse	<b>Pr. Pierre-Olivier Angrand</b> Laboratoire Hétérogénéité, Plasticité et Résistance aux Thérapies des Cancers (CANTHER), Université de Lille



# REMERCIEMENTS

---

Je souhaite remercier les membres du Jury, Monsieur le Docteur **Patrick Blader**, Madame la Docteur **Christelle Golzio**, Madame la Docteur **Karima Kissa**, Monsieur le Docteur **Jean-Pierre Levraud**, Madame la Professeur **Florence Petit**, pour avoir accepté d'évaluer ce travail et d'avoir été jusqu'au bout de ce manuscrit. Une mention spéciale pour Madame la Professeur **Xuefen Le Bourhis**, examinatrice mais surtout cheffe de la team plasticité, je vous remercie d'avoir transmis ma candidature de M1 à **Pierre-Olivier**.

Je souhaite également remercier mon directeur de thèse, le Professeur **Pierre-Olivier Angrand**. Ce n'était pas évident de miser sur moi au départ. Mais comme le veut l'adage, le jean troué ne fait pas le moine, et vous avez choisi de croire en moi. Pour cela merci. Vous m'avez quelques fois poussé dans mes retranchements, poussé à travailler mieux, à réfléchir mieux. Et surtout, vous m'avez offert une place dans la dream team zebra. Merci pour tout.

Ma chère **Pamela**, je pourrais écrire une seconde thèse pour te remercier. La vie au laboratoire n'aurait certainement pas la même saveur sans toi. Cette thèse, c'est en majeure partie parce que tu étais là pour me former, me guider et me poser les questions justes. Mais tu es avant tout une très belle personne, pleine de sensibilité. Merci d'être toi, merci pour les cafés noisette pour sauver les mauvaises journées, merci pour cette sortie RESAMA, merci pour les stickers emoji, merci pour la garde de Mimi chat.

**Mariette**, j'espère être là pour te voir soutenir, car je ne doute pas que tu réussisses ta thèse au moins aussi bien que ton M2. Cette année de M2 a été riche en émotions, mais surtout en rires. Je ne pense pas qu'on puisse citer la moitié des dossiers classifiés dans ces remerciements, mais je ne les oublierai pas, promis. Je te remercie pour les Wham en période de Noël, pas pour les Sunlights des tropiques de Gilbert Montagnier. Je te souhaite de réussir à mener à bien ton projet et de t'épanouir dans ce doctorat.

**Wouroud**, merci pour les journaux dans un premier temps, on commence toujours mieux une journée en lisant l'horoscope. Tu nous as fait découvrir la culture tunisienne, ses délicieuses pâtisseries et son merveilleux couscous. Tu es très courageuse de partir loin de ta famille pour venir dans la grisaille du nord poursuivre tes rêves. Je te souhaite également toute la réussite possible pour ta thèse.

Je tiens à remercier l'entièreté des membres de la team Plasticité, déjà pour avoir supporté mes vocalises régulières dans les labos, ce n'était pas une mince affaire. Mais je les remercie surtout pour la pertinence de leurs remarques et questions sur mes travaux, qui m'ont permis de progresser durant



cette thèse. Je remercie particulièrement : **Eric** pour l'apprentissage des mots fléchés du 20 min et de m'avoir appris la patience derrière le micro-onde et surtout l'aide précieuse sur des problématiques scientifiques complexes ; **Roland** pour votre jovialité et vos anecdotes croustillantes ; **Chann** pour l'organisation des événements de convivialité au laboratoire et ta bonne humeur ; **Samuel** et **Alessandro** pour votre expertise et votre bienveillance ; **Robert-Alain** pour la gestion du labo ; **Valérie** pour ta disponibilité et ton aide dans les démarches administratives ; **Véronique** pour ton travail indispensable et ta bonne humeur. **Isa**, nos sorties « noisettes » me manquent et je te remercie pour les vidéos TikTok dont tu nous fais l'honneur.

Je remercie également tous les étudiants de la team. J'ai d'abord une pensée pour tous les anciens, **Romain, Jordan, Clément, Nadège, Ihsan**. Vous avez instauré un esprit d'équipe, d'amitié et de transmission dans ce laboratoire, et cette ambiance participe à la réussite de tous. Je n'oublie pas votre aide précieuse pour la préparation au concours de l'ED. Je remercie tous les étudiants de l'équipe pour m'avoir suivie dans des projets aussi fous que l'organisation d'un atelier de médiation scientifique à la fête de la Science, et ce à deux reprises. Je vous suis reconnaissante pour votre confiance et votre investissement qui ont permis d'expliquer notre travail au grand public et qui sait, peut-être éveiller des vocations chez les plus jeunes ? Merci aussi d'avoir fait vivre ces réunions un peu moins scientifiques que sont les mardis en folie, renommer depuis peu les mercredis en folie. Le travail en équipe est toujours plus facile et agréable quand des relations de camaraderie unissent les protagonistes. **Marie**, tu es un exemple de persévérance et je te souhaite de trouver un travail ou un post-doc dans lequel tu puisses t'épanouir. **Sarah**, j'espère que tu arriveras à trouver la paix dans ton travail. Mathilde, une année supplémentaire pour finir cette thèse, et te permettre de trouver les derniers horcruxes dans tes cellules souches cancéreuses. Je n'oublie pas la solidarité, l'année de notre M2 avec **Evodie**, et nos (très) nombreuses répétitions qui nous ont permis de passer l'année avec succès. **Julien**, tu es bien évidemment le meilleur, c'est indiscutable, alors merci pour la mise à profit de tes skills et pour les petites gourmandises du Japon. **Edouard**, ceux-ci ne sont pas des remerciements, ceci est un contrôle de Police. **Martine**, merci pour la découverte du resto la famille et pour ces minis aventures en fiat 500. Sir **Alexandre van Outryve**, je te remercie pour ton aide dans l'organisation de Doct'O'Lille, j'espère que tu sauras instaurer le même esprit de camaraderie qui règne dans l'équipe à ONCOLille. **Geoffrey**, merci pour ces bons moments au CNO. **Thomas, Julia** et **Aurélien**, vous avez animé les pauses déjeuner. **Mayar**, merci pour ta bonne humeur et les pâtisseries qui sont toujours délicieuses. **Amina**, tu persévères dans la joie et la bonne humeur, je ne doute pas que tu réussisses ce travail brillamment. Les doctorants de la team des 7 nains, je regrette de ne pas avoir eu plus de temps pour mieux vous connaître. **Anaïs, Éloïse, Flavie, Joséphine, Kekely, Paul**, je vous remercie pour les bons moments passés autour d'un barbecue ou dans un bar à jouer à des jeux. Je

vous remercie tous d'avoir été là pour m'accompagner sur cette course de fond que représente la thèse.

Je tiens à remercier les étudiants de la team Mucines, pour leur accueil dans leur open-space à ONCOLille, leur bonne humeur, et le soin de ma petite plante papillon en mon absence.

Ma chère **Élise**, ce travail, je te le dois aussi. J'avais cru être préparée en te voyant affronter la tempête avant moi, mais je m'étais trompée. Plus que ton aide, tu m'as offert une oreille attentive, une épaule, un soutien. Sans toi à la barre, je ne pense pas que je sois un jour arrivée à terminer ce périple. Je sais à quel point cela a été difficile pour toi, à quel point ça l'est toujours actuellement. L'accalmie arrive et je vous souhaite, à toi et **Cédric**, le meilleur, vous le méritez.

A mes amis, je vous remercie d'avoir été là durant ce parcours du combattant. **Moana**, l'âne TroTro ainsi que tes blagues plutôt bancales sur les hiboux me suivront toute ma vie grâce à toi. Je suis contente de t'avoir rencontré, un peu moins de t'avoir suivie pour cet Halloween mythique sur le thème du harcèlement de rue. Mais à la fin, on finit toujours par en rire, même en attendant des frites pendant 45 min. Mes collègues rémois, **Antoine, Sarah, Amélie**, quand je repense à nos cours de géologie et à mon travail actuel, j'ai l'impression que ça s'est passé il y a une éternité, mais les nouvelles régulières changent les idées. **Lucile**, même parcours, mêmes galères à l'autre bout de la France. Parfois, voir qu'on traverse plus ou moins tous les mêmes épreuves rend les choses moins compliquées. Ton oreille attentive et tes anecdotes ont été les bienvenues.

Enfin, je remercie les personnes qui m'ont construite et qui me soutiennent depuis toujours, la famille. Ma belle-famille, vous m'avez toujours accueillie comme une membre à part entière. Belle-maman, je ne peux même pas faire toutes les blagues sur les belles mères, puisque vous m'avez toujours considéré comme votre propre fille. Merci à tous d'avoir supporté mon stress durant ces derniers mois.

A mes cousins et cousines, **Anaïs, Camille, Jérémy, Julie, Rudy**, et leur conjoint ou conjointe, merci à vous d'avoir été là également. J'espère que de nombreux pic-nics au bord du lac, avec une guitare gitane, et des pétales salés nous attendent. Merci à vous d'être là, ensemble on ne voit pas le temps qui passe et surtout on ne vieillit pas. A mes parrains et marraine, **Nono** et **Pepette**, je vous dois beaucoup, des anges gardiens sur qui on peut toujours compter. A toutes les tatas et tous les tontons de la famille Ferrari, vous me manquez tous et j'ai hâte de fêter ça avec vous, il est grand temps qu'on réunisse la famille. Les **Diaz**, merci aussi pour les parties de pétanque l'été. **Romain**, mon loupiot, merci d'avoir été là, merci pour tous les montages photos complètement fous, merci pour les rires salvateurs. Tu es le plus beau cadeau que les parents m'aient fait. Je suis très fière de toi, de l'homme que tu es devenu.

**Papa**, même si ce n'était pas facile, tu as toujours fait l'effort d'essayer de comprendre mon travail. Toujours là pour moi, que ce soit des heures au téléphone, pour un déménagement ou parce que je n'arrive pas à changer ma poulie Dampier. Tu nous as transmis cette curiosité pour la nature et la volonté de tout comprendre, qui me sont indispensables aujourd'hui. Ce travail est aussi un peu le tien.

**Orel**, des années qu'on traverse les pires épreuves et les meilleurs moments ensemble. Une vie ne sera surement pas suffisante pour te dire combien je te dois. Je remercie pour ton aide et ton amour au quotidien. Je te remercie aussi pour ta patience pendant les moments de ralage extrême, mes crises d'angoisse, mes coups de folie. Mais le meilleur nous attend encore et pour longtemps. Love.

**Maman**, cela n'a pas été facile sans toi, mais j'espère que tu es fière. Tu nous manques à tous, pas une journée ne passe sans que nos pensées ne t'accompagnent.

*« Le doute est une force.*

*Une vraie belle force.*

*Veille simplement qu'elle te pousse toujours en avant. »*

**Pierre Bottero**

# RESUME

---

La répression Polycomb exerce une régulation transcriptionnelle par deux complexes protéiques distincts, les complexes PRC1 et PRC2. Ce contrôle épigénétique est médié par la marque épigénétique H2AK119ub1 déposée par PRC1 et la marque H3K27me3 catalysée par PRC2. Découvert d'abord chez la drosophile, les deux complexes ont été conservés et se sont diversifiés au cours de l'évolution. Une multiplication du nombre de protéines de la famille Polycomb est observée chez les vertébrés. Ces protéines peuvent moduler l'action des complexes et participer au recrutement sur des cibles génomiques spécifiques. La répression Polycomb applique une régulation négative sur de nombreux gènes impliqués dans des grands processus biologiques mais aussi dans l'apparition et la progression de tumeurs. Cette étude s'intéresse donc au rôle de la répression Polycomb dans le développement et dans le cancer, sur le modèle poisson zèbre.

Les gènes *cbx* sont conservés chez les vertébrés et plus particulièrement chez les téléostéens. Les gènes *cbx* sont tous exprimés au cours du développement et au stade adulte chez le poisson zèbre, avec des profils d'expression similaires, même si on dénote quelques différences. Si il y a donc des redondances fonctionnelles, il semble exister des spécificités fonctionnelles. L'histoire évolutive de *Ezh1* et *Ezh2* a été étudiée par des analyses *in silico*, et *Ezh1* semble être le fruit d'une duplication de *Ezh2*. L'expression de *ezh2* et *ezh1* est inversement proportionnelle au cours du développement. Une étude fonctionnelle de *Ezh2* et *Ezh1* révèle que la compensation de la perte de fonction de *Ezh2* par *Ezh1* est minime. L'activité totale du complexe PRC2 dans le développement du poisson zèbre a été analysé par la génération d'un nouveau mutant *eed*. La mutation est létale autour de 11 jpf en raison d'altérations importantes du système digestif. La perte de la marque H3K27me3 altère la différenciation des cellules neurales. Les larves *eed*<sup>-/-</sup> reproduisent le paysage épigénétique observé dans certains cancers dont l'activité PRC2 est inhibée. La lignée mutante pour *eed* est donc un modèle pertinent pour l'étude de cancers impliquant la répression Polycomb, et notamment les DIPGs.

Enfin, deux revues de littérature ont été écrites, afin de rassembler les connaissances actuelles sur les expérimentations de transplantations de cellules cancéreuses chez le poisson zèbre, ainsi que l'utilisation de l'ingénierie génétique dans la modélisation de cancers sur le modèle poisson zèbre.



# ABSTRACT

---

The Polycomb repression exerts transcriptional control by two distinct complexes, PRC1 AND PRC2. This epigenetic control is ensured by H2AK119ub1 and H3K27me3 deposition. The polycomb repression was firstly discovered in *Drosophila*. The analyses of genic and proteic sequences revealed that the Polycomb protein family has been conserved and diversified through evolution. A high number of Polycomb family proteins is present in vertebrates. These proteins can modulate the activity the action of the complexes and take part to their recruitment to specific genomic loci. The Polycomb repression negatively regulates many genes involved in major biological processes, but also in the initiation and progression of cancers. Therefore, this study focuses on the role of Polycomb repression in development and in cancer, based on the zebrafish model.

The *cbx* genes are conserved in vertebrates and more particularly in teleosts. The *cbx* genes are all expressed during development and in different adult organs. The expression patterns are very similar between the orthologues, although there are some differences. These results suggest the existence of functional specificity. The evolutionary history of Ezh1 and Ezh2 has been studied by *in silico* analyses. Ezh1 appeared to be the result of an Ezh2 duplication. The expression of *ezh2* and *ezh1* is inversely proportional during development. A functional study of Ezh2 and Ezh1 reveals that the compensation for the loss of function of Ezh2 by Ezh1 is slight. The total activity of the PRC2 complex in the zebrafish development was analyzed thanks to the generation of a new *eed* mutant line. The mutation is lethal around 11 dpf due to significant alterations of digestive organs. Loss of the H3K27me mark impairs neural cell differentiation. The *eed*<sup>-/-</sup> larvae reproduce the epigenetic landscape observed in tumors with an inhibition of PRC2 activity. The mutant line for *eed* is therefore a relevant model for the study of cancer involving the Polycomb repression.

Finally, two reviews were written in order to bring together current knowledge on cancer cell transplantation experiments, as well as the use of genetic engineering in the modeling of cancers on the zebrafish model.



# ABREVIATIONS

---

## A

---

**AB** : souche **AB**  
**AEBP2** : **AE Binding Protein 2**  
**ADN** : Acide **désoxyribonucléique**  
**AP** : **Antéro-postérieur**  
**ANRIL** : **Antisense RNA in the INK4 Locus**  
**ARF** : **ADP Ribosylation Factor**  
**ARN** : **Acide ribonucléique**  
**ARNm** : **Acide ribonucléique messenger**  
**ARNnc** : **Acide ribonucléique non codant**  
**AUTS2** : **Autisme Susceptibility Gene 2**

## B

---

**BAM** : **β-addition motif**  
**BCOR** : **BCL6 Corepressor**  
**BMI1** : **PCGF4, B Lymphoma Mo-MLV Insertion Region 1 Homolog**  
**BMP** : **Bone morphogenetic protein**

## C

---

**CBX** : **Chromobox**  
**CCL2** : **C-C motif chemokine ligand 2**  
**CCNA2** : **Cyclin A2**  
**CCND1** : **Cyclin D1**  
**CFP** : **Cyan Fluorescent Protein**  
**ChIP** : **Chromatin Immunoprecipitation**  
**CKII** : **Casein Kinase 2**  
**CpG** : **dinucléotide cytosine-guanine**  
**CRISPR/cas9** : **Clustered regularly interspaced short palindromic repeats/CRISPR associated protein 9**

## D

---

**DIPG** : **Diffuse Intrinsic Pontine Glioma**  
**DNMT** : **DNA méthyl transférase**  
**DP1/2** : **E2F Dimerization Partner 1/2**

## E

---

**E2F6** : **E2F Transcription Factor 6**  
**EBF1** : **Early B Cell Factor 1**

**E(z)** : **Enhancer of zeste**  
**EBD** : **EED binding domain**  
**EED** : **Ectoderm Endoderm Development**  
**EHMT** : **Euchromatic histone lysine methyltransferase**  
**ENL** : **MLLT1, Mixed-Lineage Leukemia Translocated To 1**  
**EPOP** : **Elongin BC and Polycomb repressive complex 2 associated protein**  
**ER** : **Estrogen receptor**  
**Esc** : **Extra sex combs**  
**ESC** : **Cellule souche embryonnaire**  
**EZH** : **Enhancer of Zest Homolog**

## F

---

**FBRS** : **Fibrosin**  
**FGF** : **Fibroblast Growth Factor**  
**FT** : **Facteur de transcription**

## G

---

**G9A** : **EHMT2, Euchromatic Histone Lysine Methyltransferase 2**  
**GFAP** : **Glial Fibrillary Acidic Protein**  
**GFP** : **Green Fluorescent Protein**  
**GLP** : **EHMT1**

## H

---

**HER2** : **Human epidermal receptor 2**  
**Hes** : **Hairy And Enhanced of Split**  
**HpF** : **Heure post-fécondation**  
**Hp1** : **Heterochromatin Protein 1**  
**HSC** : **Cellule souche hématopoïétique**

## I

---

**IFNG** : **Interferon gamma**  
**IL** : **Interleukin**  
**INK4** : **CDKN2A, Cyclin Dependant Kinase Inhibitor 2A**  
**IP** : **Immunoprecipitation**

## J

---

**JARID2** : **Jumonji And AT-Rich Interaction Domain Containing 2**



**jmC** : jumonji C  
**Jpf** : Jour post-fécondation

---

## K

---

**KD** : Knock down  
**KDM2B** : Lysine demethylase 2B  
**KO** : Knock out

---

## L

---

**L3MBTL2** : Lethal(3) Malignant Brain Tumor-Like Protein 2  
**LCOR** : Ligand Dependant Nuclear Receptor Corepressor

---

## M

---

**MAG** : Myelin Associated Glycoprotein  
**MAX** : MYC Associated Factor X  
**MGA** : MAX Gene-Associated Protein  
**MHB** : Frontière mésencéphalique  
**MLL** : Mixed-lineage leukemia  
**MPNST** : Malingnat Peripheric Nerve Sheat Tumor  
**MTF** : Metal Regulatory Transcription Factor

---

## N

---

**NAFLD** : Non-Alcoolic Fatty Liver Disease  
**Ngn** : Neurogenin

---

## O

---

**OLIG2** : Oligodendrocyte Transcription Factor 2

---

## P

---

**PAL1** : LCOR  
**Pb** : Paire de bases  
**Pc** : Polycomb  
**PcG** : Polycomb group gene  
**PCGF** : Polycomb group ring finger  
**PCL** : Polycomb-like protein  
**Pdx1** : Pancreatic and Duodenal Homeobox 1  
**Ph** : Polyhomeotic  
**PHC** : Polychomeotic-like protein  
**PHF** : PHD finger protein  
**PHD** : Plant Homeodomain

**PHO** : PleioHomeotic  
**PR** : Progesteron receptor  
**PRC1** : Polycomb repressive complex 1  
**PRC1c** : PRC1 canonique  
**PRC1nc** : PRC1 non-canonique  
**PRC2** : Polycomb repressive complex 2  
**PRE** : Polycomb Response Element  
**Psc** : Posterior sex combs  
**PUFD** : PCGF Ub-like fold discriminator

---

## R

---

**RA** : Retinoic Acid  
**RAWUL** : Ring finger and WD40-associated ubiquitin like  
**RB** : Retinoblastoma protein  
**RBBP** : RB binding protein  
**REST** : RE1 Silencing Transcription Factor  
**RFP** : Red Fluorescent Protein  
**RING1B** : Ring Finger Protein 1B, RNF2  
**RISC** : Complexe d'extinction RNA-induit  
**RNF2** : Ring Finger Protein 2  
**RYBP** : RING1 and YY1 binding protein

---

## S

---

**SAL** : Set-activation loop  
**SAM** : S-adénosylméthionine  
**SANT** : SWI3, ADA2, N-CoR and TFIIB domain  
**SBD** : SANTI1 binding domain  
**Sc** : Sex combs extra  
**SCM** : Sex Comb on Midleg  
**SCMH** : Scm Homolog  
**SET** : Su(var)3-9, Enhancer-of-zeste and Trithorax domain  
**Shh** : Sonic hedgehog  
**SKP1** : S-Phase Kinase Associated Protein 1  
**SRM** : Stimulation-responsive motif  
**Su(z)12** : Suppressor 12 of zeste protein  
**SUZ12** : Suppressor of zeste 12 protein homolog

---

## T

---

**TALEN** : Transcription activator-like effector nuclease  
**TNF** : Tumor Necrosis Factor  
**TrxG** : Protéines du groupe Trithorax  
**TU** : souche Tübingen  
**UV** : Rayons ultra violets

---

**V**

---

**VEFS** : VRN2-EMFS2-FIS2-Su(z)12 domain  
**VIH** : Virus de l'Immunodéficience Humaine

---

**W**

---

**WBD** : **WD40 binding domain**  
**WNT** : **WINT**, Wingless-Type MMTV Integration Site Family

---

**X**

---

**XIST** : **X Inactive Specific**

# TABLE DES MATIERES

---

REMERCIEMENTS .....	3
RESUME.....	7
ABSTRACT.....	9
ABREVIATIONS .....	11
TABLE DES MATIERES.....	4
LISTE DES FIGURES .....	17
LISTE DES TABLES .....	21
INTRODUCTION.....	23
CHAPITRE I : LA REPRESSION POLYCOMB DE LA DECOUVERTE AUX QUESTIONS DE RECHERCHE .....	27
<b>1. Le complexe PRC1.....</b>	<b>29</b>
1.1 <i>Le complexe PRC1 canonique.....</i>	30
<u>1.1.1 Composition des complexes canoniques.....</u>	<u>30</u>
<u>1.1.2 PRC1.2 et PRC1.4 .....</u>	<u>33</u>
1.2 <i>Les complexes PRC1 non-canoniques.....</i>	34
<u>1.2.1 PRC1.1 .....</u>	<u>34</u>
<u>1.2.2 PRC1.3 et PRC1.5 .....</u>	<u>35</u>
<u>1.2.3 PRC1.6 .....</u>	<u>36</u>
1.3 <i>Recrutement des complexes PRC1 .....</i>	36
<u>1.3.1 Recrutement par le chromodomain des CBX .....</u>	<u>37</u>
<u>1.3.2 Recrutement par des ARNnc .....</u>	<u>38</u>
<u>1.3.3 Recrutement par les facteurs de transcription.....</u>	<u>39</u>
<u>1.3.4 Recrutement par les îlots CpG.....</u>	<u>39</u>
<b>2. Le complexe PRC2.....</b>	<b>41</b>
2.1 <i>Composition du complexe .....</i>	41
<u>2.1.1 Les sous-unités catalytiques EZH1 et EZH2.....</u>	<u>41</u>
<u>2.1.2 SUZ12 .....</u>	<u>44</u>
<u>2.1.3 EED .....</u>	<u>44</u>
<u>2.1.4 RBBP4/7.....</u>	<u>46</u>
<u>2.1.5 Les protéines accessoires .....</u>	<u>47</u>
2.2 <i>Recrutement du complexe .....</i>	48
<u>2.2.1 Par le complexe PRC1.....</u>	<u>49</u>
<u>2.2.2 Par les ARNnc .....</u>	<u>49</u>
<u>2.2.3 Par les protéines.....</u>	<u>50</u>
<u>2.2.4 Par les histones modifiées.....</u>	<u>50</u>
<u>2.2.5 Par les îlots CpG.....</u>	<u>51</u>

2.3	<i>Les modifications post-traductionnelles de la lysine 27 de l'histone H3</i> .....	52
2.3.1	<u>La triméthylation</u> .....	52
2.3.2	<u>La diméthylation</u> .....	52
2.3.3	<u>La monométhylation</u> .....	53
2.3.4	<u>L'acétylation</u> .....	54
CHAPITRE II : LA REPRESSION POLYCOMB DANS LE DEVELOPPEMENT .....		55
1.	<b>Le maintien des cellules souches et différenciation.</b> .....	61
2.	<b>Le développement embryonnaire précoce.</b> .....	61
3.	<b>Le développement neural.</b> .....	62
4.	<b>L'hématopoïèse.</b> .....	63
5.	<b>Le système digestif.</b> .....	65
CHAPITRE III : LA REPRESSION POLYCOMB DANS LE CANCER .....		67
1.	<b>Le locus INK4a-ARF-INK4b.</b> .....	68
2.	<b>Les altérations des PcGs.</b> .....	68
3.	<b>Les rôles de la régulation Polycomb dans les hallmarks du cancer.</b> .....	70
4.	<b>Les protéines Polycomb comme marqueur prédictif.</b> .....	70
CHAPITRE IV : LE POISSON ZEBRE, UN MODELE AQUATIQUE QUI A LE VENT EN POUPE .....		69
1.	<b>Le poisson zèbre</b> .....	69
1.1	<i>L'historique</i> .....	69
1.2	<i>La taxonomie</i> .....	70
1.3	<i>L'habitat</i> .....	71
1.4	<i>L'alimentation</i> .....	72
1.5	<i>Cycle de vie</i> .....	72
1.6	<i>Comportements du poisson zèbre</i> .....	73
1.7	<i>Élevage du poisson zèbre en laboratoire</i> .....	74
2.	<b>Le développement du poisson zèbre</b> .....	75
2.1	<i>Le système nerveux central</i> .....	76
2.2	<i>Les organes digestifs</i> .....	80
2.2.1	<u>L'intestin</u> .....	80
2.2.2	<u>Le foie</u> .....	83
2.2.3	<u>Le pancréas</u> .....	85
3.	<b>Les complexes PRC1 et PRC2 chez le poisson zèbre</b> .....	88
3.1	<i>Le complexe PRC1</i> .....	89
3.2	<i>Le complexe PRC2</i> .....	91
OBJECTIFS DE LA THESE .....		92
RESULTATS .....		94
PARTIE I : LA REPRESSION POLYCOMB.....		95

<b>1. Les gènes cbx: The Polycomb Orthologues in Teleost Fishes and Their Expression in the Zebrafish Model</b> .....	<b>95</b>
1.1 <i>Contexte scientifique</i> .....	95
1.2 <i>Article</i> .....	96
1.3 <i>Conclusion</i> .....	120
<b>2. Ezh1 arises from Ezh2 gene duplication but its function is not required for zebrafish development</b> <b>121</b>	
2.1 <i>Contexte scientifique</i> .....	121
2.2 <i>Article</i> .....	122
2.3 <i>Conclusion</i> .....	139
<b>3. Eed : Loss of Polycomb Repressive Complex 2 Function Alters Digestive Organ Homeostasis and Neuronal Differentiation in Zebrafish</b> .....	<b>140</b>
3.1 <i>Contexte scientifique</i> .....	140
3.2 <i>Article</i> .....	141
3.3 <i>Conclusion</i> .....	164
<b>PARTIE II : LE POISSON ZEBRE EN CANCEROLOGIE</b> .....	<b>165</b>
<b>1. La transplantation de cellules tumorales chez le poisson zèbre : de la recherche translationnelle à la médecine personnalisée</b> .....	<b>165</b>
1.1 <i>Contexte scientifique</i> .....	165
1.2 <i>Article</i> .....	166
1.3 <i>Conclusion</i> .....	179
<b>2. Enhancement of Breast Cancer Cell Aggressiveness by lncRNA H19 and Its Mir-675 Derivative: Insight into Shared and Different Actions</b> .....	<b>180</b>
2.1 <i>Contexte scientifique</i> .....	180
2.2 <i>Article</i> .....	181
2.3 <i>Conclusion</i> .....	203
<b>3. Genetic Engineering of Zebrafish in Cancer Research</b> .....	<b>204</b>
3.1 <i>Contexte scientifique</i> .....	204
3.2 <i>Article</i> .....	205
3.3 <i>Conclusion</i> .....	242
<b>DISCUSSION</b> .....	<b>243</b>
<b>1. Le poisson zèbre comme modèle d'étude de la répression Polycomb</b> .....	<b>243</b>
1.1 <i>La conservation des gènes cbx</i> .....	243
1.2 <i>Le complexe PRC2 chez le poisson zèbre</i> .....	245
1.2.1 <u>Ezh1 et Ezh2</u> .....	245
1.2.2 <u>Eed</u> .....	246
<b>2. Répression Polycomb et oncologie sur le modèle poisson zèbre : la modélisation de la mutation H3.3K27M des DIPGs</b> .....	<b>255</b>
<b>CONCLUSIONS</b> .....	<b>261</b>

RÉFÉRENCES : ..... 263



# LISTE DES FIGURES

---

- Figure 1 : Formation du nucléosome et organisation de la chromatine.** **A.** Octamère d’histone H2A, H2B, H3 et H4. Les modifications post-traductionnelles des histones interviennent sur les parties c-terminales des séquences protéiques. **B** Représentation schématique de l’euchromatine. **C.** Représentation schématique de l’hétérochromatine. La conformation spatiale de la chromatine restreint l’accès à la machinerie transcriptionnelle. Le passage d’un état à l’autre est possible par le biais de protéines qui vont être capable de se lier aux marques épigénétiques, de les supprimer et d’en appliquer de nouvelles. .... 24
- Figure 2 : Diversité des protéines du complexe PRC2 au cours de l’évolution.** Les organismes issus du règne *Plantae* sont encadrés en vert et les espèces représentantes du règne *Fungi* sont encadrées en orange. Les organismes du règne *Animalia* sont encadrés en jaune pour les invertébrés et en bleu pour les vertébrés. Figure adaptée de Vijayanathan *et al.*, 2022. .... 28
- Figure 3 : Organisation du complexe PRC1 chez la drosophile et l’homme.** Les sous-unités du complexe PRC1 de la drosophile sont conservées chez l’homme, avec plusieurs orthologues pour chacune des protéines. La fonction du complexe PRC1 est également conservée. La sous-unité catalytique du complexe permet le transfert d’un groupement ubiquitine sur la lysine 119 de l’histone H2A ..... 29
- Figure 4 : Les protéines de la famille CBX chez l’homme.** **A.** Les protéines de la famille CBX, orthologues de Pc, qui participent à la formation des complexes PRC1 canoniques. **B.** Les protéines orthologues de Hp1 forment des agrégats protéiques par homo- et hétérodimérisation. Les triangles bleus représentent les chromodomains, les triangles bleus clairs représentent les shadow chromodomains, les rectangles oranges représentent les domaines C-box. Les représentations graphiques ont été obtenues sur le site de modélisation [http://smart.embl-heidelberg.de/smart/show\\_motifs](http://smart.embl-heidelberg.de/smart/show_motifs). .... 31
- Figure 5 : Représentation graphique des domaines conservés des protéines PCGF.** Les rectangles oranges représentent les domaines RING finger, les rectangles verts représentent les domaines RAWFUL et les bleus les domaines PS. Figure adaptée de Connelly & Dykhuizen, 2017. .... 32
- Figure 6 : Les complexes PRC1 canoniques et non canoniques.** L’inclusion d’une protéine de la famille PCGF dirige la formation du complexe. Les interactions entre le complexe PRC1 et les protéines accessoires sont majoritairement portées par les PCGF. Adaptée de Vidal et Starowicz, 2017. .... 33
- Figure 7 : Recrutement des complexes PRC1.** **A.** Chez la drosophile, le complexe PRC1 est recruté au niveau de des PRE **B.** Le modèle de recrutement hiérarchique nécessite la déposition de la marque H3K27me3 par le complexe PRC2. Cette marque est ensuite reconnue par les protéines CBX. **C.** Les complexes PRC1c et PRC1nc peuvent être recrutés par des ARNnc. **D.** Après fixation sur les séquences RE1, le facteur de transcription REST peut recruter le PRC1 via la fixation des protéines CBX. **E.** La protéine KDM2B reconnaît les îlots CpG et permet le recrutement de PRC1.1. En revanche la méthylation des îlots CpG (ici en rond vert) inhibe le recrutement de PR1.1. Adapté de Connelly & Dykhuizen, 2016 ; Gil & O’Loghlen, 2014. ... 40
- Figure 8 : Composition du complexe PRC2 chez la drosophile et chez les mammifères.** La composition et l’organisation du complexe PRC2 de la drosophile sont conservées chez les mammifères. Les sous-unités catalytiques E(z) et EZH1/EZH2 ajoutent trois groupements méthyle sur la lysine 27 de l’histone H3. .... 41



<b>Figure 9 : Structure schématique de la protéine ezh2.</b> La protéine ezh2 est composée de 951 acides aminés. L'activité enzymatique est portée par le domaine SET. Les autres domaines, et les motifs à Zinc qu'ils contiennent permettent les interactions entre les sous-unités du complexe ainsi que des interactions intramoléculaires. Figure de Jiao & Liu, 2015. ....	<b>43</b>
<b>Figure 10 : Structure de EED. A.</b> La protéine EED est organisée en 7 domaines WD40. Chaque domaine est composé de 4 feuillets $\beta$ antiparallèle. Les 7 WD40 forment une plateforme dédiée aux interactions avec les autres protéines et avec la marque H3K27me3. L'interaction avec la lysine 27 triméthylée est représentée (au centre en gris) schématisée. <b>B.</b> La cage aromatique est composée de 4 acides aminés aromatiques. Ils forment des interactions de van der Waals avec la lysine triméthylée en orange. Réalisée sur PDB ( <a href="https://www.rcsb.org">https://www.rcsb.org</a> ), numéro d'accès : 3JZG. ....	<b>45</b>
<b>Figure 11 : Composition des complexes PRC2.1 et PRC2.2</b> .....	<b>48</b>
<b>Figure 12 : Représentation schématique du contrôle de l'expression des gènes Hox le long de l'axe AP.</b> Les PcGs et TrxGs déterminent le profil d'expression des gènes le long de l'axe antéro-postérieur au cours du développement embryonnaire. Un set de gènes spécifiques est exprimé pour déterminer sa position le long de cet axe. Figure de Piunti & Shilatifard, 2021. ....	<b>62</b>
<b>Figure 13 : Régulation de l'hématopoïèse par les complexes Polycomb.</b> Les complexes PRC1 et PRC2 interviennent à tous les niveaux de la différenciation hématopoïétique. La composition biochimique des complexes varie en fonction du stade de différenciation, ce qui leur permet de réprimer spécifiquement les programmes géniques. HSC : cellule souche hématopoïétique ; MPP : progéniteur multipotent ; progéniteur lymphoïde multipotent ; CMP : progéniteur myéloïde commun ; CLP : progéniteur lymphoïde commun ; GMP : cellule progénitrice de granulocyte-macrophage ; MEP : progéniteur mégakaryocyte-érythroïde ; CD4 : lymphocyte T CD4+ ; CD8 : lymphocyte T CD8+ ; B : lymphocyte B ; NK : cellule natural killer ; Mono : monocyte ; Gran : granulocyte ; Ery : érythrocyte ; Mega : mégakaryocyte. Figure de Di Carlo <i>et al.</i> , 2018. ....	<b>65</b>
<b>Figure 14 : Rôle oncogénique du complexe PRC1 sur le locus INK4a-ARF-INK4b.</b> Lorsque le locus est transcriptionnellement actif, les protéines INK4a et INK4b se lient aux cyclines D. Cette interaction entraîne la phosphorylation de pRb, et maintient la cellule en phase S. La protéine ARF permet d'inhiber l'activité de MDM2, qui consiste en l'ubiquitination de p53 et sa dégradation. L'expression de ces gènes conduit à un arrêt du cycle cellulaire. Lorsque le locus est réprimé par PRC1, les cyclines activent la reprise du cycle cellulaire et la dégradation de p53 lève les freins à la prolifération cellulaire et prévient l'apoptose cellulaire. ....	<b>68</b>
<b>Figure 15 : Répartition géographique des captures de <i>Danio rerio</i> en milieu sauvage au cours du temps.</b> Figure de Engeszer <i>et al.</i> , 2007. ....	<b>69</b>
<b>Figure 16 : Représentation schématique du cycle de vie du poisson zèbre.</b> Les principales divisions cellulaires et mouvements morphogénétiques sont observables en seulement quelques heures. Le temps de génération très court du poisson zèbre en fait un modèle vertébré précieux en laboratoire. hpf : heures post-fécondation ; MBT : Transition midblastula ; Prim-5 : position 5 du primordium de la ligne latérale postérieure ; jpf : jour post-fécondation ; Long-Pec : Long pectoral fins stage. Figure adaptée de Richards, 2011. ....	<b>73</b>

- Figure 17 : Représentation schématique du pattern de neurogenèse au stade neural plate.** Les régions de différenciation neurale sont séparées par des régions d'inhibition de la neurogenèse. Les inhibiteurs sont différents en fonction du positionnement sur la plaque neurale. MH :domaine cerveau moyen/postérieur ; MN : motoneurone spinal ; n : notochorde ; r4 : neurone moteur et sensoriel du rhombomère 2 ; s : neurone sensoriel spinal ; tg : placode trigéminal ; vcc : cluster ventro-caudal. Figure de Bally-Cuif & Hammerschmidt, 2003). ..... **77**
- Figure 18 : Comparaison de l'architecture de l'épithélium intestinale entre les mammifères en (a) et le poisson zèbre en (b).** L'épithélium intestinal des mammifères est composé de villi et de cryptes. Les cryptes contiennent des cellules sécrétrices de Paneth, qui ont un rôle dans le maintien de l'homéostasie de l'intestin et les échanges avec le microbiote. Les cryptes contiennent également des cellules souches qui se divisent de manière asymétrique et génèrent les entérocytes, les cellules calciformes et les cellules entéroendocrines. Les cellules de Paneth et les cryptes de Lieberkühn sont absentes de l'intestin des poissons zèbres. Néanmoins les cellules souches sont présentes à la base des villosités et suivent le même schéma de différenciation en se rapprochant du sommet des villosités. Les cellules immunitaires M sont également absentes du système intestinal des poissons zèbres. La similarité architecturale et fonctionnelle de l'intestin fait du poisson zèbre un très bon modèle d'étude concernant le développement et des pathologies intestinales. Figure de Flores *et al.*, 2019. .... **82**
- Figure 19 : Représentation de l'organogenèse hépatique en fonction du développement.** La phase de croissance est détectée par l'expression de la RFP dans la lignée transgénique Tg(fabp10 : dsred). Figure de Chu & Sadler, 2009. .... **83**
- Figure 20 : Coupe histologique de pancréas de poisson zèbre et humain.** La coloration hématoxyline et éosine révèle un cytoplasme foncé riche en mucus des cellules pancréatiques exocrines. Les cellules claires rassemblées en îlots sont les cellules endocrines. L'organisation du pancréas est similaire chez le poisson zèbre et l'homme. La barre d'échelle représente 1 mm. Figure de Bordeira-Carrico *et al.*, 2022. .... **85**
- Figure 21 : Représentation schématique de la cascade de signalisation à l'origine de l'induction, la spécification et la régionalisation des organes du système digestif.** Les lignes fléchées et barrées indiquent respectivement les relations négatives et positives. Le recrutement de facteurs spécifiques évolue en fonction du développement. Figure de Tiso *et al.*, 2009 ..... **86**
- Figure 22 : Expression relative des ARNm des 8 protéines CBX dans plusieurs types de cancers dans Oncomine.** Les surexpressions et gains de copies sont indiquées en rouges, les sous-expressions ou perte de copie en bleues. Les nombres correspondent au nombre d'études qui ont passé les critères de thresholds. Figure de Li *et al.*, 2020..... **244**
- Figure 23 : L'Injection de morpholinos chez le poisson zèbre. A.** Photographies d'embryons injectés à 72 hpf. La barre d'échelle représente 72 hpf. **B.** Représentation graphique des phénotypes observés en fonction de chaque condition. Le taux de mortalité de plus de 40% parmi les embryons contrôles témoigne d'une ponte de mauvaise qualité. Tous les embryons qui présentent un phénotype sont classés dans la catégorie « malformés » sans discrimination de sévérité. .... **248**
- Figure 24 : Mise en place du cartilage chez les mutants Polycomb. A.** Coloration au bleu alcian de larves sauvages et *rnf2<sup>-/-</sup>* à 72 hpf. ch : ceratohyal ; ep : ethmoid plate ; hys : hyosymplectic ; m : Meckel's cartilage ; pc : parachordal ; pq palatoquadrate ; tc : trabeculae. Figure de van der Veldern *et al.*, 2013. **B.**

Coloration au bleu alcian de larves sauvages et <i>eed<sup>-/-</sup></i> à 9 jpf. Les cartilages sont correctement mis en place chez les larves mutantes. ....	250
<b>Figure 25 : Observation de l'apoptose par immunofluorescence.</b> Les cellules apoptotiques sont marquées à l'aide d'un anticorps dirigé contre la forme active de la caspase 3. La barre d'échelle représente 500 µm. ....	251
<b>Figure 26 : Construction schématique du système GAL4/UAS.</b> Le facteur GAL4 est inséré sous le contrôle d'un promoteur spécifique de la lignée neurale. L'expression de GAL4 permet l'expression de l'oncogène uniquement dans cette lignée cellulaire. Figure adaptée de Scheer & Ortega, 1999. ....	256
<b>Figure 27 : Construction du vecteur cible.</b> Tol2 : Pied de transposition. Is : Isolateur. UAS : Upstream Activating Sequence. HA : Héماغlutinine. RFP : Red Fluorescent Protein. CFP : Cyan Fluorescent Protein. BC : Promoteur de la Beta-cristalline.....	257
<b>Figure 28 : Embryon de 2 jpf après injection du vecteur contenant le variant H3.3 humain.</b> L'expression de la CFP dans le cristallin permet une sélection rapide des embryons. La barre d'échelle représente 500µm	257
<b>Figure 29 : Expression transitoire du vecteur contenant H3F3A. (A).</b> Expression de GAL4 (vert) et du transgène H3F3A (rouge) 2 jours après l'injection. Les barres d'échelle représentent 500µm. <b>(B)</b> Expression de GAL4 (vert) 4 jours après l'injection. L'expression de H3F3A n'est plus visible. Les barres d'échelle représentent 200 µm. ....	258
<b>Figure 30 : Embryons Tg(GAL 4 ; UAS :GFP) 4 jours après injection du vecteur contenant le gène MYCN.</b> L'expression de GAL 4 (vert), de la CFP dans le cristallin (bleu) et de MYCN (rouge) sont toujours visibles 4 jours après l'injection. LA barre d'échelle représente 200 µm. ....	259

# LISTE DES TABLES

---

<b>Table 1 : La nomenclature Brno. La nomenclature indique en premier l'histone concernée par la modification.</b> Le résidu modifié d'acide aminé est référencé par le code à 1 lettre suivi de sa position dans la séquence protéique. Les abréviations des modifications sont ajoutées en minuscule pour les distinguer des acides aminés. Adapté de Turner, 2005. ....	<b>24</b>
<b>Table 2 : Recensement des lignées mutantes et les phénotypes associés chez le modèle murin et les phénotypes observés chez l'homme. ....</b>	<b>56</b>
<b>Table 3 : Classification phylogénétique de <i>Danio rerio</i>.</b> Cette classification s'appuie sur la mise à jour de la classification des poissons osseux de 2017 (Betancur <i>et al.</i> , 2017).....	<b>71</b>
<b>Table 4 : Orthologues des gènes de la famille Polycomb humain chez le poisson zèbre .....</b>	<b>88</b>



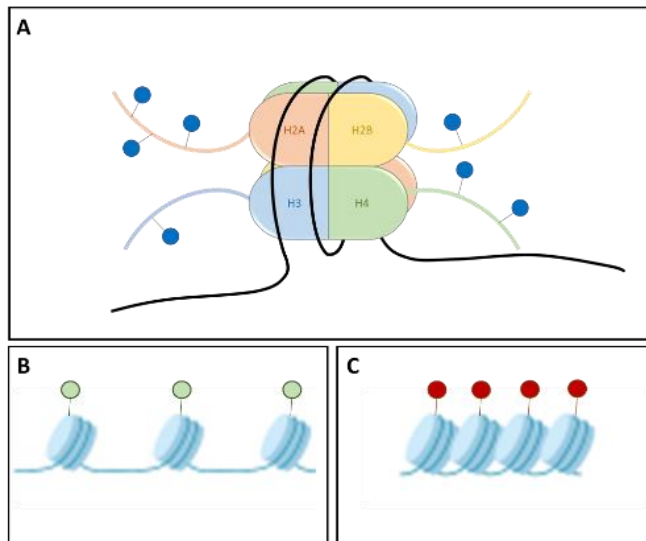
# INTRODUCTION

---

L'information génétique est commune à chaque cellule d'un même organisme et est transmissible à la descendance. Toutefois, si le patrimoine génétique contient effectivement une majorité des informations nécessaires à la vie cellulaire, lui seul ne peut expliquer comment un zygote peut se diversifier en centaines de types cellulaires aux phénotypes totalement différents, ni comment deux patients atteints de la même maladie monogénique répondront de manière différente à un traitement. La complexification de la notion initiale « un génotype = un phénotype » est en partie expliquée par l'épigénétique (Strohman, 1995).

Introduit la première fois en 1940 par Conrad Waddington, le terme épigénétique se définit aujourd'hui comme l'étude des changements transmissibles de la fonction d'un gène sans variation de la séquence génomique (Wu & Morris, 2001). Ces changements regroupent plusieurs mécanismes distincts. Plus de 20 000 ARN non codants (ARNnc) ont été identifiés (<https://www.genecodegenes.org/human/stats.html>). Ils possèdent de nombreuses fonctions biologiques, notamment l'application d'une régulation post-traductionnelle des ARNs messagers (ARNm). Les ARNm qui sont reconnus parfaitement par un ARNnc seront soumis à dégradation par le complexe RISC. Si la complémentarité entre les ARNnc et les ARNm est partielle, ces ARNm ne sont pas traduits. Les ARNnc peuvent se lier avec des protéines et favoriser par exemple le recrutement de méthylases de l'ADN (Zhang *et al.*, 2019). La méthylation de l'ADN, un autre mécanisme de régulation épigénétique, intervient presque exclusivement sur les cytosines suivies d'une base guanine chez les mammifères. Ces modifications des résidus cytosine sont réalisées par une famille de méthylases spécifiques appelées DNMT, qui catalysent le transfert d'un groupement méthyl à partir d'un donneur, le S-adenosylméthionine (SAM) (Moore *et al.*, 2013). Enfin, l'un des grands mécanismes de régulation épigénétique est la modification post-traductionnelle des histones et leurs variants. Au sein du noyau, l'ADN est enroulé autour d'une structure particulière appelée nucléosome. Ce nucléosome est un octamère constitué de deux fois chaque histone : H2A, H2B, H3 et H4 (Figure 1 A). On trouve dans le génome humain plusieurs variants de chaque histone, avec une spécificité et une fonction propre à chacun. Le variant H2A.X par exemple est intégré pour faciliter la réparation de l'ADN. L'intégration d'un variant d'histone peut modifier l'expression génique et favoriser l'expression ou la répression de programmes géniques développementaux (Santoro & Dulac, 2015). Les nucléosomes permettent un enroulement de 146pb d'ADN (McGhee & Felsenfeld, 1980). L'ensemble de l'ADN, des nucléosomes et des protéines associées forment la chromatine et permet l'organisation structurale de l'information génétique au sein du noyau. Mais plus que structurel, la chromatine peut avoir un impact sur la transcription des gènes. En effet, la chromatine peut être dans deux états : un état lâche et ouvert,

facilitant l'accès aux facteurs de transcription et la machinerie transcriptionnelle, l'euchromatine (Figure 1 B) et un état fermé et compact, l'hétérochromatine (Figure 1 C).



**Figure 1 : Formation du nucléosome et organisation de la chromatine.** A. Octamère d'histone H2A, H2B, H3 et H4. Les modifications post-traductionnelles des histones interviennent sur les parties c-terminales des séquences protéiques. B Représentation schématique de l'euchromatine. C. Représentation schématique de l'hétérochromatine. La conformation spatiale de la chromatine restreint l'accès à la machinerie transcriptionnelle. Le passage d'un état à l'autre est possible par le biais de protéines qui vont être capable de se lier aux marques épigénétiques, de les supprimer et d'en appliquer de nouvelles.

La détermination de l'état de la chromatine est dépendante des modifications post-traductionnelles des histones sur leur partie N-terminal, effectuées par des protéines spécialisées : les protéines dites « writers » capables d'appliquer les marques épigénétiques, les protéines « readers » qui lisent ces marques et les protéines « erasers » qui seront capables de supprimer ces marques. Ces modifications peuvent être très variées et regroupent principalement des acétylations, des méthylations, des phosphorylations et des ubiquitinations sur des résidus arginine et lysine (Zhao & Garcia, 2015). Les modifications post-traductionnelles des histones seront écrites suivant la nomenclature Brno (Turner 2005) et résumée dans la Table 1.

**Table 1 : La nomenclature Brno. La nomenclature indique en premier l'histone concernée par la modification.** Le résidu modifié d'acide aminé est référencé par le code à 1 lettre suivi de sa position dans la séquence protéique. Les abréviations des modifications sont ajoutées en minuscule pour les distinguer des acides aminés. Adapté de Turner, 2005.

Groupement ajouté	Résidu d'acide aminé modifié	Niveau de modification	Abréviation de la modification	Exemple
Acétyl-	Lysine	mono-	ac	H3K27ac
Méthyl-	Lysine	mono-	me1	H3K27me1
	Lysine	di-	me2	H3K27me2
	Lysine	tri-	me3	H3K27me3
Ubiquityl-	Lysine	mono-	ub1	H2AK119ub1
Phosphoryl-	Sérine	mono-	ph	H3S10ph

Les modifications post-traductionnelles des histones sont effectuées par des enzymes spécialisées. Ces enzymes sont parfois regroupées au sein de complexes. La répression Polycomb regroupe deux

complexes protéiques, les complexes PRC1 (Polycomb repressive complex 1) et PRC2 (Polycomb repressive complex 2). Ces deux complexes modifient les histones. Le complexe PRC1 ajoute un groupement ubiquityl sur la lysine 119 de l'histone H2A et le complexe PRC2 ajoute un groupement méthyl sur la lysine 27 de l'histone H3.

Ces modifications conduisent à la compaction de la chromatine et ainsi la répression de nombreux gènes. Ces deux complexes ont donc un rôle clé dans des processus physiologiques variés.



## CHAPITRE I : LA REPRESSION POLYCOMB DE LA DECOUVERTE AUX QUESTIONS DE RECHERCHE

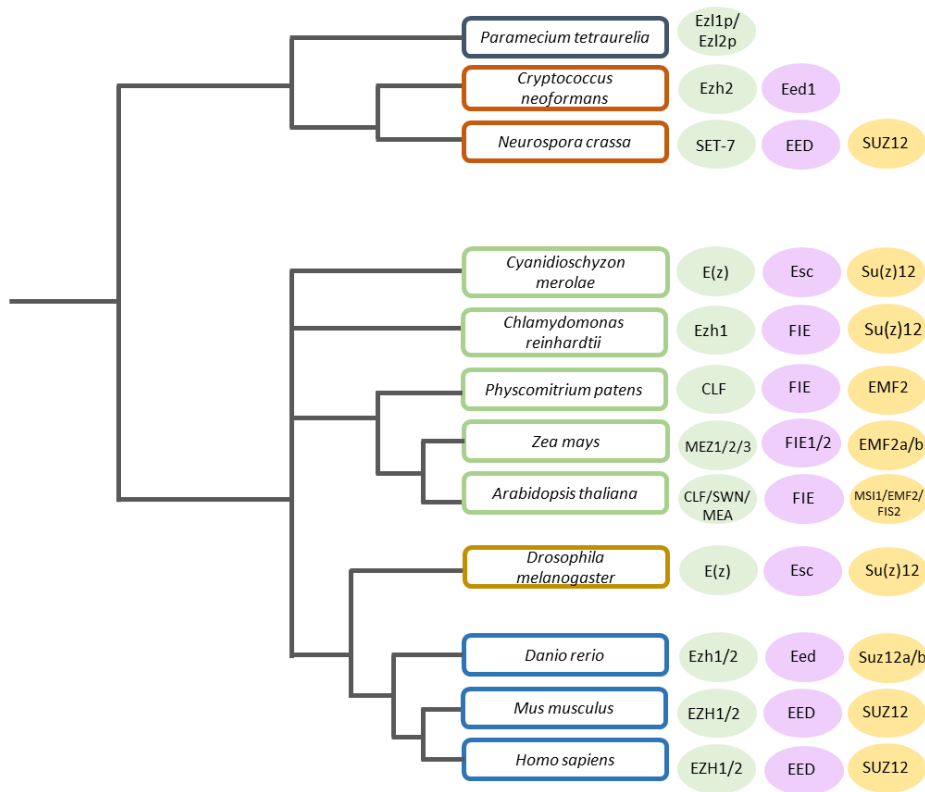
---

Le premier gène *Polycomb* (*Pc*) a été découvert chez la drosophile (*Drosophila melanogaster*) dès la fin des années 40 (Lewis, 1947). La perte de fonction de *Pc* se traduit par une modification des segments thoraciques et abdominaux au stade larvaire en réponse à une expression ectopique des gènes *Hox*, suggérant un rôle de régulation spatiale de *Pc* sur les gènes *Hox* au cours du développement (Lewis, 1978). Des études supplémentaires sur des lignées mutantes de drosophiles ont permis l'identification de mutations avec des phénotypes très similaires à celui de *Pc*, telles que les mutations du gènes *Esc* ou *Ph* (Struhl, 1981 ; Dura *et al.*, 1985), menant à la dénomination de groupe de gènes Polycomb (PcG) (Jürgens, 1985). Les gènes *E(z)* et *Su(z)12* ont été isolés initialement au cours d'investigations sur les interactions des gènes *white* et *zeste* ( $z^1$ ) pour la détermination de la couleur des yeux des drosophiles. Les lignées mutantes pour *E(z)* ou *Su(z)12* ont un effet antagoniste sur l'expression du gène  $z^1$  mais des phénotypes supplémentaires sont observés. La disposition du dimorphisme sexuel est altérée avec la présence de peignes sexuels supplémentaires (Wu *et al.*, 1989). Bien que la concordance des phénotypes observés en cas de mutations et les expériences de colocalisation suggéraient une action régulatrice de concert des protéines PcG, ce n'est qu'en 1992 qu'une interaction a été démontrée par immunoprécipitation de PC et PH (Franke *et al.*, 1992). De nombreuses interactions ont été par la suite mises en évidence (Jones *et al.*, 1998), jusqu'à l'identification du Polycomb Repressive Complex 1 (PRC1) (Shao *et al.*, 1999), puis du Polycomb Repressive Complex 2 (PRC2) (Ng *et al.*, 2000). Ces complexes permettent de modifier et compacter la chromatine, et ainsi répriment la transcription de programmes géniques particuliers.

En parallèle des découvertes effectuées sur la drosophile, plusieurs équipes ont identifié les homologues des protéines PcG chez la souris (Pearce *et al.*, 1992 ; Nomura *et al.*, 1994 ; Hobert *et al.*, 1996a) et chez l'homme (Alkerma *et al.*, 1993 ; Abel *et al.*, 1996). La conservation des domaines protéiques SET et CXC est importante entre les homologues protéiques des différentes espèces (Hobert *et al.*, 1996b). Cette conservation des séquences entre des espèces aussi éloignés que la drosophile et la souris témoigne de l'importance des complexes de répression Polycomb dans la répression génique au cours de l'évolution. La répression Polycomb est apparue très tôt chez les eucaryotes, puisque des organismes unicellulaires comme les paramécies possèdent les protéines EZL1 et EZL2, avec la capacité de réaliser les mêmes modifications épigénétiques. Des travaux récents ont identifié les PcG dans le génome de nombreux groupes taxonomiques différents (Figure 2).

Au regard du nombre de protéines homologues et d'orthologues, il semble que la complexité des PcG aient suivie la complexification du plan d'organisation du corps des organismes au cours de l'évolution (Vijayanathan *et al.*, 2022).

Les protéines du complexe PRC1 sont également très conservées chez les espèces animales, mais sont absentes du règne des champignons (Schuettengruber *et al.*, 2017). Chez les plantes, la répression médiée par le complexe PRC1 est nécessaire au développement, mais la composition du complexe est différente. Il semble que les deux complexes PRC2 et PRC1 soient issus de deux processus évolutifs différents : l'évolution du complexe PRC2 aurait été façonnée par des duplications du génome successives, tandis que le complexe PRC1 serait le résultat d'une évolution convergente (Vijayanathan *et al.*, 2022).

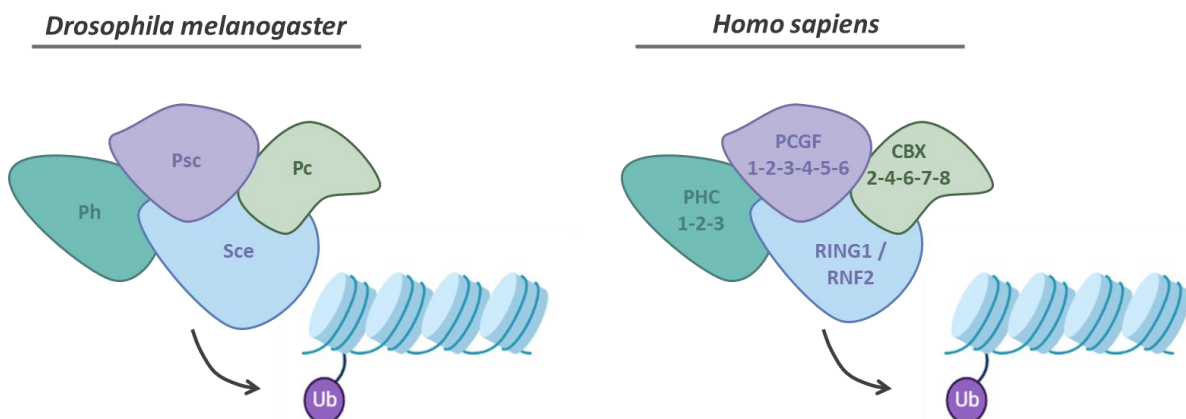


**Figure 2 : Diversité des protéines du complexe PRC2 au cours de l'évolution.** Les organismes issus du règne *Plantae* sont encadrés en vert et les espèces représentantes du règne *Fungi* sont encadrées en orange. Les organismes du règne *Animalia* sont encadrés en jaune pour les invertébrés et en bleu pour les vertébrés. Figure adaptée de Vijayanathan *et al.*, 2022.

## 1. Le complexe PRC1

Chez la drosophile, le complexe PRC1 est formé des protéines Pc, Ph, Psc (Posterior sex comb) et Sce (Sex combs extra). La sous-unité catalytique de ce complexe, Sce, permet le transfert d'un groupement ubiquitine sur le résidu de lysine 119 de l'histone H2A. Cette activité est dépendante de la présence des autres sous-unités du complexe. Le dépôt de la marque H2AK119ub1 compacte la chromatine et participe à la répression transcriptionnelle des gènes (Francis *et al.*, 2004 ; Wang *et al.*, 2004).

Chez les mammifères, le complexe PRC1 est composé des protéines CBX2/4/6/7/8, RING1/RNF2, PHC1-3 et BM1/PCGF1-6. L'organisation du complexe est similaire (Figure 3).



**Figure 3 : Organisation du complexe PRC1 chez la drosophile et l'homme.** Les sous-unités du complexe PRC1 de la drosophile sont conservées chez l'homme, avec plusieurs orthologues pour chacune des protéines. La fonction du complexe PRC1 est également conservée. La sous-unité catalytique du complexe permet le transfert d'un groupement ubiquitine sur la lysine 119 de l'histone H2A

Les différences majeures observées entre les mammifères et la drosophile se situent au niveau de la multiplication des orthologues pour les différentes protéines du complexe. Cette variabilité ainsi que la redondance des protéines permet l'assemblage d'une multitude de complexes et une adaptation à des environnements et stimuli diverses (Gil & O'Loghlen, 2014).

Les complexes PRC1 sont classés en deux catégories en fonction de leur mode de recrutement et leur composition. Les complexes PRC1 canoniques sont recrutés sur la chromatine par la marque H3K27me3, déposée par le complexe PRC2. Le recrutement des complexes PRC1 non-canoniques est dépendant de la composition de ces complexes (Morey *et al.*, 2012).

## 1.1 Le complexe PRC1 canonique

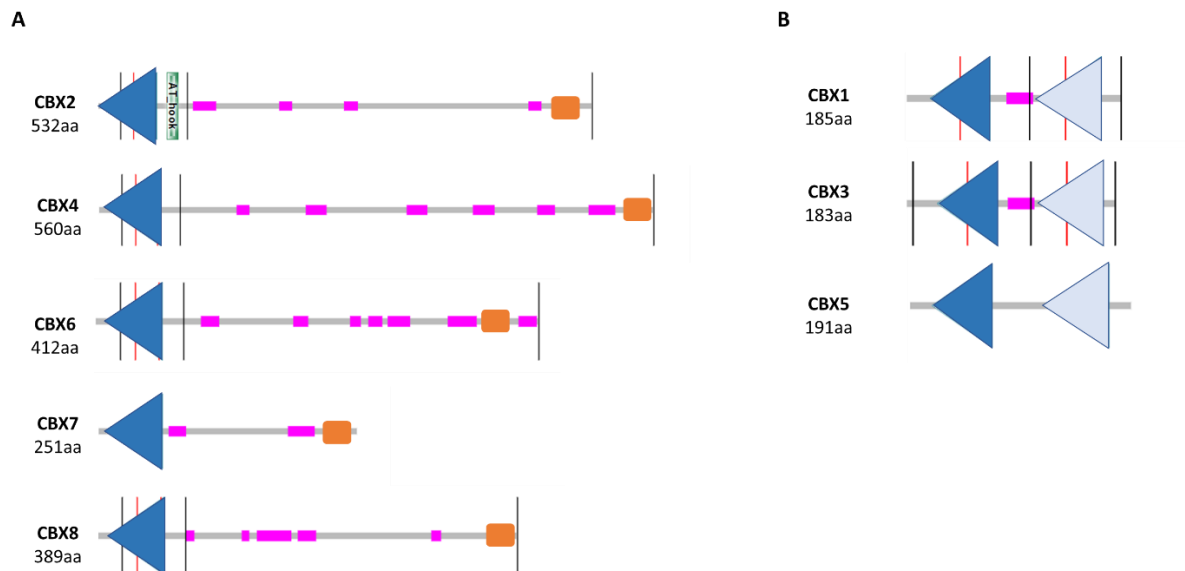
### 1.1.1 Composition des complexes canoniques

**RING1 et RNF2.** Les deux orthologues de la protéine Sce de la drosophile, RING1 (RING1A) et RNF2 (ou RING1B), portent l'activité catalytique du complexe PRC1. Ces protéines sont des E3-ubiquitin ligase. Le transfert d'un groupement ubiquitine sur le résidu lysine 119 de l'histone H2A restreint l'accès la polymérase et réprime la transcription (Stock *et al.*, 2007). Ces protéines possèdent un domaine RING finger en N-terminal qui permet une fixation à l'ADN (Lovering *et al.*, 1993). Un domaine WD40-associated ubiquitin-like (RAWFUL) est également conservé en position C-terminal (Connelly & Dykhuizen, 2017). Ces protéines ont un rôle fondamental dans le développement précoce et la perte de fonction de RNF2 provoque un arrêt précoce du développement au moment de la gastrulation (Voncken *et al.*, 2003).

**PHC.** Trois orthologues de Ph existent chez les mammifères, PHC1, PHC2 et PHC3. Ces protéines participent à l'assemblage du complexe. Elles possèdent un large domaine en position C-terminale appelé initialement SPM (Peterson *et al.*, 1997). Ce domaine est maintenant connu sous le nom de SAM (Sterile Alpha Motif) pour sa conformation dans l'espace. Ce domaine permet une oligomérisation des protéines et recrute d'autres complexes PRC1. L'assemblage des complexes forme des structures macromoléculaires, visibles en microscopie. Ce mécanisme de recrutement du complexe PRC1 applique une répression transcriptionnelle coordonnée sur une large portion génomique (Isono *et al.*, 2013).

**CBX.** Chez les mammifères, 5 orthologues de la protéine Pc de la drosophile sont conservés : CBX2, CBX4, CBX6, CBX7 et CBX8. Le domaine chromodomaine en N-terminal de chaque protéine a donné son nom à cette famille de protéines. Cette famille comporte trois autres protéines CBX1, CBX3 et CBX5 qui possèdent également ce chromodomaine (Figure 4). Ce domaine permet la reconnaissance spécifique de la triméthylation de la lysine 27 de l'histone H3 (Fischle *et al.*, 2003). Un autre domaine est conservé en C-terminale, le domaine C-box également appelé PcR-box (Senthilkumar *et al.*, 2009 ; Kim & Kingston 2020). Ce domaine permet l'intégration de la protéine CBX dans le complexe PRC1. Ce domaine est absent des protéines CBX1, CBX3 et CBX5, ce qui les privent d'interaction directe avec les complexes PRC1 canoniques. Ces protéines sont les orthologues vrais de la protéine de la drosophile Hp1 et ont une fonction différente des autres membres de la famille CBX. La présence d'un shadow chromodomain autorise des homodimérisations et des hétérodimérisations, et forment des agrégats protéiques (Figure 4). Elles participent néanmoins à la formation et à la fonction de certains complexes PRC1 non-canoniques. Les comparaisons des séquences des protéines CBX montrent un domaine

supplémentaire en comparaison avec leur orthologue Pc, le domaine AT-Hook pour CBX4, CBX6, CBX7 et CBX8, et le domaine AT-Hook Like pour CBX2. Ces motifs AT-Hook ne sont pas exclusifs aux protéines CBX. Ils permettent une interaction avec des séquences d'ADN.

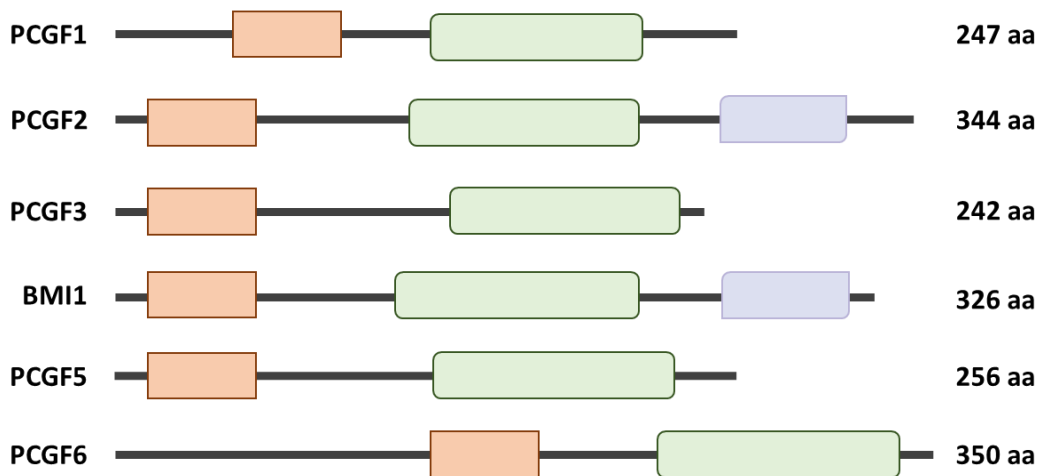


**Figure 4 : Les protéines de la famille CBX chez l'homme. A.** Les protéines de la famille CBX, orthologues de Pc, qui participent à la formation des complexes PRC1 canoniques. **B.** Les protéines orthologues de Hp1 forment des agrégats protéiques par homo- et hétérodimérisation. Les triangles bleus représentent les chromodomains, les triangles bleus clairs représentent les shadow chromodomains, les rectangles oranges représentent les domaines C-box. Les représentations graphiques ont été obtenues sur le site de modélisation [http://smart.embl-heidelberg.de/smart/show\\_motifs](http://smart.embl-heidelberg.de/smart/show_motifs).

**SCMH.** La protéine Scm a deux orthologues chez l'homme, SCM1 et SCM2. Ces protéines possèdent un domaine Sterile Alpha Motif (SAM). Ce domaine permet des interactions avec des protéines diverses, et notamment les sous-unités du complexe PRC2. Les protéines SCM1 partagent une homologie de séquence en C-terminale avec les protéines PH au niveau du domaine SAM. Ce domaine particulier permet la formation d'interactions homo- et hétérotypiques avec des protéines d'autres complexes PRC1. Cette particularité participe à la répression transcriptionnelle en formant de large domaine de recrutement de complexe Polycomb (Peterson *et al.*, 1997 ; Piunti & Shilatifard, 2021). La liaison des protéines SCM1 avec les autres interactants du complexe PRC1 est plus lâche. Leur présence n'est pas indispensable à l'activité du complexe. De ce fait elles sont classées comme protéines accessoires du complexe PRC1 canonique (Schuettengruber *et al.*, 2017).

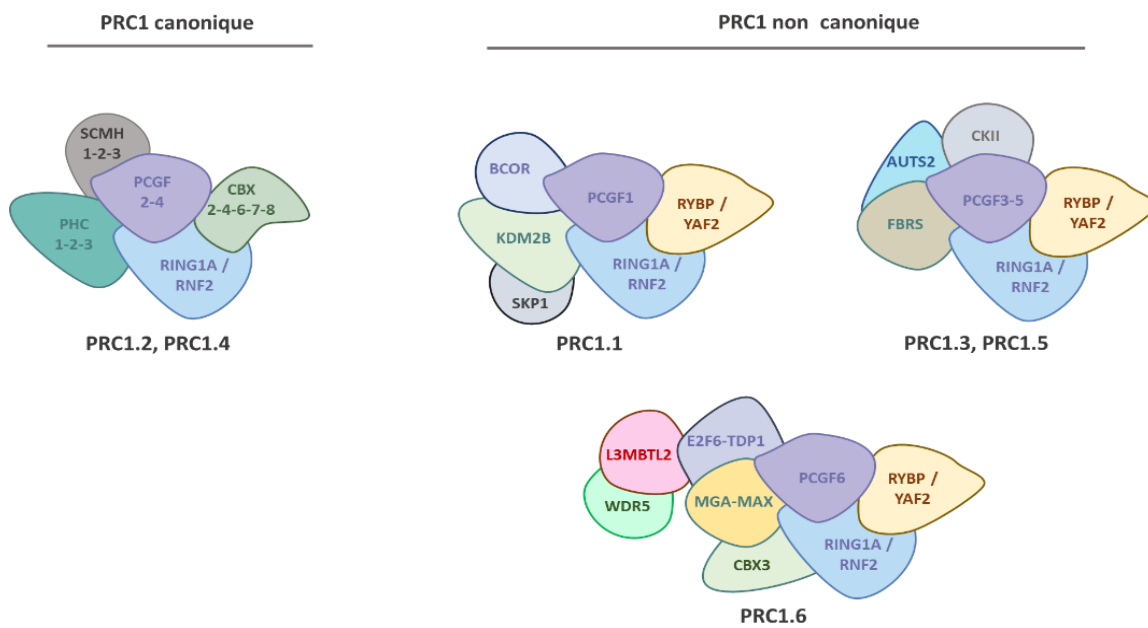
**PCGF.** Chez les mammifères, de nombreux orthologues existent pour la protéine Psc : PCGF1, PCGF2, PCGF3, BMI1, PCGF5, PCGF6. La famille de protéines PCGF possède deux domaines protéiques avec un fort taux de conservation : un domaine RING finger, avec une homologie importante des motifs RING de RING1A/B, et un domaine RAWFUL (Figure 5), dont l'homologie est moins importante mais reste

supérieure à 50%. Ce domaine RAWFUL fixe le domaine PUFD de BCOR, une autre protéine de la famille Polycomb (Junco *et al.*, 2013). Les protéines de la famille Psc sont exprimées dans des tissus variés, mais leur expression est dépendante du type cellulaire et du contexte biologique (Sánchez-Beato *et al.*, 2016). De façon intéressante, il semble aussi que la fonction des protéines PCGF se soit spécialisée au cours de l'évolution. PCGF1 est indispensable pour la différenciation des cellules souches embryonnaires (ESC) (Yan *et al.*, 2017). En revanche, la perte fonctionnelle de PCGF1 chez le poisson zèbre n'est pas dommageable pour le développement précoce des larves, mais induit un vieillissement prématuré des poissons adultes (Dupret *et al.*, 2016).



**Figure 5 : Représentation graphique des domaines conservés des protéines PCGF.** Les rectangles oranges représentent les domaines RING finger, les rectangles verts représentent les domaines RAWFUL et les bleus les domaines PS. Figure adaptée de Connelly & Dykhuizen, 2017.

Dans un premier temps, une dichotomie a été instaurée dans la classification des complexes PRC1. Les complexes canoniques possèdent un représentant de la famille CBX et sont recrutés par la marque H3K27me3 tandis que les complexes non-canoniques ne possèdent pas de protéines CBX dans leur assemblage. Plus tardivement, 6 complexes PRC1 distincts ont été mis en évidence en fonction de l'orthologue PCGF présents dans le complexe (Figure 6) (Gao *et al.*, 2012).



**Figure 6 : Les complexes PRC1 canoniques et non canoniques.** L'inclusion d'une protéine de la famille PCGF dirige la formation du complexe. Les interactions entre le complexe PRC1 et les protéines accessoires sont majoritairement portées par les PCGF. Adaptée de Vidal et Starowicz, 2017.

Les complexes canoniques sont constitués des variants PRC1.2 et PRC1.4. Seules les protéines PCGF2 et PCGF4 composent les PRC1c. Contrairement aux autres PCGF, ces deux protéines possèdent un domaine supplémentaire en N-terminal, un domaine riche en sérine PS (Geng & Gao, 2020). Ce domaine faciliterait les interactions avec les autres protéines du complexe.

### 1.1.2 PRC1.2 et PRC1.4

La composition des complexes contenant PCGF2 et PCGF4 sont identiques (Gao *et al.*, 2012). La spécificité de ciblage et d'action de ces complexes dépend de l'assemblage des autres unités du complexe. Les CBX par exemple peuvent être recrutées très spécifiquement sur des loci génomiques et recrutent chaque unité du complexe pour conduire à la répression Polycomb (Morey *et al.*, 2013). Bien que les complexes PRC1.2 et PRC1.4 puissent être identiques dans leur composition, leur fonction n'est pas entièrement redondante. En effet, PCGF4 (communément appelé BMI1) a un rôle prépondérant dans le maintien des cellules souches. L'expression de *BMI1* et *PCGF2* dans les cellules souches hématopoïétiques révèlent une expression concomitante, mais dans des proportions inégales. Lorsque *BMI1* est fortement exprimé, l'expression du gène *PCGF2* est réduite, et réciproquement. La perte de fonction de chacun de ces gènes mène à l'apparition de phénotypes adverses. La perte de *BMI1* inhibe l'auto-renouvellement des cellules souches tandis que la perte de *PCGF2* inhibe la différenciation (Kajiume *et al.*, 2009). Les complexes PRC1.2 et PRC1.4 semblent donc posséder des fonctions spécifiques, non compensables par un autre orthologue.

## 1.2 Les complexes PRC1 non-canoniques

Les complexes PRC1c possèdent tous les protéines CBX. Le premier modèle de recrutement mis en évidence du PRC1 était le recrutement dépendant du complexe PRC2. Depuis, de nombreuses études pointent le rôle de complexes PRC1 variés et un recrutement indépendant de la marque H3K27me3, car dépourvus de sous-unités CBX. D'autres acteurs protéiques sont trouvés dans les complexes PRC1. L'augmentation du nombre d'interactants accroît le nombre de combinaison de protéines pour former le complexe PRC1. Certaines études hypothétisent que la majorité de l'activité de l'activité du complexe PRC1, l'ubiquitination de la lysine 119 de l'histone H2A, est effectuée par des complexes non-canoniques (Conway & Bracken, 2017).

Le rôle moléculaire de ces complexes PRC1 non-canoniques (PRC1nc) est identique à celui des PRC1c : l'ubiquitination de la lysine 119 de l'histone H2A. Cette activité enzymatique est d'ailleurs portée par les mêmes protéines, RING1A et RNF2. Les orthologues de Psc sont également retrouvés dans les complexes PRC (Figure 6).

**RYBP.** La protéine RYBP (YY1 binding protein) ou son homologue YAF2 (YY1-associated factor 2) interagissent avec RING1 et le facteur de transcription YY1 via un domaine en C-terminal (García *et al.*, 1999). Ces deux protéines possèdent un domaine en doigt de zinc avec un domaine particulier NZF qui possède une activité de liaison à l'ubiquitine. Ce domaine permet donc la liaison des protéine RYBP et YAF2 à la marque H2AK119ub (Arrigoni *et al.*, 2006). Dans les cellules souches embryonnaires (ESC) cette protéine permet un recrutement du PRC1nc a des loci spécifiques. La comparaison entre la fixation de CBX7 et celle de RYBP révèle des zones de chevauchements des complexes PRC1c et PRC1nc, mais également des zones d'exclusivités. La présence de RYBP permet également une modulation de l'activité du complexe PRC1nc (Morey *et al.*, 2013). La fixation de RYBP sur la marque H2AK119ub facilite le recrutement du complexe PRC1, son activité et ainsi la diffusion de cette marque répressive. La délétion de cette protéine chez le modèle murin a non seulement affecté la déposition de la marque H2AK119ub, mais aussi celle de la marque H3K27me3. Ces résultats mettent en lumière le rôle primordial de RYBP pour l'activité des complexes PRC1nc, mais également pour la connexion entre les complexes Polycomb dans l'établissement de la répression transcriptionnelle (Rose *et al.*, 2016). RYBP ou YAF2 sont présents dans tous les complexes PRC1nc (Figure 6).

### 1.2.1 PRC1.1

**PCGF1.** La protéine PCGF1 est essentielle dans l'assemblage de ce complexe variant PRC1. Elle permet le recrutement des protéines accessoires et de la sous-unité catalytique à la chromatine. Bien que des



expériences de KO de PCGF1 se soit révélées sans effet sur le niveau d'ubiquitination en raison d'une compensation par les autres orthologues, l'incorporation de cet orthologue permet de cibler des gènes spécifiques (Scelfo *et al.*, 2019).

**KDM2B.** Cette protéine possède une activité enzymatique différente de l'activité du PRC1. Le domaine JmjC (Jumonji C) de KDM2B possède une activité déméthylase spécifique de la modification du résidu lysine 36 de l'histone H3. La marque H3K36me3 est une marque activatrice de la transcription. KDM2B possède un motif riche en résidu leucines appelé F-box. Ce motif assure la liaison à PCGF1 et RING1B (Wu *et al.* 2013), ainsi que BCOR (Wong *et al.*, 2016). Ces interactions semblent séquentielles : un premier assemblage avec SKP1 permet ensuite à KDM2B d'intégrer un complexe PCGF1-BCOR. La protéine KDM2B possède également un domaine à doigt de zinc CxxC, qui reconnaît et fixe les îlots CpGs. Ce domaine autorise le recrutement du complexe PRC1.1 au niveau des îlots riches en CpGs afin de médier la répression Polycomb (Farcas *et al.*, 2012).

**BCOR.** La protéine BCOR (BCL6 co-repressor) possède un orthologue chez les mammifères, BCORL1. Ces deux protéines conservent des domaines avec un fort taux d'homologie mais ont des profils d'expression et des fonctions différentes. Les deux protéines possèdent un domaine en N-terminal PUFD qui permet l'interaction avec le domaine RAWUL de PCGF1 et ainsi la formation du complexe PRC1.1. La stabilité des interactions est assurée par un domaine ANK (Ankyrine), qui consiste en une répétition d'acides aminés qui forment deux hélices alpha séparées par une boucle (Wong *et al.*, 2020). La formation du complexe PRC1nc contenant BCOR est nécessaire au maintien de l'état souche des ESC (Wang *et al.*, 2018).

### 1.2.2 PRC1.3 et PRC1.5

Les protéines PCGF3 et PCGF5 interagissent toutes les deux avec les mêmes protéines : AUTS2 (Autism Susceptibility Candidate 2), FBRS et CKII (Caséine Kinase 2). La façon dont ses variants PRC1 sont recrutés à la chromatine n'a pas été entièrement élucidée. La protéine AUTS2 permet la liaison de PCGF3 et PCGF5 à CKII. La kinase CKII phosphoryle RING1B sur le résidu de sérine 168 et cette modification post-traductionnelle inhibe l'activité du complexe PRC1.3 et PRC1.5 (Gao *et al.*, 2014). De plus, CKII semble permettre le recrutement d'autres facteurs de transcription telle que P300. Cet ensemble de protéines crée un environnement chromatinien permissif à la transcription et est nécessaire à la différenciation des cellules neurales (Geng & Gao, 2020). Ces résultats nuancent le rôle fondamental des complexes PRC1, la répression transcriptionnelle médiée par la déposition de la marque H2AK119ub1. En revanche, des expériences de KO des deux protéines PCGF3 et PCGF5 ont montré une baisse globale de cette marque épigénétique accompagnée d'une létalité très précoce des

embryons murins. Ces deux protéines se lient au long ARNnc XIST. Cette fixation entraîne le recrutement des complexes PRC1nc à la chromatine et ainsi la mise sous silence du chromosome X chez les embryons femelles (Almeida *et al.*, 2017). Ainsi ces résultats révèlent la dualité des complexes PRC1.3 et PRC1.5 et l'adaptation de leurs activités en fonction des stimuli.

### 1.2.3 PRC1.6

Au sein de ce complexe PRC1.6, PCGF6 se lie aux protéines MGA (MAX gene-associated protein), E2F6 (E2F Transcription Factor 6) et L3MBTL2 (Histone Methyl-Lysine Binding Protein 2).

**MGA.** Cette protéine est un facteur de transcription qui agit sous forme d'hétérodimère avec MAX (MYC-associated factor X). Ces deux protéines contiennent deux domaines de liaisons à l'ADN, un domaine T-box et un domaine bHLH (basic helix-loop-helix). L'hétérodimère MGA-MAX est capable de se fixer sur des E-boxes, une séquence de quelques nucléotides, ce qui permet le recrutement du complexe PRC1.6 principalement en amont du promoteur des gènes (Stielow *et al.*, 2018).

**L3MBTL2.** Cette protéine possède une capacité de fixation aux résidus de lysine méthylés des histones H3 et H4 par une répétition de domaines MBT (Yoo *et al.*, 2010). Cette activité permet le recrutement du complexe PRC1.6 et de renforcer un environnement répressif par l'apposition de la marque H2AK119ub1. Une modification post-traductionnelle de L3MBTL2 par SUMOylation affecte son activité et permet un ciblage de gènes spécifiques. La protéine mutante L3MBTL2<sup>K675/700R</sup> inhibe la sumoylation de la protéine ce qui résulte en la dérégulation de gènes pro-inflammatoires (Stielow *et al.*, 2014).

**E2F6.** Contrairement aux autres membres de la famille de protéine E2F, E2F6 ne possède ni de domaine de fixation aux protéines RB, ni de domaine de transactivation. E2F6 possède un domaine de liaison à l'ADN ainsi qu'un domaine d'hétérodimérisation aux protéines DP1/DP2 (Transcription factor Dp-1), à l'instar de la famille E2F (Gaubatz *et al.*, 1998). Le domaine de liaison à l'ADN autorise la fixation de E2F6 sur des éléments de réponses aux protéines E2F. L'absence du domaine de transactivation et l'association au complexe PRC1.6 permet une répression des gènes cibles, et notamment certains gènes impliqués dans le cycle cellulaire (Stielow *et al.*, 2018).

## 1.3 Recrutement des complexes PRC1

Le premier mécanisme de recrutement des protéines PcG a été identifié chez la drosophile, sur les gènes *Hox*. Les protéines Polycomb se fixent sur des séquences d'ADN en amont de la séquence codante des gènes *Hox*. Ces motifs ont été inclus au sein d'éléments transposables et ont créé de nouveaux sites de fixation pour les protéines PcG (Chan *et al.*, 1994). Ces séquences particulières qui

permettent le recrutement des complexes PRC1 et PRC2 sont appelés Polycomb Responsive Element (PRE). Parmi les différentes espèces de *Drosophile*, les motifs protéiques de reconnaissance des PREs sont bien conservés, tandis que la position des PREs dans le génome et leur séquence ne le sont pas. Ces résultats montrent que les PREs pourraient être des séquences qui évoluent rapidement, rendant leur identification chez les vertébrés compliquée (Hauenschild *et al.*, 2008 ; Bauer *et al.*, 2016). Le nombre important de protéines paralogues pour chacune des sous-unités permet de créer des complexes uniques, qui sont distribués au travers du génome, avec un mode de recrutement et une action propre à chaque complexe. La complexification des complexes de répression Polycomb suggère que si les PREs sont conservés chez les vertébrés, la variabilité de leur séquence s'est accrue également. Chez les mammifères, de nombreuses équipes ont essayé d'identifier des séquences consensus fixant les PcG qui pourraient être des PREs et leurs résultats sont encore source de débats (Bauer *et al.*, 2016). Des parallèles ont toutefois été dressés entre les PREs de la drosophile et les potentielles séquences identifier chez les mammifères : ce sont des séquences d'ADN non méthylées, situés aux promoteurs et au niveau de régions intergéniques, ayant la capacité de recruter les complexes PcG.

### 1.3.1 Recrutement par le chromodomain des CBX

Un premier modèle hiérarchique de recrutement du complexe PRC1 a été montré en 2004. Les protéines Pho ou Phol se fixent sur les PREs. Cette fixation entraîne le recrutement du complexe PRC2 qui triméthyle la lysine 27 de l'histone H3. La marque H3K27me3 est reconnue par le chromodomain de Pc (Figure 7 A) et active le complexe PRC1 canonique (Wang *et al.*, 2004). Parmi les orthologues de Pc chez les mammifères, toutes les protéines CBX ont un chromodomain avec un haut degré de conservation. Chaque protéine est intégrée de manière unique dans le complexe PRC1 (Vandamme *et al.*, 2011). La perte de fonction d'une protéine de la famille CBX ne conduit pas à un arrêt précoce du développement et des transformations homéotiques, ce qui suggèrent une compensation par les autres protéines de la famille. Dans les cellules ES murines, CBX2 et CBX4 partagent la fonction d'orientation de la différenciation des cellules. En revanche, CBX7 tend à inhiber l'expression de gènes de différenciation (Morey *et al.*, 2012). Au locus du gène suppresseur de tumeur INK4a, CBX7 et CBX8 agissent sur les mêmes régions génomiques mais un knock-down réalisé contre CBX7 inhibe la protéine CBX8 et inversement. L'action des protéines CBX est donc redondante pour certaines cibles, mais les différents complexes formés par les paralogues sont également interdépendants (Maertens *et al.*, 2009). À l'inverse, au sein des cellules souches hématopoïétiques, CBX7 et CBX8 partagent les mêmes cibles génomiques, mais sont en compétition. L'incorporation de l'un des paralogues oriente le devenir

des cellules souches après la division. La présence de CBX8 dans les complexes PRC1 permet la répression de gènes impliqués dans le maintien de l'état souche. La protéine CBX8 est principalement exprimée dans les progéniteurs, tandis que CBX7 est exprimée dans les cellules souches, où elle maintient réprimé les gènes de différenciation et préserve les capacités d'auto-renouvellement (Klauke *et al.*, 2013).

Les complexes PRC1c sont donc recrutés via la reconnaissance de la marque H3K27me3 par les protéines CBX. Néanmoins des analyses d'immunoprécipitation de la chromatine (ChIP) ont permis d'identifier des régions génomiques dénuées de H3K27me3 qui fixent les protéines CBX (Schaaf *et al.*, 2013). Ces résultats pourraient être en partie expliqués par l'aspécificité des protéines CBX. En effet, les chromodomaines présentent un degré d'homologie important entre les paralogues et avec l'orthologue Pc chez la drosophile. Pourtant, certaines protéines CBX sont capables de reconnaître la triméthylation de la lysine 27 et de la lysine 9 (Bernstein *et al.*, 2006). Les séquences de résidus d'acides aminés qui encadrent les lysines 9 et 27 sont très similaires. La protéine CBX4 semble même avoir une affinité plus grande avec la marque H3K9me3. Mais les deux marques épigénétiques H3K9me3 et H3K27me3 sont des marques de l'hétérochromatine et sont retrouvées globalement sur les mêmes régions du génome. L'affinité des chromodomaines pour cette seconde marque ne suffit donc pas à expliquer la présence des orthologues de Pc sur des régions dépourvues de H3K27me3.

### 1.3.2 Recrutement par des ARNnc

L'équipe de Bernstein a montré que les chromodomaines des protéines CBX pouvaient également fixer des ARNs de manière aspécifique (Bernstein *et al.*, 2006). L'activité de CBX7 est même fortement réduite après traitement à la RNase. Le locus *INK4a-ARF* est une cible de régulation bien documentée du complexe PRC1. Ce locus produit le long ARNnc ANRIL. La conformation de ce long ARNnc permet une fixation au chromodomaine de CBX7 (Figure 7 B). Cette interaction CBX7/ANRIL est nécessaire à la fixation du PRC1 sur la marque H3K27me3, et ainsi à la répression du locus *INK4* (YAP *et al.*, 2010). L'action du complexe PRC1 dans ce contexte permet un contrôle de la sénescence cellulaire. De la même façon, CBX4 est impliqué dans le contrôle de gènes de croissance cellulaire par le biais de son interaction avec TUG1 et MALAT1 (Yang *et al.*, 2011).

De la même façon, d'autres ARNnc sont capables de recruter les PRC1 variants. Très tôt après la fécondation, un des deux chromosomes X doit être inactivé. Cette inactivation est en partie médiée par la répression Polycomb. Les deux complexes PRC1 et PRC2 sont nécessaires à l'inactivation du chromosome X. En revanche, dans le cas d'une perte d'activité de PRC2, PRC1 est toujours en capacité d'initier et d'appliquer la répression Polycomb (Schoeftner *et al.*, 2006). L'ARN XIST est composé de

plusieurs modules de séquences répétés. Les modules B et C sont indispensables à l'interaction avec PCGF3 et PCGF5 (Bousard *et al.*, 2019). Les protéines PCGF3 et PCGF5 stimulent alors l'activité de PRC1.3 et PRC1.5 respectivement.

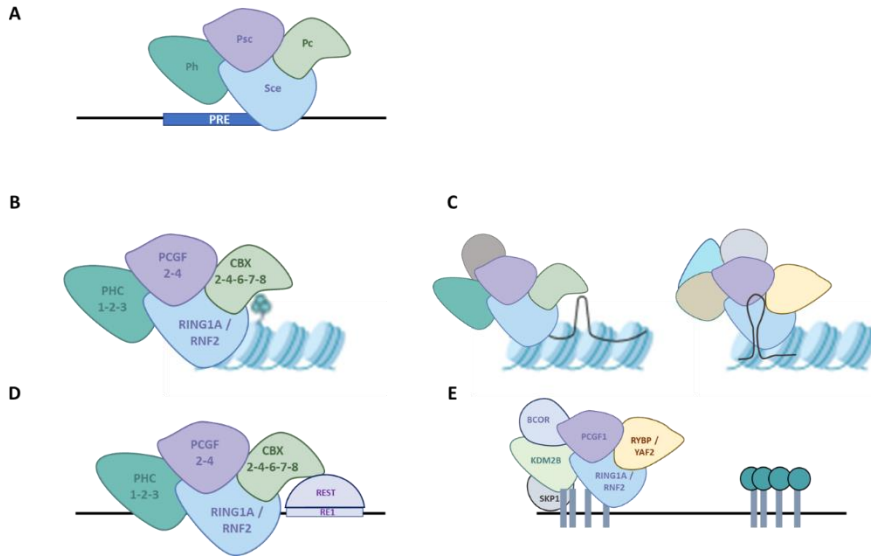
### 1.3.3 Recrutement par les facteurs de transcription

Les facteurs des transcriptions sont très spécifiques et se fixent sur des éléments de réponse présents dans les séquences nucléotidiques, ou agissent par interaction protéiques. Certains facteurs de transcriptions possèdent des domaines qui leur permettent de se lier aux protéines Polycomb. Par exemple le facteur de transcription AF9 interagit avec CBX8 lors du développement de leucémies aigues. Cette interaction n'implique ni le chromodomain ni le domaine C-box. L'interaction avec AF9 s'effectue avec les résidus 202 à 333 de CBX8. Aucune séquence similaire n'est retrouvée dans les séquences des autres paralogues (Hemenway *et al.*, 2001). Malgré une apparente redondance de fonction, les interactions de facteurs de transcription avec un orthologue unique permettent le ciblage précis des complexes PRC1 au travers du génome et une adaptation au cours du développement et au gré de l'environnement. De nombreux autres facteurs de transcription ont la capacité de se lier aux protéines CBX et de modifier leur activité. La protéine REST se fixe sur les protéines CBX et modifie leur position sur certains loci (Ren & Kerppola, 2011). L'interaction de la protéine CBX8 avec ENL inhibe son activité de répression par le complexe PRC1 (Maethner *et al.*, 2013). Les facteurs de transcription tels que SNAIL, RUNX1, ou encore YY1 établissent des liaisons avec un nombre important d'unités de PRC1, et permettent d'induire une répression génique médiée par les complexes Polycomb au cours de du développement puis de la vie d'un organisme (Schuettengruber *et al.*, 2017).

### 1.3.4 Recrutement par les îlots CpG

La protéine KDM2B possède un motif en doigt de zinc qui lui assure une fonction de reconnaissance des îlots CpG. D'autres domaine de KDM2B fixent les unités du complexes PRC1, ce qui assure un recrutement de PRC1.1 sur des îlots CpG non méthylés (Farcas *et al.*, 2012). L'établissement de la marque H2AK119ub1 via ce phénomène a un rôle très important au cours du développement précoce. En effet, KDM2B se fixent sur des îlots CpG qui ne sont pas méthylés et ne sont pas transcrits activement. Dans les ESCs murines, un KD de KDM2B conduit à une surexpression de facteurs de transcription impliqués dans la mise en place des grands programmes de développement, tels que *GATA6*, *SOX7*, *OCT4* ou *GATA4*. Dans ce contexte, le recrutement par KDM2B au niveau des CpGs apparait prépondérant pour maintenir le statut des cellules souches (He *et al.*, 2013). La marque H2AK119ub1 prévient la méthylation de l'ADN sur ces loci, ce qui permet de pouvoir modifier

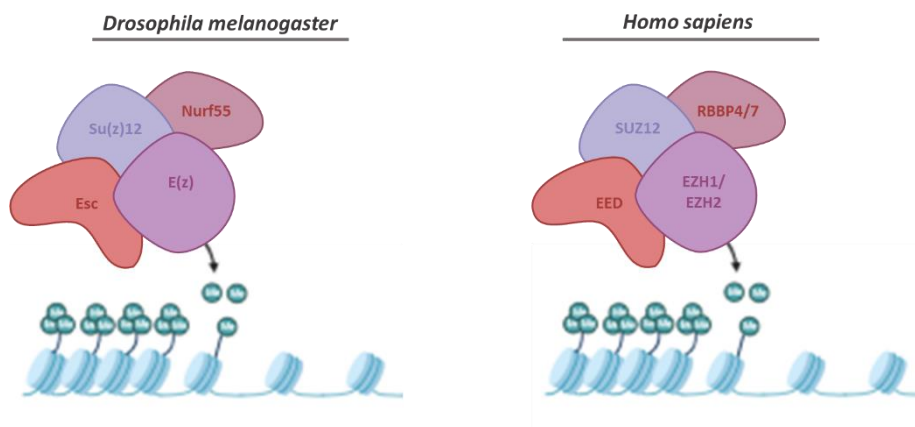
l'environnement chromatinien rapidement si la cellule reçoit des signaux de différenciation. L'application de la marque H2AK119ub1 sur les îlots CpG permet également le recrutement du complexe PRC2 et la déposition de la marque H3K27me3. Ce mode de recrutement bouscule le modèle hiérarchique de recrutement des complexes Polycomb et suggère que PRC1 et PRC2 peuvent initier la répression Polycomb *de novo* et recruter d'autres protéines readers et writers afin de renforcer la répression (Blackledge *et al.*, 2014).



**Figure 7 : Recrutement des complexes PRC1.** **A.** Chez la drosophile, le complexe PRC1 est recruté au niveau de des PRE **B.** Le modèle de recrutement hiérarchique nécessite la déposition de la marque H3K27me3 par le complexe PRC2. Cette marque est ensuite reconnue par les protéines CBX. **C.** Les complexes PRC1c et PRC1nc peuvent être recrutés par des ARNnc. **D.** Après fixation sur les séquences RE1, le facteur de transcription REST peut recruter le PRC1 via la fixation des protéines CBX. **E.** La protéine KDM2B reconnaît les îlots CpG et permet le recrutement de PRC1.1. En revanche la méthylation des îlots CpG (ici en rond vert) inhibe le recrutement de PR1.1. Adapté de Connelly & Dykhuizen, 2016 ; Gil & O’Loghlen, 2014.

## 2. Le complexe PRC2

Le complexe de répression Polycomb 2 (PRC2) a également été découvert chez la drosophile. Ce complexe est composé de 4 unités essentielles à son fonctionnement Esc (Extra sex comb), E(z) (Enhancer of zeste), Nurf55 et Su(z)12 (Suppressor of zeste 12) (Figure 8). Ce complexe est très conservé au cours de l'évolution et de nombreux orthologues sont également retrouvés chez les espèces animales au cours de l'évolution (Figure 1). Chez les mammifères, le complexe PRC2 est composé des acteurs suivants : EZH1/2 (Enhancer of Zeste Homolog 1/2), EED (Embryonic Ectoderm Development), SUZ12 et RBBP4/7 (RB Binding Protein 4/7) (Figure 8). L'activité de répression de ce complexe est médiée par la déposition de 3 groupements méthyle sur la lysine 27 de l'histone H3 (Figure 8) (Kuzmichev *et al.*, 2002). À l'instar de la marque H21K119ub1, cette modification post-traductionnelle de l'histone H3 compacte la chromatine et crée un environnement hostile à la transcription génique.



**Figure 8 : Composition du complexe PRC2 chez la drosophile et chez les mammifères.** La composition et l'organisation du complexe PRC2 de la drosophile sont conservées chez les mammifères. Les sous-unités catalytiques E(z) et EZH1/EZH2 ajoutent trois groupements méthyle sur la lysine 27 de l'histone H3.

### 2.1 Composition du complexe

#### 2.1.1 Les sous-unités catalytiques EZH1 et EZH2

Le gène *E(z)* a été identifié la première fois chez la drosophile, lors de l'étude des loci *zeste* et *white* qui dirige la couleur des yeux des individus (Kalisch & Rasmuson, 1974). La protéine E(z) est exprimée de façon ubiquitaire dans tous les tissus embryonnaires. Sa localisation intracellulaire est uniquement nucléaire (Carrington & Jones, 1996). Chez la souris, deux orthologues sont retrouvés, ENX-1 (EZH1) et ENX-2 (EZH2) (Hobert *et al.*, 1996). La protéine EZH2 est exprimée de façon ubiquitaire durant les stades très précoces du développement embryonnaire, ce qui révèle l'importance de l'activité de PRC2

dans la mise en place des premiers programmes de développement. Les protéines EZH1 et EZH2 possèdent de nombreux domaines conservés. Le domaine SET possède l'activité catalytique des protéines et présentent un fort taux d'homologie. Ce domaine fixe le S-Adenosylméthionine, le donneur du groupement méthyle qui est ensuite transféré sur la lysine 27 de l'histone H3. Deux domaines SANT permettent les liaisons avec l'histone H3.

**EZH1.** C'est une protéine de 747 acides aminés, avec une homologie de séquence à E(z) supérieure à 50% et de près de 80% au niveau du domaine catalytique SET (Abel *et al.*, 1996). Il catalyse l'ajout de groupements méthyl sur la lysine 27 de l'histone H3. Le gène *EZH1* est exprimé sous forme de deux transcrits majoritaires à l'âge adulte et au stade fœtale. Un phénomène d'épissage alternatif est à l'origine de la différence de taille des transcrits, et pourrait permettre la régulation de sa fonction. L'équipe de Bodega a montré par exemple la régulation du complexe PRC2 par une isoforme de EZH1 dépourvue de domaine catalytique. Cette isoforme demeure cytosolique et permet la rétention de EED dans le cytoplasme des cellules et inhibe l'assemblage du complexe PRC2. Sous l'effet d'un stress oxydatif, l'isoforme EZH1 $\beta$  est dégradée, et permet à EED d'intégrer le complexe PRC2. Ce phénomène favorise une adaptation rapide des cellules en fonction d'un stress environnemental. L'isoforme EZH1 $\beta$  a ici un rôle de senseur de l'environnement cellulaire (Bodega *et al.*, 2017).

Chez la souris, EZH1 est retrouvée sur les mêmes cibles génomiques que EZH2. Le KD de EZH1 ne semble pas diminuer le niveau de méthylation de lysine 27 de l'histone H3 au cours du développement. En revanche, EZH1 semble être plus exprimée dans les tissus et organes adultes. Le rôle de EZH1 semble donc être le maintien de la répression Polycomb dans des cellules non prolifératives, au sein de tissus adultes et même en absence de EZH2 (Margueron *et al.*, 2008). Dans les myocytes cardiaques, EZH1 et EZH2 sont nécessaires à la formation du cœur, en revanche seul EZH1 est nécessaire pour la régénération cardiaque. Les expériences de CHIP ont identifié les positions de EZH1 au travers du génome, et la protéine est retrouvée associée à des loci transcriptionnellement actifs. EZH1 aurait un rôle d'activateurs de la transcription, indépendamment de PRC (Ai *et al.*, 2018).

**EZH2.** La protéine *ezh2* est composée de 751 acides aminés. Contrairement aux autres histone-méthyltransférases porteuses du domaine SET, EZH2 (ainsi que EZH1) ne sont actifs qu'au sein du complexe PRC2, suggérant que la conformation après association du complexe permet l'activation du domaine SET. La cristallisation de EZH2 révèle que le domaine SET a la même configuration que les autres domaines SET canoniques, avec une partie centrale qui comprend le domaine actif, un I-SET qui lie la protéine au substrat et enfin un groupement d'acides aminés aromatiques qui forme un canal de liaison méthyl-lysine. Au sein de la protéine EZH2, le domaine SET est encadré en N-terminal par le



domaine CXC et un domaine post-SET en C-terminal. Lorsque la protéine EZH2 est dissociée de ses partenaires, le repliement des domaines CXC et post-SET camoufle le site de liaison au SAM. Seule la fixation des autres protéines Polycomb et des cofacteurs permet de modifier la conformation tridimensionnelle et activer EZH2 (Wu *et al.*, 2013). Les mutations qui modifient certains des acides aminés de la séquence de EZH2 peuvent ainsi se révéler désastreuses, inhibant ou activant totalement la protéine, quelques soient ses interactions. De nombreux autres domaines sont présents sur la séquence protéique de EZH2 (Figure 9) et nécessaires à sa fonction par les interactions qu'ils permettent. Le domaine SANT1L participe à la conformation de la protéine en se fixant sur le domaine SBD (SANT1L-binding domain). Le domaine EBD (EED-binding domain) fixe la protéine EED. Le repliement de EZH2 via l'interaction SANT1L-SBD forme une ceinture, formée des domaines EBD, BAM ( $\beta$ -addition motif), SAL (SET activation loop) et SRM (stimulation-responsive motif) autour de des domaines WD40 de EED.

Le domaine CXC, riche en cystéine, interagit avec les autres protéines du complexe et est directement lié au domaine SET (Jiao & Liu, 2015). L'ensemble de cette structure met en lumière l'importance de la conformation du complexe pour son activité. Ainsi, toutes mutations qui modifient même légèrement cette structure va avoir des conséquences dramatiques sur la fonction de EZH2.

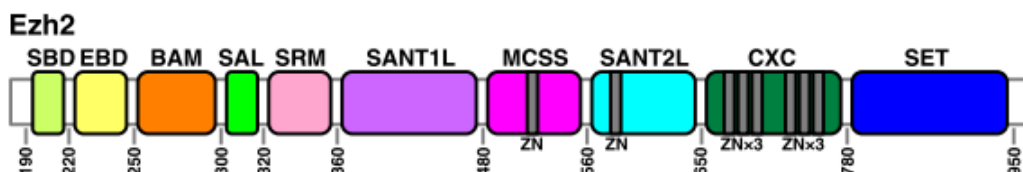


Figure 9 : Structure schématique de la protéine ezh2. La protéine ezh2 est composée de 951 acides aminés. L'activité enzymatique est portée par le domaine SET. Les autres domaines, et les motifs à Zinc qu'ils contiennent permettent les interactions entre les sous-unités du complexe ainsi que des interactions intramoléculaires. Figure de Jiao & Liu, 2015.

L'action majoritaire de EZH2 est la déposition de la marque épigénétique H3K27me3. Cependant EZH2, en association avec le PRC2, peut méthyliser des protéines différentes de l'histone H3. Elle peut méthyliser la lysine 116 de JARID2, qui se fixe alors à EED et stimule l'activité catalytique du complexe. Ici, la méthylation de EZH2 permet une boucle de régulation positive de PRC2 (Sanulli *et al.*, 2015). Le complexe PRC2 peut également méthyliser de nombreux facteurs de transcriptions et activer la transcription de programmes géniques (Park *et al.*, 2021). Enfin, on distingue quelques fonctions de EZH2, indépendantes du complexe PRC2, et même de son activité catalytique. En effet, dans certains cancers, en se liant avec d'autres facteurs, EZH2 peut exercer un rôle dans la régulation de la traduction ou de la réparation à l'ADN (Yi *et al.*, 2021 ; Koyen *et al.*, 2020)

Chez la souris, l'inactivation de *Ezh2* est létale au cours de la gastrulation. Le gène *Ezh1* n'est pas exprimé dans les blastocystes et ne peut donc compenser la perte de *Ezh2* (O'Carroll *et al.*, 2001). La

protéine EZH2, et son activité dans le complexe PRC2 sont indispensables dans le développement très précoce des mammifères. Chez l'homme, des mutations de *EZH2* ont été identifiées dans le syndrome de Weaver. Ce syndrome est caractérisé par une croissance excessive avant et après la naissance, une macrocéphalie, des retards développementaux et intellectuels. La majorité des patients portent des mutations dites « perte de fonction » dans le domaine SET de EZH2 (Cyrus *et al.*, 2019).

### 2.1.2 SUZ12

La protéine SUZ12 participe à l'assemblage du complexe PRC2. La délétion de *Suz12* dans les embryons murins conduit à une létalité très précoce, de la même façon que les embryons *ezh2*<sup>-/-</sup>. Cette mortalité est associée avec une perte des marques H3K27me3 et H3K27me2. Ces résultats démontrent que SUZ12 est indispensable à l'activité de PRC2 et ainsi participe aux processus de développements précocement (Pasini *et al.* 2004).

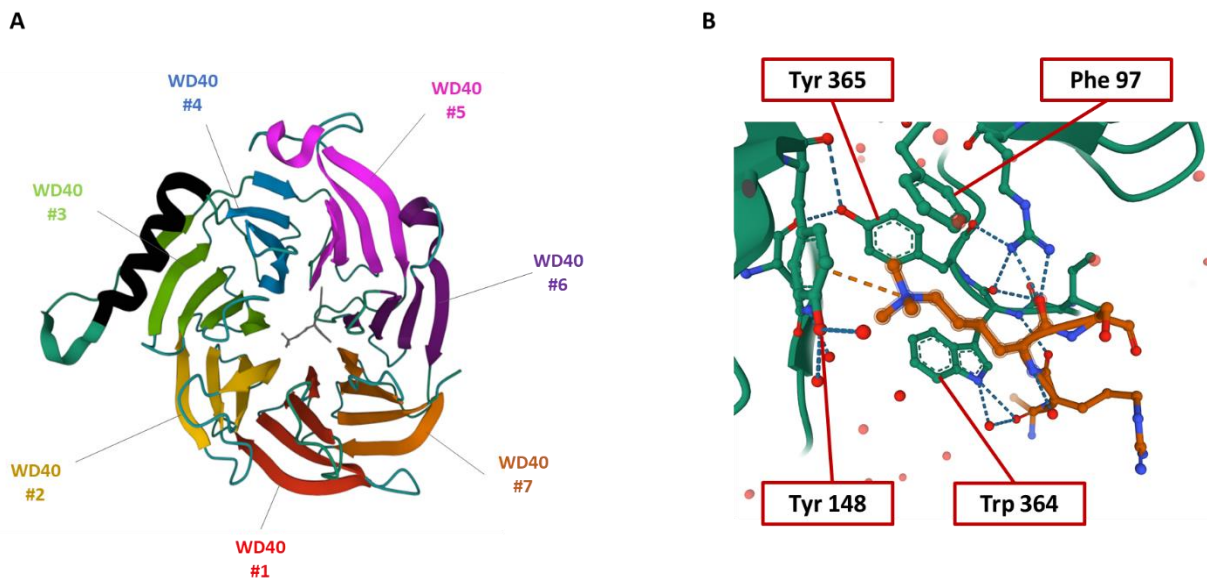
Initialement, SUZ12 était décrit avec un large domaine VEFS (VRN2-EMFS2-FIS2-Su(z)12) box dans la région C-terminale, caractérisé par un cluster d'acides aminés acides, et une séquence riche en tryptophane/méthionine. Ce domaine forme une plateforme pour les interactions avec EED et EZH2/1. Le domaine VEFS de SUZ12 seul semble être suffisant pour la formation du complexe. Deux autres domaines ont été identifié, un domaine WDB (WD40-binding domains) et un domaine à doigt de Zinc (Ciferri *et al.*, 2012). Mais la structure de SUZ12 et ses domaines d'interaction se sont révélés plus complexes. La résolution de la structure du complexe PRC2 montre que SUZ12 interagit avec RBBP4 par 4 domaines distincts : deux domaines WDB, un domaine à doigt de Zinc, et un domaine de fixation au doigt de Zinc. Un domaine C2 a été mis en évidence plus tardivement. Les domaines C2 sont nécessaires pour la liaison aux phospholipides et des interactions protéiques. Les protéines AEBP2 et PHF19 sont en compétition pour se fixer sur ce domaine de SUZ12. Ce domaine de SUZ12 permet donc l'assemblage de complexe PRC2 différents, ce qui modifie le recrutement de ces complexes (Chen *et al.*, 2018).

De la même façon que EZH2, des mutations de SUZ12 sont également recensées dans des troubles du développement chez l'homme. Ces troubles sont en revanche moins bien connus, et moins représentés puisque seulement 3 cas ont été répertoriés à ce jour (Cyrus *et al.*, 2019).

### 2.1.3 EED

L'orthologue de *Esc* chez les mammifères est *EED* (*Embryonic ectoderm development*). La comparaison de séquence de *Eed* et *Esc* a révélé l'orthologie de ces séquences avec une identité supérieure à 50%

et de près de 80% au niveau des domaines protéiques (Schumacher *et al.*, 1996). La séquence de EED est composée d'un domaine de liaison protéine-protéine en N-terminal et d'une répétition de 7 domaines WD40 (Denisenko & Bomsztyk, 1997 ; Han *et al.*, 2007). Les domaines WD40 sont des courtes séquences d'une quarantaine d'acides aminés terminées par un motif Trp-Asp (WD). La répétition de ces motifs favorise une conformation tridimensionnelle en hélice d'avion (Figure 10 A). Les domaines WD40 forment ainsi une plateforme pour créer de nombreuses interactions avec des protéines mais aussi des séquences nucléotidiques. Le domaine EBD de EZH2 par exemple est reconnu par les domaines WD40 (Han *et al.*, 2007) Les domaines WD40 reconnaissent certains motifs de EZH2/EZH1 (Han *et al.*, 2007) et l'ensemble des domaines EBD, SBD BAM et SANTL1L de EZH2 entourent la protéine EED (Jiao & Liu, 2015). Les résidus aminés aromatiques Phe97, Tyr148, Trp364 et Tyr365 sont organisés dans l'espace pour former une cage aromatique qui interagit avec la marque épigénétique H3K27me3 (Figure 10 B) (Xu *et al.*, 2010). Dans ce sens, EED permet la propagation de cette marque en activant le complexe pour asseoir la répression Polycomb. Néanmoins, sa présence est toujours nécessaire dans le complexe PRC2 pour son activité catalytique. En effet, l'enroulement de EZH2 autour de EED et la connexion entre les domaines SBD-SANT1L est nécessaire à l'activité enzymatique de EZH2 (Jiao *et al.*, 2015).



**Figure 10 : Structure de EED.** A. La protéine EED est organisée en 7 domaines WD40. Chaque domaine est composé de 4 feuillets  $\beta$  antiparallèle. Les 7 WD40 forment une plateforme dédiée aux interactions avec les autres protéines et avec la marque H3K27me3. L'interaction avec La lysine 27 triméthylée est représentée (au centre en gris) schématisée. B. La cage aromatique est composée de 4 acides aminés aromatiques. Ils forment des interactions de van de Walls avec la lysine triméthylée en orange. Réalisée sur PDB (<https://www.rcsb.org>), numéro d'accèsion : 3JZG.

Une lignée de souris avec une mutation hypomorphe de *Eed* présente des transformations antéro-postérieures tout au long de la colonne vertébrale, ce qui est similaire à la mutation de *Esc* chez la

drosophile (Schumacher *et al.*, 1996). De la même façon que les protéines SUZ12 et EZH2, un KO de EED chez la souris est létal durant la gastrulation, avec un mauvais positionnement du mésoderme dans les régions extra-embryonnaires (Faust *et al.*, 1998). Les cellules ES délétées de EED ont un taux de H3K27me3 réduit. Malgré une co-occupation de nombreux loci par PRC1 et PRC2, des programmes géniques sont déréprimés et les cellules s'engagent dans des voies de différenciation (Boyer *et al.*, 2006). Ces résultats montrent que EED est indispensable à l'activité du complexe, et que le complexe PRC2 est nécessaire pour le maintien du statut souche des cellules. Chez l'homme, le gène *EED* est impliqué dans des troubles du développement. Les patients avec le syndrome de Cohen-Gibson portent une mutation autosomique dominante de *EED* et présentent un tableau clinique proche du syndrome de Weaver (Cohen & Gibson, 2016). Néanmoins les deux syndromes dénotent quelques différences. Les déficiences intellectuelles sont plus fréquentes et sévères dans le syndrome de Cohen-Gibson que dans le syndrome de Weaver (Cyrrus *et al.*, 2019), ce qui interroge sur le rôle différentiel de EZH2, de EED et ainsi du PRC2 dans le développement neural. Enfin, chez la souris, la perte de EED dans les ovocytes conduit à une croissance anormale des embryons, une augmentation de l'adiposité ainsi que de la densité minérale osseuse (Prokopuk *et al.*, 2018). Ces résultats suggèrent donc un rôle prépondérant de EED dans le développement précoce.

### 2.1.4 RBBP4/7

Les protéines RBBP4 et RBBP7 (Retinoblastome binding protein 4 et 7) sont des protéines à fort taux d'homologie. Elles sont constituées de 7 domaines WD40 et deux hélices  $\alpha$  en position C- et N-terminal. Ces domaines leur permettent de se lier avec les sous-unités du complexe PRC2, mais également avec les histones H3 et H4 (Murzina *et al.*, 2008 ; Verreault *et al.*, 1996). Les protéines RBBP4/7 interagissent également avec de nombreux autres acteurs, et comme des « erasers ». L'interaction entre BCL11A et RBBP4 permet le recrutement du complexe NuRD. Ce complexe retire les acétylations des cibles génomiques et autorise la triméthylation de la lysine 27 de l'histone H3 médiée par le PRC2 (Moody *et al.*, 2018). La protéine RBBP4 est également nécessaire, en association avec SUZ12, pour assurer la liaison des protéines accessoires AEBP2 et PHF19 et recruter le complexe PRC2 de manière spécifique (Chen *et al.*, 2018). Des expérimentations de KO chez la souris ont déterminé une redondance de fonction entre RBBP4 et RBBP7. La perte de fonction de chacune des protéines n'a pas d'effet sur le développement précoce des souris. En revanche, la double mutation de RBBP4/7 provoque un arrêt du développement au moment avant le stade blastocyste (Xiao *et al.*, 2021).

### 2.1.5 Les protéines accessoires

De nombreuses protéines accessoires ont été identifiées associées au complexe PRC2. Dans un premier temps, les études ont identifié des partenaires qui interagissent ensemble, ou qui sont mutuellement exclusifs (Alekseyenko *et al.*, 2014 ; Hauri *et al.*, 2016). Un nouveau paradigme a alors été mis en avant, la formation de deux types de complexes différents, les complexes PRC2.1 et PRC2.2 (Figure 11) (Holoch & Margueron, 2017). Même si les deux types de complexes cooccupent une majorité des loci cibles de PRC2, la variabilité de composition des complexes permet de réguler différemment le recrutement et l'activité enzymatique des complexes (Healy *et al.*, 2019).

**Les complexes PRC2.1.** Ils comportent au moins un représentant des Polycomb-Like Proteins (PCLs), la protéine EPOP (Elongin BC and Polycomb repressive complex 2-associated Protein) et LCOR (Ligand Dependent Nuclear Receptor Corepressor) Les PCLs regroupent les protéines PHF1, MTF2 et PHF19. Chacune d'entre elle possède un domaine Tudor et deux domaines à doigt PHD qui reconnaissent les lysines méthylées. L'insertion des PCLs au sein du complexe PRC2 est exclusive. Les PCLs permettent un recrutement spécifique en fonction de l'intégration de l'une ou l'autre des protéines (Holoch & Margueron, 2017). Le domaine Tudor des PCLs a une affinité particulière pour la marque H3K36me3, qui est une marque associée à des loci actifs transcriptionnellement. Les protéines PCLs permettent donc de recruter le complexe PRC2 sur des loci actifs afin d'initier la répression épigénétique (Ballaré *et al.*, 2012 ; Cai *et al.*, 2013). La protéine EPOP, autre constituante des complexes PRC2.1, interagit avec ELOB (Elongin BC) en N-terminal et le complexe PRC2 en C-terminal, ce qui pourrait permettre de relier le PRC2 au complexe d'élongation de la transcription. La protéine EPOP permet de maintenir un niveau basal de transcription des cibles génomiques de PRC2 dans les cellules ES et ainsi de maintenir une plasticité des cellules pluripotentes (Beringer *et al.*, 2016). Enfin, la présence des protéines LCOR/LCORL dans le complexe, auparavant connue comme PALI1/2, est mutuellement exclusive avec EPOP (Aleksenyenko *et al.*, 2014). LCOR/LCORL permettent le recrutement de corépresseurs et contrôlent la régulation de l'identité cellulaire au cours du développement (Conway *et al.*, 2018). Au cours de la différenciation, l'intégration des protéines accessoires varie, et ces phénomènes sont particulièrement dynamiques et nécessaires pendant la différenciation neurales (Kloet *et al.*, 2016).

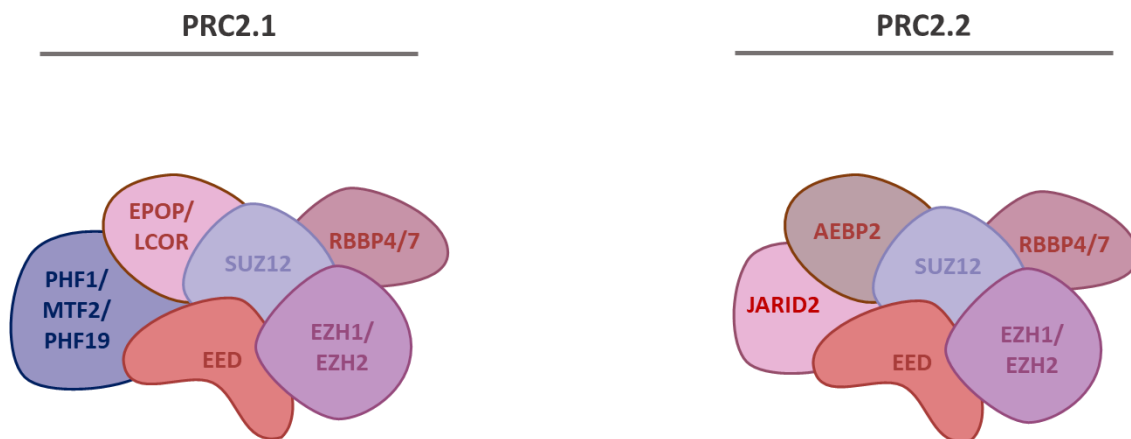


Figure 11 : Composition des complexes PRC2.1 et PRC2.2

**PRC2.2.** Les complexes PRC2.2 sont composés des protéines AEBP2 (AE Binding Protein 2) et JARID2 (Jumonji, AT rich interactive domain 2). Contrairement aux protéines accessoires des complexes PRC2.1, les deux protéines sont intégrées de manière équivalente dans le complexe (Smits *et al.*, 2013). La protéine AEBP2 possède un domaine à doigt de zinc qui interagit avec les autres constituants du complexe PRC2 ainsi que des motifs de liaison à l'ADN (Kim *et al.*, 2009). Cette protéine stimule le PRC2 en le stabilisant dans une conformation favorable à l'activité enzymatique (Ciferri *et al.*, 2012). La protéine JARID2 contient un domaine histone déméthylase JmjC sans activité catalytique mais l'activité de JARID2 réside dans les domaines ARID (AT-rich interaction domain) et à doigt de zinc qui fixent les nucléosomes (Li *et al.*, 2010). Le complexe PRC2 est capable de triméthyliser la lysine 116 de JARID2. Cette modification post-traductionnelle de JARID2 est reconnue par EED et stimule l'action du PRC2. Cette boucle de contrôle positive régule la répression Polycomb durant la détermination des lignages cellulaires (Sanulli *et al.*, 2015). La protéine JARID2 permettrait également de recruter le complexe PRC2 en se fixant sur des régions génomiques spécifiques, avec une affinité particulière pour les séquences riches en GC (Kim *et al.*, 2003). De manière générale, les protéines AEBP2 et JARID2 sont stimulantes pour l'activité du complexe PRC2.

## 2.2 Recrutement du complexe

De la même façon que pour le PRC1, chez la drosophile le complexe PRC2 est principalement recruté sur les PRE via son interaction avec les protéines Pho/Phol (Pleiohomeotic/Pleiohomeotic like) (Fritsch *et al.*, 1999). Chez les mammifères la protéine YY1 semblent reprendre les fonctions de Pho en contactant les complexes PRC1 et PRC2 (Thomas & Seto, 1999). Il existe donc une interconnexion importante entre les complexes Polycomb, mais cette connexion seule ne peut expliquer les mécanismes de recrutement du PRC2.

### 2.2.1 Par le complexe PRC1

Historiquement le modèle hiérarchique affirmait que le complexe PRC2 déposait la marque H3K27me<sub>3</sub>, et cette marque était utilisée comme signal de recrutement par les protéines CBX du PRC1. Depuis, ce dogme a largement été remis en cause, par les nombreux autres mécanismes de recrutement du PRC1 ainsi que la répartition des marques de H3K27me<sub>3</sub> et H2AK119ub1 (Boyer *et al.*, 2006 ; Ku *et al.*, 2008). Du fait de la capacité des complexes PRC1, et particulièrement les complexes PRC1.1, PRC1.3 et PRC1.5, à catalyser l'ubiquitination de la lysine 119 de l'histone H2A indépendamment de PRC2, l'équipe de Blackledge s'est demandé si cette activité pouvait recruter le PRC2 en retour. Ils ont montré que la marque H2AK119ub1, majoritairement médiée par le complexe PRC1.1 était effectivement nécessaire pour le recrutement de PRC2 et la formation de domaine de répression Polycomb (Blackledge *et al.*, 2014). Les protéines JARID2 et AEBP2 du complexe PRC2.2 reconnaissent la lysine 119 de l'histone H2A. Lorsque cette lysine est ubiquitinée, l'activité du PRC2 est stimulée, et le dépôt de la marque H3K27me<sub>3</sub> est augmenté (Kalb *et al.*, 2014).

### 2.2.2 Par les ARNnc

Le recrutement de PRC2 par les ARNnc a été mise en cause la première fois dans le cadre de l'inactivation du chromosome X. La sous-unité catalytique EZH2 fixe l'ARN RepA (Zhao *et al.*, 2008) tandis que la protéine JARID2 interagit avec le lnc Xist et engage le complexe PRC2 dans l'inactivation du chromosome X (da Rocha *et al.*, 2014). Le long ARNnc HOTAIR, issu du cluster de gène HOXC, est nécessaire à la répression du locus HOXD. Pour cela, il recrute le complexe PRC2 afin d'appliquer la marque épigénétique H3K27me<sub>3</sub> (Rinn *et al.*, 2007). Le premier long ARNnc découvert H19 interagit également avec le complexe PRC2. En plus d'appliquer une régulation de l'expression de *EZH2* (Lin *et al.*, 2016), le recrutement de EZH2 par H19 favorise la croissance tumorale et les propriétés métastatiques de plusieurs cancers (Chen *et al.*, 2019 ; Fazi *et al.*, 2018). Dans le cancer du poumon, plus d'une dizaine d'ARNnc interagissent avec EZH2 pour promouvoir les capacités des cellules tumorales (Su *et al.*, 2018). Des mécanismes de recrutement et de régulation fonctionnelle par les ARNs sur plusieurs autres sous-unités du PRC2 ont ainsi été mis en évidence (Almeida *et al.*, 2020). Enfin, des expériences d'immunoprécipitation d'ARN révèlent dans les cellules ES la fixation aspécifique de nombreux ARN sur le PRC2 (Zhao *et al.*, 2010). L'activité du complexe PRC2 est négativement impactée par la fixation de ces ARNs sur EZH2, mais l'affinité de EZH2 pour ces ARNs est fortement diminuée lorsque la protéine est incluse dans le complexe (Cifuentes-Rojas *et al.*, 2014). Ce mécanisme de fixation aspécifique des ARNs sur le PRC2 permettrait une évaluation à l'échelle du

génomique de l'activité transcriptionnelle des gènes et ainsi un recrutement de PRC2 en fonction des besoins (van Kruijsbergen *et al.*, 2015).

### 2.2.3 Par les protéines

De nombreuses études ont mis en avant la colocalisation des PcGs et des facteurs de transcriptions (FTs) au travers du génome, ce qui suggère une relation entre les PcGs et ces protéines (Bernstein *et al.*, 2006 ; Boyer *et al.*, 2006 ; Lee *et al.*, 2006), et plus précisément que ces facteurs de transcription seraient capables de recruter le PRC2 sur des loci particuliers (Benveniste *et al.*, 2014). Les facteurs de transcription STAT3, OCT3-4 et NANOG par exemple sont essentiels pour le maintien des capacités de pluripotence des cellules ES. Des expériences de CHIP révèlent que STAT3 et OCT3-4 se fixent au niveau du promoteur de *Eed* et donc semblent pouvoir favoriser sa transcription. Ces FTs peuvent également recruter EED, et ainsi le PRC2, sur les loci de gènes de différenciation cellulaire, ce qui révèle l'importance du complexe PRC2 dans l'auto-renouvellement des cellules souches et le maintien de leur capacité (Ura *et al.*, 2008). L'interaction physique des FTs avec les PcGs a d'ailleurs été démontrée de nombreuses fois. Des séquences de réponse au facteur REST corrélaient avec une forte présence de H3K27me3 (Dietrich *et al.*, 2012). De la même façon, des séquences contenant les sites de fixation au facteur SNAIL sont suffisants pour induire le recrutement du PRC2 (Arnold *et al.*, 2013). La présence de ces motifs orchestre la répression Polycomb et dirige l'expression de gènes nécessaires à la neurogénèse (Arnold *et al.*, 2013 ; Tien *et al.*, 2015).

### 2.2.4 Par les histones modifiées

La modification post-traductionnelle de l'histone H2A n'est pas la seule qui puisse moduler le recrutement et l'activité du complexe PRC2 via JARID2. La marque H3K27me3 elle-même peut réguler l'activité du complexe. En effet, reconnue par la sous-unité EED, elle est utilisée comme plateforme pour le PRC2 afin de propager la répression Polycomb (Xu *et al.*, 2010 ; Jiao *et al.*, 2015). Un unique complexe PRC2 peut de ce fait se fixer sur un nucléosome dont la lysine 27 de l'histone H3 est triméthylée et déposer cette marque épigénétique sur un autre nucléosome (Poepfel *et al.*, 2018). La protéine EED a également une bonne affinité pour plusieurs modifications post-traductionnelles d'histone. La triméthylation de l'histone H1 sur la lysine K26 a un effet inhibiteur sur l'activité du PRC2. En revanche, la fixation de EED sur d'autres marques peut avoir un effet stimulant sur le PRC2. La marque H3K9me3 semblent pouvoir recruter le complexe PRC2 (Xu *et al.*, 2010). Le complexe PRC2 est capable d'interagir directement avec la méthyltransférase G9a/GLP, qui catalyse la triméthylation de la lysine 9 de l'histone H3 (Mozetta *et al.*, 2014). Le recrutement de PRC2 au niveau de la marque



H3K9me3 peut également être indirecte, via la protéine HP1 afin d'établir l'hétérochromatine (Boros *et al.*, 2014).

Les protéines PCLs possèdent des domaines Tudor avec une affinité particulière pour la marque d'euchromatine H3K36me3 (Ballaré *et al.*, 2012). Les raisons de ce recrutement de PRC2 est encore sujet à controverse. La première hypothèse était la possibilité d'appliquer une répression Polycomb au niveau de site transcriptionnellement actifs. Néanmoins, PRC2 est inhibé lorsque PHF1 reconnaît la marque H3K36me3 (Yuan *et al.*, 2011 ; Musselman *et al.*, 2012) et pourrait permettre de réduire l'activité du PRC2 au niveau de régions génomiques qui doivent rester transcriptionnellement actives. Enfin, à l'époque de la découverte de la découverte des PcGs chez la drosophile, un autre groupe de protéine a été identifié, appartenant au complexe Trithorax (TrxG). Les TrxGs triméthylent la lysine 4 de l'histone H3, et cette marque épigénétique est activatrice, ce qui place l'action des PcGs et des TrxGs en opposition. Les marques H3K4me3 et H3K27me3 sont mutuellement exclusives et antagonistes, ce qui permet de maintenir la capacité de switch ON/OFF stablement (Reinig *et al.*, 2020 ; Loh & Veentra, 2022).

### 2.2.5 Par les îlots CpG

Dans un premier temps, le recrutement de PRC2 au niveau des îlots CpGs non-méthylés fut un recrutement indirect. La sous-unité KDM2B reconnaît ces îlots et recrutent le complexe PRC1.1 sur ces parties du génome. La marque H2AK119ub1 est déposée et est reconnue par JARID2 au sein du complexe PRC2. Dans un second temps, un recrutement via les protéines PCL de PRC2.2 a été proposé. Les protéines PHF1 et MTF2 possèdent une structure particulière dans leur partie N-terminal, le domaine EH. Ce domaine en hélice possède une affinité pour les îlots CpG et permet un recrutement direct de PRC2.2 (Li *et al.*, 2017 ; Perino *et al.*, 2018). Mais la structure et la densité des CpGs semblent jouer un rôle très important. La densité des CpG ainsi que le pourcentage de G/C semble être importante pour l'état non méthylé de l'ADN, mais également l'état de la chromatine. Le dépôt de marque épigénétique comme H3K27me3 ou H3K4me3 au niveau de ses régions est l'état par défaut (Wachter *et al.*, 2014). Chez les mammifères, il existe une grande richesse de CpG au niveau des promoteurs des gènes, tandis que c'est beaucoup plus rare chez les vertébrés non mammifères. Néanmoins il existe des îlots CpG très bien conservés au niveau de promoteur de certains gènes (Long *et al.*, 2013), et la triméthylation de la lysine 27 de l'histone H3 au niveau de ces régions hypométhylées est également bien conservée (van Kruijsbergen *et al.*, 2015). Ces îlots semblent particulièrement déterminant dans le recrutement de PRC2 puisque des analyses *in silico* avec des logiciels entraînés

pour prédire les sites de déposition de la marque H3K27me3 seulement à partir de la densité de CpG ont identifié correctement les régions, et ce sur des séquences issues de génomes de plusieurs espèces de vertébrés (Van Heeringen *et al.*, 2014).

### 2.3 Les modifications post-traductionnelles de la lysine 27 de l'histone H3

Il existe 4 modifications post-traductionnelles de la lysine 27 de l'histone H3. Chacune de ces marques est répartie précisément dans le génome, et est associée avec un état transcriptionnel propre. Chacune de ses marques épigénétiques a donc un rôle dédié.

#### 2.3.1 La triméthylation

La marque H3K27me3 est une marque de l'hétérochromatine facultative. Récemment des études ont levé le voile sur le rôle de cette marque dans l'établissement de l'empreinte génomique. En effet, le mécanisme d'empreinte le plus connu est effectué par méthylation de l'ADN. Selon les loci, l'allèle maternel ou paternel est alors méthylé pour ne plus être exprimé. Au cours du développement précoce des souris, des domaines portant la marque H3K27me3 sont établis principalement sur l'allèle d'origine maternelle et porte un mécanisme d'empreinte génomique non canonique (Raas *et al.*, 2022). De par le rôle répressif de la marque H3K27me3, elle est impliquée dans de grands processus de maintien des capacités d'auto-renouvellement, dans la détermination des lignages cellulaires et dans le processus de différenciation complet de types cellulaires variées (Loh & Veentra, 2022 ; Iwagawa & Watanabe, 2019). La triméthylation de la lysine 27 est une modification post-traductionnelle relativement longue à effectuer car elle nécessite l'application de groupements méthyl de façon séquentielle. C'est aussi une marque épigénétique particulièrement stable au cours du temps (Zee *et al.*, 2010). La marque H3K27me3 est principalement distribuée sur les promoteurs des gènes et le long de gènes réprimés (Højfeldt *et al.*, 2018)

#### 2.3.2 La diméthylation

Bien que la marque H3K27me2 soit considérée comme une marque épigénétique répressive, elle a été dans un premier temps considérée comme une marque transitoire, qui prévient l'acétylation de cette lysine (Margueron & Reinberg, 2011). Cette marque est l'état de méthylation le plus fréquent, environ 50% des histones H3 sont diméthylées sur leur lysine 27 (Ferrari *et al.*, 2014), au niveau de lysines dépourvues des marques H3K27me1 ou H3K27me3 (Højfeldt *et al.*, 2018).]. Au sein du génome des ESCs, les marques H3K27me2 et H3K26me2 sont colocalisées, ce qui suggère que la marque H3K27me2 soient présentes sur des gènes actifs transcriptionnellement (Streubel *et al.* 2018). Parmi les gènes qui

portent la marque H3K27me2, plus de 60% d'entre eux sont transcrits activement, et sont impliqués dans le métabolisme. Néanmoins, d'autres gènes, impliqués notamment dans le système immunitaire, sont mis sous silence. Le ratio entre la présence de H3K27me2 et H3K27me3 permet une régulation fine de l'identité cellulaire et de l'induction de la différenciation au cours du développement. La lysine 27 déméthylée peut facilement être triméthylée, ou au contraire être acétylée. La marque H3K27me2 autorise l'expression ou la répression génique de façon dynamique, en fonction des stimuli (Juan *et al.*, 2016).

### 2.3.3 La monométhylation

La marque épigénétique H3K27me1 possède un rôle de régulation génique propre et différent de celui des marques H3K27me2/3. En effet, la répartition de ces marques au travers du génome est spécifique (Laugesen *et al.*, 2019). La marque H3K27me3 est principalement observée au niveau des promoteurs des gènes, dans des îlots CpGs. En revanche, si la présence de H3K27me1 dépend du type cellulaire et ainsi des gènes exprimés, dans les cellules érythroïdes, la marque est enrichie au niveau du corps de gènes activement transcrits (Steiner LA *et al.*, 2011 ; Ferrari *et al.*, 2014). Les histones H3 péri-centromériques sont également monométhylées (Peters *et al.*, 2003).

L'enzyme responsable de la monométhylation de la lysine 27 de l'histone H3 est source de débat depuis plusieurs années. La première enzyme reportée pour cette activité est EZH1 (Shen *et al.*, 2008). Alors que la perte de fonction de EZH2 n'impacte pas le taux de H3K27me1, le double mutant *Ezh1*<sup>-/-</sup> *Ezh2*<sup>-/-</sup> abolit totalement la présence de cette marque, ce qui suggère un rôle prépondérant de EZH1 dans la méthylation de la lysine 27 de l'histone H3 dans les cellules ES. Par la suite, le rôle des deux sous-unités EZH1 et EZH2 dans la monométhylation de la lysine 27 de l'histone H3 a été démontré (Ezhkova *et al.*, 2011 ; Ferrari *et al.*, 2014). Seulement, d'autres méthyltransférases ont été proposées pour la réalisation de la monométhylation de H3K27 : G9a (EHMT2) et Glp (EHMT1). Ces deux méthyltransférases déposent trois groupement méthyls sur la lysine 9 de l'histone H3 (Tachibana *et al.*, 2005). Les interactions entre G9a/GLP et le PRC2 permettent de réguler l'activité du PRC2 (Mozzeta *et al.*, 2014) et semble permettre aux protéines G9a/GLP la méthylation de H3K27 (Wu *et al.*, 2011). L'inhibition de G9a dans les cancers de la vessie semble diminuer le taux de H3K27me1 dans les cellules cancéreuses (Segovia *et al.*, 2019) alors que la perte de fonction n'affecte pas la méthylation de H3K27 dans les ESCs (Mozzeta *et al.*, 2014), ce qui pourrait suggérer une action de G9a seulement dans des conditions spécifiques qui restent à déterminer.

### 2.3.4 L'acétylation

Les protéines CBP (CREB-Binding Protein) possèdent un bromodomaine qui permet une fixation des résidus de lysines acétylées, ainsi qu'une activité histone acétyltransférase. Ces deux protéines sont impliquées dans l'acétylation de nombreuses histones, y compris l'histone H3, sur sa lysine 27 (Kouzarides, 2007 ; Jin *et al.*, 2011). L'acétylation de la lysine 27 de l'histone H3 est une marque épigénétique activatrice, associée à l'euchromatine. La marque h3K27ac est préférentiellement observée au niveau des promoteurs des gènes (Creyghton *et al.*, 2010 ; Lavarone *et al.*, 2019). En compétition avec la triméthylation, la perte de l'activité du PRC2 provoque une hausse globale du taux d'acétylation (Pasini *et al.*, 2010). La marque H3K27ac remplace les endroits précédemment occupés par H3K27me3/2 en cas de perte de fonction de Ezh2. La diffusion de cette marque participe à la dérégulation génique en cas de perte de H3K27me3 et semble favoriser l'apparition des phénotypes très tôt dans le développement (Lavarone *et al.*, 2019). La signature épigénétique et la balance entre les positions des marques H3K27ac et H3K27me3 sont particulièrement importantes dans le développement précoce des membres de la souris. Selon les tissus et le stade de développement, la présence de ces marques autorise l'expression d'un set prédéfini de gènes, nécessaire à la formation des membres (Cotney *et al.*, 2012).



## CHAPITRE II : LA REPRESSION POLYCOMB DANS LE DEVELOPPEMENT

---

Les protéines Polycomb ont été identifiées par des études fonctionnelles du développement de la drosophile. L'apparition de phénotypes proches en cas de perte de fonction des protéines a permis de les regrouper en famille. Ces phénotypes apparaissent chez la drosophile très précocement, au stade larvaire, ce qui suggère un rôle prépondérant des protéines de la famille Polycomb dans le développement précoce. De la même façon, de nombreuses mutations des PcGs sont impliquées dans l'apparition de troubles du développement chez l'homme (Cyrrus *et al.*, 2019). Afin d'identifier le rôle précis de chacune des protéines dans le développement des mammifères, des modèles murins KO ont été développés ces dernières années (Table 2). Ces modèles révèlent une implication large des PcGs dans les processus développementaux précoces et au cours de la vie des souris.

## CHAPITRE II : LA REPRESSION POLYCOMB DANS LE DEVELOPPEMENT

Table 2 : Recensement des lignées mutantes et les phénotypes associés chez le modèle murin et les phénotypes observés chez l'homme.

### Canonical PRC1

Gène	Phénotypes	Références	Pathologies humaines reportées	Références
<i>Ring1</i>	Transformation homéotique du squelette axial	del Mar Lorente et al., 2000	Syndrome neurodéveloppemental, malformations squelettiques, anomalies du système immunitaires ; perturbation de la décidualisation des tissus utérins	Pierce et al., 2018 ; Bian et al., 2016
<i>Rnf2</i>	Transformations homéotiques létales. Différenciation prématurée des cellules neurales et hématopoïétiques	Voncken et al., 2003 Leeb et al., 2007	Impliqué dans l'apparition et la progression de types variés de cancers ; surexpression dans les fibroses pulmonaires ; prévient l'apparition d'hypertrophie cardiaque associée à la sénescence ; Syndrome neurodéveloppemental ; spécification des cellules souches embryonnaires ; différenciation précoce des cellules neurales	Yan et al., 2021 ; Pan et al., 2022 ; Chen et al., 2022 ; Luo et al., 2021 ; Desai et al., 2020
<i>Cbx2</i>	50% de mortalité périnatale, transformations homéotiques du squelette, défauts de l'hématopoïèse et de maturation de la rate, développement incomplet de la glande surrénale, Inversion du sexe male à femelle	Katoh-Fukui et al., 1998 Katoh-Fukui et al., 2012 Coré et al., 1997	Dysgénésie gonadique ; implication dans la progression tumorale	Sproll et al., 2019 ; Eid et al., 2015 ; Naqvi et al., 2023
<i>Cbx4</i>	Hypoplasie thymique, défaut de maintenance et répliation des cellules souches épidermiques	Luis et al., 2011 Liu et al., 2013	Perturbation de la décidualisation des tissus utérins ; contrôle de la sénescence ; participe à la réaction inflammatoire allergique asthmatique ; implication dans la progression tumoral	Bian et al., 2016 ; Ren et al., 2019 ; Luis et al., 2014 ; Ling et al., 2022 ; Naqvi et al., 2023
<i>Cbx6</i>	Diminution du poids du corps et de la masse graisseuse totale, troubles métaboliques	Sharnes et al., 2011	Implication dans la progression tumoral	Naqvi et al., 2023
<i>Cbx7</i>	Privilégie le développement de cancers du foie et du poumon, diminution de la taille des glandes sébacées	Forsati et al., 2012 White et al., 2013	Participe au stress du réticulum endoplasmique induit par une lésion d'ischémie-reperfusion	Zhang et al., 2020

## CHAPITRE II : LA REPRESSION POLYCOMB DANS LE DEVELOPPEMENT

<i>Cbx8</i>	Initiation altérée et maintenance de la transformation leucémique	Tan et al, 2011		
<i>Pcgf2</i>	Retard de croissance, transformations homéotiques, défaut de différenciation des cellules cardiaques et immunitaires	Akasaka et al., 1996 Akasaka et al., 1997	Syndrome Turnpenny-Fry (retard de développement, dysmorphie faciale, anomalies squelettiques)	Ercoskun et al., 2021
<i>Bmi1</i>	Problème de maintenance des cellules souches	van der Lugt et al., 1994 Orguro et al., 2010	Progression de la maladie d'Alzheimer ; prévient l'apparition d'hypertrophie cardiaque associée à la sénescence ; impliqué dans la progression de nombreux cancers	Chen et al., 2022 ; Flamier et al., 2018 ; Bhattacharaya et al., 2015
<i>Scml2</i>	Défaut du développement testiculaire et de la spermatogénèse, infertilité des males	Hasegawa et al., 2015		
<i>Phc1</i>	Spécification de l'axe antéro postérieur, développement cardiaque anormal, défaut de différenciation de la lignée cellulaire des lymphocytes B	Takahara et al., 1997 Isono et al., 2005 Shirai et al., 2002	Microcéphalie primaire autosomique récessive ; développement de lymphomes	Faheem et al., 2015 ; Raaphorst et al., 2004
<i>Phc2</i>	Spécification de l'axe antéro postérieur, transformations du squelette	Isono et al., 2005 Isono et al., 2013		

### PRC1.1

Gène	Phénotypes	Références	Pathologies humaines reportées	Références
<i>Pcgf1</i>	Diminution des capacités de différenciation des ESCs, spécification ectodermique et mésodermique	Yan et al., 2017	Enrichissement de cellules souches cancéreuses colorectales	Ji et al., 2021
<i>Rybp</i>	Létalité précoce, défaut du développement du système nerveux central et des cardiomyocytes, diminution de la maintenance des lymphocytes B	Pirity et al., 2005 Ujhelly et al., 2015 Calés et al., 2016.	Prévalence pour le développement de rhinosinusite chronique, sensibilité à l'insuline, progression de plusieurs types de cancers	Zhang et al., 2012 ; Lee et al., 2017 ; da Silva et al., 2018
<i>Kdm2b</i>	Létalité périnatale partielle, défaut du développement cérébrale, anomalies crânielles,	Fugua et al., 2011 Boulard et al., 2016	Favorise l'inflammation lors de sinusites chroniques avec polypes nasaux ; syndrome neurodéveloppemental ; impliqués dans des cancers variés	Liu et al., 2020 ; van Jaarsveld et al., 2023 ;



## CHAPITRE II : LA REPRESSION POLYCOMB DANS LE DEVELOPPEMENT

	non-inactivation du chromosome X			Vacik et al., 2018
<i>Bcor</i>	Défaut de formation des tissus extra-embryonnaires, létalité embryonnaire des mâles, retard d'activation de gènes de différenciation, défaut de formation du cœur	Huynh et al., 2000 Wamstad et al., 2008	Microphthalmie liée à l'X et syndrome oculofaciocardiodental ; épilepsie ; évènements de duplications et de fusions conduisant à la formation de néoplasmes	Ragge et al., 2019 ; Li et al., 2022 ; Astolfi et al., 2019
<i>Skp1</i>	Augmentation de la mort cellulaire des cellules neurales, diminution de la prolifération des cellules dans les organes lymphoïdes	Mandel et al., 2009	Maladie de Parkinson ; Syndrome de Sjögren ; cible thérapeutique potentielle pour bloquer la prolifération des cellules cancéreuses	Mandel et al., 2012 ; Sandhya & Danda, 2014 ; Hussain et al., 2016

### PRC1.3 et PRC1.5

Gène	Phénotypes	Références	Pathologies humaines reportées	Références
<i>Pcgf3</i>	Défaut de différenciation du mésoderme, absence de rate	Almeida et al., 2017 Zhao et al., 2017	Favorise la progression du cancer du poumon	Hu et al., 2021
<i>Pcgf5</i>	Défaut de différenciation du mésoderme	Almeida et al., 2017 Zhao et al., 2017	Favorise l'infection par la bactérie <i>Ehrlichia chaffeensis</i>	Mitra et al., 2018
<i>CKII</i>	Problème de maintien des cellules souches, défaut du développement cardiaque et des cellules hématopoïétiques	Huillard et al., 2010 Lou et al., 2008	Impliqué dans la maladie d'Alzheimer, Parkinson, Huntington, et des cancers variés	Roffey et al., 2021
<i>Auts2</i>	Défaut de développement du système nerveux central	Almeida et al., 2017 Gao et al., 2014	Syndrome d'haploinsuffisance de AUTS2 avec des retards neurodéveloppementaux ; Fusion avec <i>PAX5</i> à l'origine de leucémies lymphoblastiques aiguës	Sanchez-Jimeno et al., 2021 ; Denk et al., 2012

### PRC1.6

Gène	Phénotypes	Références	Pathologies humaines reportées	Références
<i>Pcgf6</i>	Diminution de la prolifération, différenciation du mésoderme, Différenciation des cellules germinales males	Endoh et al., 2017 Zdzieblo et al., 2014 Sun et al., 2015		
<i>L3mbtl2</i>	Létalité précoce, défaut de la gastrulation	Qin et al., 2012		

## CHAPITRE II : LA REPRESSION POLYCOMB DANS LE DEVELOPPEMENT

<i>E2f6</i>	Transformations homéotiques, expression ectopique des gènes spécifiques de la lignée germinale	Storre et al., 2002 Pohlers et al., 2005	Lymphome d'Hodgkin, favorise le développement oncogénique des carcinomes du nasopharynx	Sanchez-Beato et al., 2004 ; Li et al., 2021
<i>Max</i>	Létal à E5, impact sur la prolifération des cellules ES	Shen-Li et al., 2000	Progression et invasion tumorale	Mathsyaraja et al., 2021
<i>Mga</i>	Létal à E5	Washkowitz et al., 2015	Progression et invasion tumorale	Mathsyaraja et al., 2021

### PRC2

Gène	Phénotypes	Références	Pathologies humaines reportées	Références
<i>Eed</i>	Létal durant la gastrulation, déficit du mésoderme antérieur, différenciation des ESCs, croissance dérégulée en cas d'absence d'ARNm d'origine maternelle, troubles de différenciation du système nerveux central, problèmes de comportement, maintien des cellules souches intestinales, défaut de l'hématopoïèse fœtale, contrôle du développement épidermique	Faust et al., 1995 Faust et al., 1998 Obier et al., 2015 Prokopuk et al., 2018 Wang et al., 2022 Koppens et al., 2016 Wu et al., 2017 Dauber et al., 2016	Syndrome de Cohen-Gibson, interaction avec le VIH, apparition de MPNSTs et de myélomes	Cohen & Gibson, 2016 ; Violot et al., 2003, Lee et al., 2014 ; Score et al., 2012
<i>Suz12</i>	Létal durant la gastrulation, défaut du mésoderme, des cellules dendritiques, de l'homéostasie des cellules myéloïdes et des cellules souches hématopoïétiques, du développement des glandes mammaires	Hojfeld et al., 2018 Pasini et al., 2004 Zhan et al., 2021 Lee et al., 2015 Michalak et al., 2018 Dauber et al., 2016	Syndrome de Imagawa-Matsuto, apparition de MPNSTs, de sarcomes de l'endomètre	Imagawa et al., 2023; Lee et al., 2014 ; Hrzenjak, 2016
<i>Ezh2</i>	Létal durant la gastrulation, absence du maintien de l'homéostasie hépatique, du contrôle des Natural killer T cells , du métabolisme lipidique, de la différenciation des cellules neuronales dans le	O'Caroll et al., 2001 Bae et al., 2014 Tumes et al., 2019 Feng et al., 2016	Syndrome de Weaver, progression des NAFLD, mutations dans des types variés de cancer, régulation de la progression tumorale	Cyrus et al., 2019 ; Yamaguchi & Hung, 2014, Lim & Kim 2020

## CHAPITRE II : LA REPRESSION POLYCOMB DANS LE DEVELOPPEMENT

	cerebellum, contrôle du développement épidermique	Dauber et al., 2016		
<i>Ezh1</i>	Absence du maintien des cellules souches hématopoïétiques, de l'homéostasie hépatique et de la spermatogénèse	Hidalgo et al., 2012 Bae et al., 2014 Mu et al 2017 Yiew et al., 2019	Schizophrénie, cancers de la thyroïde	Johnstone et al., 2018, Calebiro et al., 2016
<i>Rbbp4</i>	Activation anormale du système immunitaire	Ishimaru et al., 2006	Progression de plusieurs types de cancers	Mladek et al., 2022 ; Li et al., 2020
<i>Rbbp7</i>	Défaut de la gastrulation et de la cardiogénèse	Cox et al., 2010		

### PRC2.1

Gène	Phénotypes	Références	Pathologies humaines reportées	Références
<i>Mtf2</i>	Létal à E15, retard de croissance important, contrôle de l'érythropoïèse	Rothberg et al., 2018 ; Perino et al., 2018	Variation d'expression à valeur pronostique dans plusieurs cancers	Tang et al., 2023
<i>Lcor/Lcorl</i>	Létalité périnatale	Conway et al., 2018	Régulation des cellules cancéreuses mammaires	Jalaguier et al., 2017
<i>Phf19</i>	Transformations homéotiques, défaut de l'hématopoïèse, prolifération des cellules souches hématopoïétiques	Vizàn et al., 2020	Hypertrophie cardiaque, régulation de la prolifération et de capacités d'invasion des cellules cancéreuses	Gu et al., 2021 ; Jain et al., 2020 ; Xiaoyun et al., 2021 ; Ren et al., 2019

### PRC2.1

Gène	Phénotypes	Références	Pathologies humaines reportées	Références
<i>Jarid2</i>	Létal à E10,5-E18,5, défaut de la formation du tube neural et du développement cardiaque.	Takeuchi et al., 1995 ; Takeuchi et al., 1999	Délétion dans les malignités myéloïdes chroniques	Puda et al., 2011
<i>Aebp2</i>	Létalité périnatale, transformations antérieures du squelette	Grijzenhout et al., 2016	Résistance de cancers ovariens à la cisplatine	Zhang et al., 2020

### 1. Le maintien des cellules souches et différenciation.

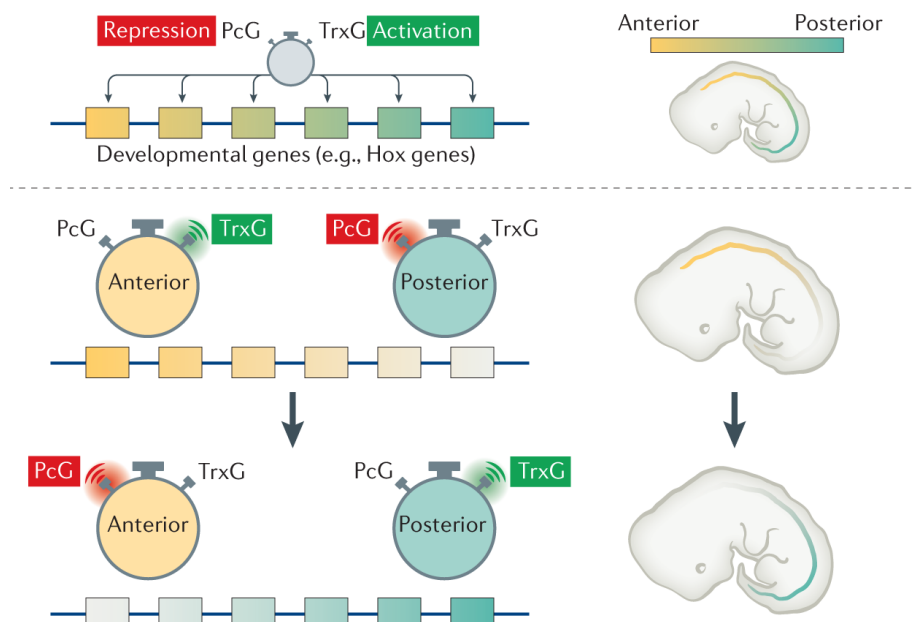
Les PcGs assurent un rôle de détermination et de maintien de l'identité cellulaire, en autorisant ou non l'expression de programme de différenciation. Chacune des sous-unités des deux complexes PRC1 et PRC2 ont d'ailleurs des positions génomiques différentes dans les cellules souches et dans les cellules différenciées (Montgomery *et al.*, 2005 ; Morey *et al.*, 2015). Dans les cellules souches, au stade embryonnaire ou adulte, de nombreux loci portent la marque H3K27me3, ainsi que la marque activatrice H3K4me3 (Shema *et al.*, 2016). La présence de ces deux marques détermine un état bivalent des gènes et du destin cellulaire. Selon les signaux de différenciation, la chromatine sera réorganisée. Certains promoteurs bivalents seront transcrits activement tandis que d'autres seront réprimés. La cellule suivra alors une voie de différenciation précise. La répartition et le recrutement des complexes sont donc fortement liés au statut de différenciation et à la lignée cellulaire (Schuettengruber *et al.*, 2017). Cette adaptation est possible par la multiplicité des PcGs et des moyens de recrutement des complexes ainsi que les interactions avec des acteurs variés, ce qui permet une adaptation dynamique. Dans les ESCs, *Cbx7* est le gène de la famille *CBX* le plus exprimé. Lorsque les cellules entrent en différenciation, *CBX7* est remplacée par un autre paralogue (Klauke *et al.*, 2013). D'autres unités des complexes PRC1 et PRC2 ont été mis en évidence dans le maintien des ESCs et leur différenciation (Table 3). Les cellules souches adultes sont également encadrées par la répression Polycomb. La protéine EZH2 réprime le locus *Ink4A-Ink4B* dans les cellules souches intestinales, ce qui maintient une capacité proliférative basale de ces cellules (Ezkhova *et al.*, 2009). La perte de la répression Polycomb induit la perte des capacités d'auto-renouvellement et de prolifération des cellules souches. Au sein du complexe PRC2, la présence de certaines protéines accessoires dictent le recrutement du complexe. L'intégration des protéines accessoires dans les complexes PRC2 est ainsi différente dans les cellules différenciées et indifférenciées (Kloet *et al.*, 2016 ; Oliviero *et al.*, 2016). Néanmoins, des questions demeurent sur la régulation de l'expression des gènes Polycomb dans les différents types cellulaires.

### 2. Le développement embryonnaire précoce.

Les pertes de fonction d'une grande partie des gènes Polycomb chez la souris sont à l'origine de l'arrêt prématuré du développement embryonnaire autour de la gastrulation. Ces résultats révèlent l'importance primordiale de la répression Polycomb dans le développement précoce. Les cellules embryonnaires suivent des mouvements morphogénétiques précis, qui conduisent à la formation des trois feuilletts embryonnaires : l'endoderme, le mésoderme et l'ectoderme. La mise en place de ces feuilletts est soumise à l'expression de programmes géniques spécifiques qui déterminent la position de chacune des cellules notamment sur l'axe antéro-postérieur (AP). La perte de fonction du gène *Eed*

inhibe la détermination de l'axe chez les souris (Faust *et al.*, 1998). Des gènes marqueurs du mésoderme sont alors retrouvés exprimés de manière ectopique dans les cellules de l'ectoderme proximal (Faust *et al.*, 1995).

De nombreuses transformations homéotiques sont également observées sur les modèles murins mutés pour les PcGs. Les gènes *Hox* permettent de déterminer l'identité cellulaire le long de l'axe AP. La régulation différentielle de l'expression des gènes *Hox* le long de l'axe AP est assurée en partie par les complexes Polycomb et TrxG, qui vont respectivement réprimer et activer les gènes *Hox* en fonction du positionnement cellulaire (Figure 12). La perte de *Ring1* et *Rnf2* conduit à une dérégulation des gènes *Hox* et ainsi à une modification de leur profil d'expression. Ce phénotype se traduit par l'apparition de modification du squelette des embryons (Voncken *et al.*, 2003 ; del Mar Lorente *et al.*, 2000).



**Figure 12 : Représentation schématique du contrôle de l'expression des gènes *Hox* le long de l'axe AP.** Les PcGs et TrxGs déterminent le profil d'expression des gènes le long de l'axe antéro-postérieur au cours du développement embryonnaire. Un set de gène spécifique est exprimé pour déterminer sa position le long de cet axe. Figure de Piunti & Shilatifard, 2021.

### 3. Le développement neural.

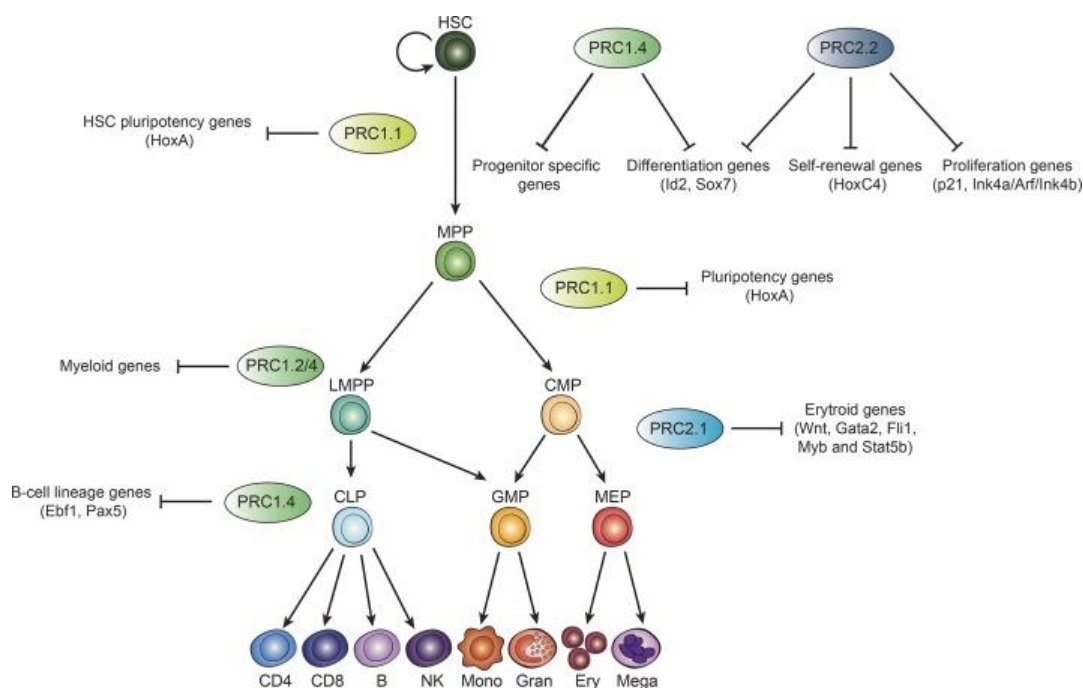
Les cellules souches neurales sont à l'origine de tous les lignages neuraux. Leur capacité d'auto-renouvellement et leur état de non-différenciation sont maintenus par la répression Polycomb, et en particulier par *Rnf2*. La perte de fonction de la sous-unité catalytique de PRC1 dans le bulbe olfactif provoque une perte de la population des cellules souches neurales.

Les cellules se différencient uniquement en neurone, via la diminution de l'activation de la voie de signalisation Notch (Román-Trufero *et al.*, 2009). La neurogénèse est un processus développemental très complexe. Elle met en jeu de nombreux acteurs moléculaires qui permettent la différenciation d'un progéniteur commun en cellules gliales et neuronales variées. Ce processus de différenciation est orchestré par une modification de l'environnement chromatinien, et en partie par la répression Polycomb (Corley & Kroll, 2015). Les progéniteurs neuraux possèdent des loci bivalents. Lorsque les progéniteurs s'engagent dans la voie de différenciation neuronale, les PcGs sont relocalisées afin de réprimer les programmes géniques associés à la différenciation gliale (Mohn *et al.*, 2008). L'absence d'activation des programmes géniques associés à la neurogénèse conduit à la gliogénèse. Les progéniteurs gliaux maintiennent également des loci bivalents jusqu'à la spécification astrocytaire ou oligodendrocytaire. Le niveau d'expression de *Ezh2* dans les progéniteurs neuraux semble conduire la cellule dans une voie de différenciation. En effet, une forte expression de *Ezh2* est observée dans la lignée des précurseurs d'oligodendrocyte et ce jusqu'à la différenciation terminale tandis qu'elle est absente de la différenciation des neurones corticaux (Pereira *et al.*, 2010). Au sein de cette lignée, *Ezh2* est principalement retrouvé au niveau de promoteurs de gènes de différenciation neuronale ou astrocytaire (Sher *et al.*, 2012). La spécification neurale nécessite des signaux de différenciation, mais également des mécanismes intrinsèques. Chacune des voies de différenciation implique des voies de signalisation et des facteurs spécifiques. Alors que la différenciation neuronale verra une activation de la voie Wnt avec l'activation de facteurs comme la *Neurogenin*, la gliogénèse activera la voie Jak-STAT pour activer des gènes associés aux astrocytes comme *GFAP*. La répression semble pouvoir réguler temporellement et spatialement les différents acteurs moléculaires de la différenciation neurale. La différence de profil d'expression de chacune des PcGs au cours du développement et dans les régions cérébrales pourraient être expliquée par une spécialisation des protéines en fonctions des types cellulaires (Corley & Kroll, 2014).

### 4. L'hématopoïèse.

La protéine *Bmi1* et son activité dans le complexe PRC1 assure le maintien de l'identité et la capacité d'auto-renouvellement des cellules souches et des cellules pluripotentes de nombreuses lignées cellulaires. Dans les cellules souches hématopoïétiques (HSCs), *Bmi1* assure également le maintien de ces capacités, et l'augmentation de l'expression de *Bmi1* favorise la prolifération des cellules souches (Iwama *et al.*, 2004). L'action du complexe PRC1.4 permet de réprimer les gènes de différenciation des cellules souches hématopoïétiques. Les cellules peuvent ensuite entrer dans une voie de différenciation. La protéine *CBX7* est d'abord fortement présente dans les HSCs, et est impliquée dans

la mise sous silence des programmes de différenciation. La transition des HSCs vers les progéniteurs est accompagnée d'une diminution de CBX7, ce qui facilite l'intégration de son paralogue CBX8 dans le PRC1. La modification biochimique du complexe permet une réorientation du complexe sur les programmes géniques non nécessaires à la différenciation (Klauke *et al.*, 2013). Le complexe PRC1.1 facilite également la mise en place de la différenciation en réprimant des gènes de pluripotence comme *HoxA* (Ross *et al.*, 2012). Après la formation de progéniteurs, la répression Polycomb reste nécessaire jusqu'à la différenciation terminale. La protéine Bmi1 permet un recrutement du complexe au niveau du promoteur des gènes *Ebf1* et *Pax5*. La répression de ces gènes engage les cellules dans la lignée des lymphocytes B (Oguro *et al.*, 2010). L'activité de Ezh2 est également nécessaire pour la maturation complète des lymphocytes B (Hidalgo *et al.*, 2012). La sous-unité Suz12, au sein du complexe PRC2, permet également la maturation des lymphocytes B et T et un contrôle de l'expression des cytokines activatrices des lymphocytes comme *Il10* ou *Ifng* (Lee *et al.*, 2015). Le complexe PRC2 intervient aussi dans la détermination du lignage myéloïde. La protéine accessoire Mtf2 recrute le complexe PRC2 sur les promoteurs de gènes de la voie Wnt. La réduction de l'activation de cette voie réduit l'expression de gènes de différenciation des érythrocytes (Rothberg *et al.*, 2018). L'ensemble des complexes PRC1 et PRC2 intervient au cours de tout le processus de différenciation des cellules hématopoïétiques (Figure 13). Néanmoins, chaque complexe possède une spécificité d'action en fonction du lignage cellulaire et du stade de différenciation. De plus, les complexes PRC1 et PRC2 peuvent avoir une action inverse. Dans les cellules souches hématopoïétiques, le complexe PRC1 tend à favoriser le maintien des cellules souches tandis que le complexe PRC2 inhibe leur capacité d'auto-renouvellement par répression du gène *HoxC4* (Di Carlo *et al.*, 2018). De nombreuses questions demeurent sur le recrutement préférentiel d'un complexe ou d'un autre dans les cellules hématopoïétiques.



**Figure 13 : Régulation de l'hématopoïèse par les complexes Polycomb.** Les complexes PRC1 et PRC2 interviennent à tous les niveaux de la différenciation hématopoïétique. La composition biochimique des complexes varie en fonction du stade de différenciation, ce qui leur permet de réprimer spécifiquement les programmes géniques. HSC : cellule souche hématopoïétique ; MPP : progéniteur multipotent ; progéniteur lymphoïde multipotent ; CMP : progéniteur myéloïde commun ; CLP : progéniteur lymphoïde commun ; GMP : cellule progénitrice de granulocyte-macrophage ; MEP : progéniteur megakaryocyte-érythroïde ; CD4 : lymphocyte T CD4+ ; CD8 : lymphocyte T CD8+ ; B : lymphocyte B ; NK : cellule natural killer ; Mono : monocyte ; Gran : granulocyte ; Ery : érythrocyte ; Mega : megakaryocyte. Figure de Di Carlo *et al.*, 2018.

## 5. Le système digestif.

Le système digestif est composé d'un ensemble de tissu, aux fonctions propres. L'assemblage de ces tissus en organes résultent de processus de différenciation complexes et spatio-temporellement régulés. La répression Polycomb permet une partie de ces régulations. La voie Wnt active la répression Polycomb et induit l'induction du mésoderme et de l'endoderme au cours du développement précoce. La protéine Ezh2 est observée associée aux éléments régulateurs de *Pdx1* afin de restreindre la spécification pancréatique dans la région intestinale. Le pancréas est constitué de cellules endocrines et exocrines. La différenciation de ces cellules fait appel à des voies et l'expression de gènes spécifiques. Sous l'influence de signaux de la voie Notch, la triméthylation de la lysine 27 des histones H3 au niveau du locus *Ngn3* dirige les cellules progénitrices vers une différenciation en cellule exocrine. A l'inverse, le maintien de l'expression des gènes *Ngn3* et *Nkx* permet une orientation des progéniteurs



vers des cellules endocrines. (Dumasia et al., 2020). De la même façon le rôle de la répression Polycomb est critique dans le développement et le maintien des structures dans les autres organes du système digestif. L'épithélium intestinal est renouvelé très régulièrement grâce à la division asymétrique des cellules souches intestinales présentes au niveau des cryptes intestinales. Chez le modèle murin, la délétion de *Eed* conduit à la perte de la marque H3K27me3. La perte de cette marque répressive provoque un arrêt du cycle cellulaire de ces cellules souches et altère le renouvellement de la paroi intestinale (Lorzadeh *et al.*, 2021). La régulation de la différenciation et du maintien des cellules souches au sein des organes digestifs est donc prépondérante. La perturbation de cette régulation est observée dans de nombreuses pathologies humaines, et peut éventuellement favoriser l'apparition de certains cancers.

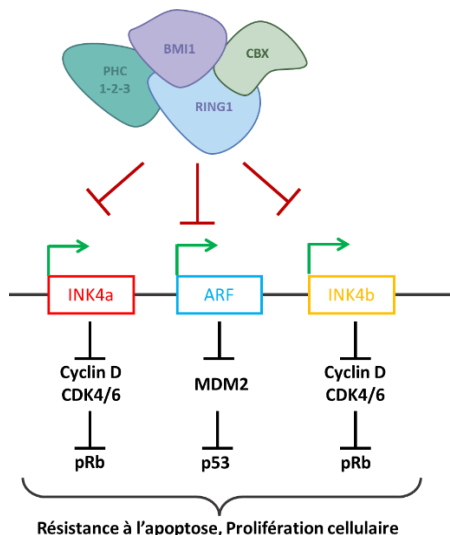
### CHAPITRE III : LA REPRESSION POLYCOMB DANS LE CANCER

---

La protéine BMI1 fut la première protéine Polycomb à être impliquée dans le cancer et découverte dans le cadre d'une coopération avec l'oncogène *myc* (van Lohuizen *et al.*, 1991). De par le rôle de régulation des grands processus biologiques tels que la prolifération, le contrôle de la différenciation cellulaire ou le maintien des cellules souches, une mutation ou une altération de l'expression des protéines PcGs peut se révéler dramatique. Les translocations MLL (Mixed lineage leukemia) sont à l'origine du développement de leucémies aiguës. Ces translocations créent des fusions avec de façon préférentielle les facteurs de transcriptions AF9 et ENL. La protéine de fusion MLL-AF9 possède un domaine CTD (C-terminal domain), qui est reconnu par CBX8. Cette interaction promeut la transformation des cellules ainsi que leur prolifération indépendamment de PRC1 (Tan *et al.*, 2011). Dans ce cadre, CBX8 a plutôt un rôle oncogénique. L'implication de CBX7 dans la cancérogénèse est dépendante du contexte cellulaire. Une diminution de l'expression de CBX7 a d'abord été observée dans les carcinomes thyroïdiens. Des IPs ont révélé que CBX7 favorisait l'expression de la E-cadhérine en se liant à un histone désacétylase. Dans ce contexte tumoral, CBX7 se comporte comme un suppresseur de tumeur et la diminution de CBX7 est facteur de mauvais pronostic (Federico *et al.*, 2009). En revanche au sein des lymphomes, *CBX7* est surexprimé et cette surexpression coïncide avec l'expression de *MYC*. La répression exercée par CBX7 favorise le transformation des lymphocytes B en cellules tumorales hautement agressives (Scott *et al.*, 2007). La protéine CBX7 résume à elle seule le caractère ambivalent des protéines du groupe Polycomb. Les protéines Polycomb ont donc un rôle ambivalent : elles peuvent avoir un caractère suppresseur de tumeur ou oncogénique.

## 1. Le locus INK4a-ARF-INK4b.

Ce locus comprend 3 gènes d'intérêt dans l'étude du cancer : le gène *INK4a* et le gène *INK4b* qui codent respectivement les deux inhibiteurs de cycline p16 et p15, et *ARF*. La protéine ARF est impliqué dans de nombreuses interactions, et ses fonctions comprennent entre autres la régulation du cycle cellulaire et l'induction de l'apoptose. Les protéines p16 et p15, en inhibant les cyclines, orientent les cellules vers la sénescence (Figure 14). Le contrôle transcriptionnel de ce locus est complexe et est assuré par de multiples composants, mais une partie de la régulation est effectuée par le complexe BMI1 et le PRC1 (Jacobs *et al.*, 1999 ; Gil & Peters, 2006). La perte d'expression des gènes du locus, et particulièrement *INK4A*, associée à d'autres mutations, est mise en cause dans de nombreux cancers hématologiques.



**Figure 14 : Rôle oncogénique du complexe PRC1 sur le locus INK4a-ARF-INK4b.** Lorsque le locus est transcriptionnellement actif, les protéines INK4a et INK4b se lie aux cyclines D. Cette interaction entraîne la phosphorylation de pRb, et maintient la cellule en phase S. La protéine ARF permet d'inhiber l'activité de MDM2, qui consiste en l'ubiquitination de p53 et sa dégradation. L'expression de ces gènes conduit à un arrêt du cycle cellulaire. Lorsque le locus est réprimé par PRC1, les cyclines activent la reprise du cycle cellulaire et la dégradation de p53 lève les freins à la prolifération cellulaire et prévient l'apoptose cellulaire.

Le complexe PRC2 peut également exercer une répression de ce locus par son activité de triméthylation de la lysine 27. La mutation *KRAS*<sup>G12V</sup>, est fréquemment observée dans les lésions néoplasiques pancréatiques. Alors que la sénescence cellulaire est promue par *INK4A*, la surexpression de *EZH2* montre une diminution de l'expression de ce locus. La perte de contrôle du cycle cellulaire favorise la croissance tumorale (Mallen-St. Clair *et al.*, 2012).

## 2. Les altérations des PcGs.

Parmi toutes les mutations de BCOR recensées dans la base de données TCGA cancer (<https://www.cbic.cbioportal.org>), la mutation N1459S est retrouvée de manière récurrente dans les lymphomes, les carcinomes de l'endomètre et les carcinomes pulmonaires. Cette mutation intervient sur le domaine ANK de BCOR, ce qui pourrait affecter la stabilité des interactions avec les autres protéines du complexe. Il existe également de nombreuses fusions entre BCOR et des protéines telles

que la cycline CCND3, ou le facteur MAML3 qui peuvent conduire à l'apparition de néoplasmes variés (Astolfi *et al.*, 2019). De la même façon, les autres protéines Polycomb sont retrouvées mutées dans un nombre non négligeable de cancers. En revanche, les liens de corrélation entre les mutations et l'initiation tumorale, la progression tumorale ou la survie n'ont pas encore été établis.

Les protéines sont impliquées dans un nombre non négligeable de cancers. La perte de fonction de CBX7 par exemple favorise le développement de tumeurs pulmonaires et hépatiques. Le KO de CBX4 semble plutôt inhiber la formation d'ostéosarcomes. Les exemples sont nombreux concernant les protéines CBX, mais leur rôle pro- ou anti-tumoral n'est pas toujours bien établi, et est très dépendant du contexte tumoral (van Wijnen *et al.*, 2020). La protéine EZH2 est particulièrement étudiée pour son rôle dans le cancer. L'altération de l'expression de *EZH2* autant que sa surexpression sont à l'origine ou favorise la croissance de nombreux cancers (Di Carlo *et al.*, 2019). Des mutations récurrentes affectent la tyrosine 641 de EZH2 dans des lymphomes (Morin *et al.*, 2011 ; Caganova *et al.*, 2013). La tyrosine 641 est située au niveau du domaine catalytique de EZH2. La mutation conduit dans les lymphomes à une hyperactivité de EZH2 et une augmentation de la déposition de la marque H3K27me3 (Sneeringer *et al.*, 2010). Des mutations des autres sous-unités du PRC2 conduisent à la tumorigenèse, mais forment d'autres types de cancers. Les mutations affectant la stabilité des domaines WD ou la cage aromatique de EED favorisent le développement de leucémies (Ikeda *et al.*, 2016 ; Ueda *et al.*, 2016). L'inactivation de PRC2 par la perte de fonction de EED ou SUZ12 est à l'origine de plus de 70% des tumeurs de la gaine des nerfs périphériques (Lee *et al.*, 2014). Des mutations de *EZH2* ne sont en revanche jamais observées. De la même façon que chaque complexe Polycomb possède une spécificité d'action au cours du développement, le développement tumoral est dépendant de la sous-unité mutée. Enfin il existe des mutations qui ne touchent pas directement les PcGs, mais impactent leur activité. Il existe plusieurs variants d'histone, qui possèdent des fonctions propres. Une mutation récurrente du variant H3.3 a été identifiée dans les gliomes diffus intrinsèques pontiques (DIPGs). C'est une mutation faux-sens qui remplace la lysine 27 par une méthionine. La mutation inhibe l'activité méthyltransférase de PRC2 et peut même retenir EZH2, ce qui conduit à une baisse globale de H3K27me3, y compris sur les variants H3 non mutés (Lewis *et al.*, 2013 ; Jiao & Liu, 2015). De façon intéressante, alors que les DIPGs aborant la mutation K27M sont associés à une survie inférieure à un an chez l'enfant, la même mutation chez l'adulte est associée avec une survie plus longue, voire meilleur pronostic (Vuong *et al.* 2022 ; Hu *et al.*, 2022). L'effet de la mutation du variant d'histone H3.3 semble très dépendant du contexte cellulaire et du niveau de maturation cérébrale du patient.

### 3. Les rôles de la régulation Polycomb dans les hallmarks du cancer.

Les hallmarks du cancer ou marques distinctives du cancer ont été édictées par Hanahan et Weinberg en 2000. Elles comprennent actuellement les capacités de réplication illimitée des cellules, la résistance à l'apoptose et au système immunitaire, l'autosuffisance aux facteurs de croissance et l'insensibilité aux inhibiteurs de croissance, une inflammation pro-tumorale, une instabilité génomique conséquente, l'induction de l'angiogenèse, la formation de métastase ainsi que la modification du métabolisme cellulaire (Hanahan & Weinberg, 2011 ; Flavahan *et al.*, 2017). Le rôle de PRC1 dans la réplication cellulaire et dans la résistance à l'apoptose a été mis en lumière avec la régulation du locus *INK4a-ARF-INK4b*. Dans le cas des gliomes pédiatriques H3.3K27M, la diminution globale de la marque H3K27me3 n'est pas observée au niveau du locus *INK4a-ARF-INK4b*. Malgré l'inhibition du PRC2 par la mutation, l'activité de PRC2 reste nécessaire à la progression tumorale (Mohammad *et al.*, 2017). La capacité métastatique dépend de l'invasion et de la migration des cellules cancéreuses. Dans les cancers prostatiques, l'altération des protéines du complexe PRC1 conduit à l'expression de plusieurs gènes, dont *CCL2*. Cette cytokine attire les macrophages et participe à la création d'un milieu inflammatoire favorisant la migration et l'invasion des cellules cancéreuses (Su *et al.*, 2019). L'intégration de CBX4 dans le complexe PRC1 permet l'ubiquitination du facteur HIF $\alpha$ . Cette modification post-traductionnelle augmente son activité. Ce facteur de transcription induit l'angiogenèse (Li *et al.*, 2014). Enfin, des modifications post-traductionnelles des sous-unités Polycomb peuvent modifier l'activité du complexe. La phosphorylation de la tyrosine 311 de EZH2 diminue l'activité totale de PRC2. Au cours du développement tumoral, les cellules situées au centre de la tumeur se retrouvent en hypoxie et en privation de nutriments. L'AMPK peut alors phosphoryler EZH2, et induire l'expression de gènes impliqués dans le métabolisme (Wan *et al.*, 2018). Que ce soit par une altération des protéines Polycomb, ou un recrutement ectopique des complexes par d'autres acteurs, les complexes PRC1 et PRC2 agissent sur chacune des particularités du cancer, ce qui en fait des cibles thérapeutiques intéressantes (Zhao & Wu, 2021).

### 4. Les protéines Polycomb comme marqueur prédictif.

Alors que l'activité des protéines Polycomb favorise ou inhibe le développement tumoral en fonction de contexte cellulaire, de nombreuses études ont essayé d'établir des liens entre les altérations de protéines, leur sur- ou sous-expression et la signification clinique selon les types des cancers. Là encore, les résultats sont très dépendants de chaque protéine, et pour une même protéine, du type de cancer (Crea *et al.*, 2012). Dans les cancers du sein, une analyse du plasma sanguin des patients a

révélé que la présence d'ARNm de *BMI1* était associée à un mauvais pronostic, et particulièrement dans les stades avancés (Silva *et al.*, 2007). De la même façon, une surexpression de *EZH2* est associée à un pronostic sombre dans les cancers du sein (Wang *et al.* 2015). Le rôle des protéines de la famille CBX a également été étudié dans les cancers du sein. L'expression des CBX est corrélée à un sous-type de cancer mammaire : l'expression de *CBX2* est associée au type Basal-like tandis que *CBX4* est associée aux sous-types luminal A et B. L'expression des protéines pourraient être utilisées pour la prédiction de la réponse aux traitements, avec une chimiorésistance des cellules exprimant *CBX2* et une chimiosensibilité des cellules exprimant *CBX4* (Liang *et al.*, 2017). Les PcGs, et en particulier les paralogues de la famille CBX pourraient jouer un rôle capital dans la détermination d'un pronostic et de la réponse aux traitements des patients en fonction du type de cancer. Une meilleure connaissance des fonctions de chaque protéine dans les tissus sains et cancers est nécessaire, afin de pouvoir lutter plus efficacement contre ces processus tumoraux.



### 1. Le poisson zèbre

#### 1.1 L'histoire

Le poisson zèbre, *Danio rerio*, a été observé la première fois dans les eaux du Ganges en 1822 par Hamilton (Hamilton-Buchanan, 1822). Dans son milieu naturel, le poisson zèbre est principalement présent au nord de l'Inde et au Bangladesh, autour des rivières du Ganges et de Brahmaputra (Spence *et al.* 2008). Plusieurs rapports d'observations au fil des années ont identifié des populations au Népal et au Pakistan (Figure 15). Néanmoins, ces observations n'ont pas été réitérées ces dernières années. L'une des hypothèses exposées fait état des impacts néfastes de la croissance démographique exponentielle dans cette région d'Asie qui limiterait les habitats viables pour les poissons (Engeszer *et al.*, 2007).

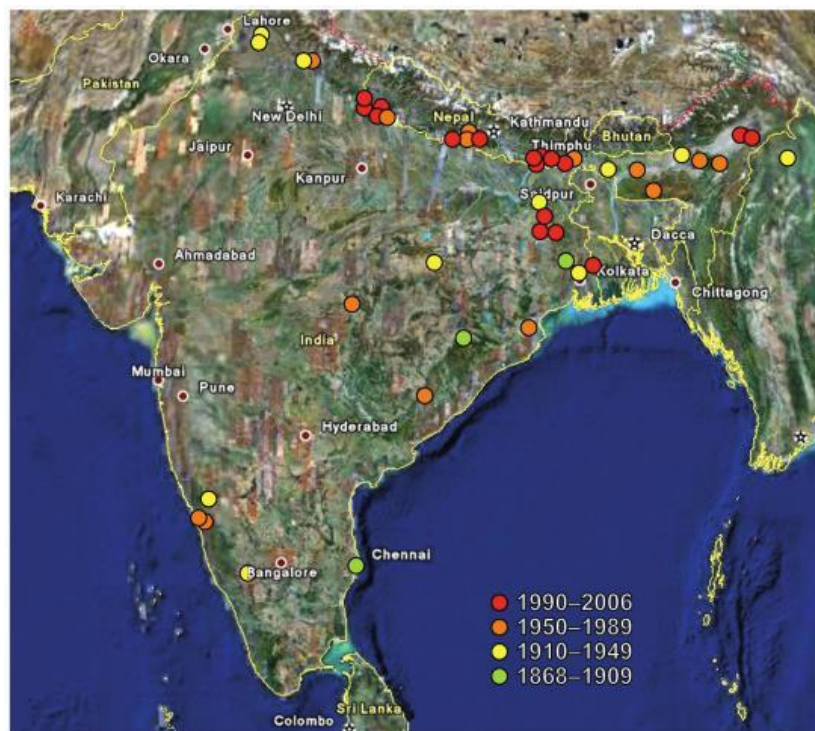


Figure 15 : Répartition géographique des captures de *Danio rerio* en milieu sauvage au cours du temps. Figure de Engeszer *et al.*, 2007.

D'abord principalement utilisé comme espèce d'ornement en aquariophilie, le poisson zèbre est vite devenu un animal d'intérêt en biologie. Dès les années 1930, Creaser (1934) décrivait les techniques d'élevage des poissons zèbres pour récupérer les œufs. Le poisson zèbre est ovipare. Les œufs sont transparents et autorisent une observation des stades les plus précoces du développement.



Le poisson zèbre est devenu un modèle de premier plan grâce aux travaux de Streisinger, et ses recherches pour obtenir un modèle vertébré modifiable facilement génétiquement. De nombreuses mutations affectant le développement sont récessives et ainsi difficilement observables. L'expression d'un allèle récessif dans la descendance nécessite un croisement entre deux parents hétérozygotes. La fécondation des poissons zèbres est externe et il est possible de récupérer les gamètes pour modifier leur ploïdie et créer des mutations récessives observables (Grunwald & Eisen, 2002). L'équipe de Streisinger a créé un protocole de fécondation d'un ovule par un spermatozoïde dont l'information génétique a été inactivée par irradiation aux UV, et permet pour la première fois la production d'un clone homozygote chez le poisson zèbre (Streisinger *et al.* 1981). Les descriptions du système nerveux au stade embryonnaire et la ressemblance phénotypique et organisationnelle des types cellulaires neuraux avec les mammifères ont entériné les embryons de poissons zèbres comme modèle d'embryologie (Kimmel, 1982 ; 1989 ; 1993 ; Kimmel *et al.*, 1990 ; Kimmel *et al.*, 1995). Les poissons zèbres suivent les mêmes processus de développement que les autres modèles animaux connus tels que le xénope ou l'embryon de poulet. De plus, les comparaisons des gènes *Hox* entre le poisson et la souris montrent des homologies de séquence importante. A l'époque, le modèle drosophile était principalement utilisé pour la réalisation d'études génétiques fonctionnelles, et les mutants générés sont nombreux. En revanche, c'est sur le poisson zèbre que le premier projet de mutation à large échelle a été réalisé, à Tübingen en collaboration avec une équipe de Boston (Haffter *et al.*, 1996 ; Driever *et al.*, 1996). Le travail conduit sur plus de 4000 mutations a permis la découverte de plusieurs centaines de gènes impliqués dans le développement. Enfin, avec l'avènement des nouvelles techniques de séquençages à haut débit, l'institut Sanger accueille le projet « Zebrafish genome project » pour le séquençage du génome entier du poisson zèbre. Le séquençage a été réalisé sur la souche Tübingen (TU), sur laquelle a été réalisé le projet de mutation à grande échelle. Les comparaisons entre les génomes de référence révèlent que près de 70% des gènes humains ont au moins un orthologue chez le poisson (Howe *et al.*, 2013). La première version de l'assemblage du génome a été distribuée sous le nom Zv9. La dernière version de l'assemblage GRC11 est disponible sur le site de l'institut Sanger ([www.sanger.ac.uk](http://www.sanger.ac.uk)).

### 1.2 La taxonomie

Le poisson zèbre, *Danio rerio*, est un vertébré de la classe des poissons osseux. Il appartient à l'ordre des Cypriniformes. Cet ordre est le plus grand ordre de vertébrés et rassemble plus de 2500 espèces (Richards, 2011). Le poisson zèbre a été identifié au début du 19<sup>ème</sup> siècle avec une dizaine d'autres espèces *Danio* (Hamilton, 1822). Ces poissons ont d'abord été regroupés au sein du sous-genre *Brachydanio* (Weber & de Beaufort, 1916). En l'absence de critères différenciants permettant leur

classification, les genres *Danio* et *Brachydanio* ont finalement été regroupés sous un genre unique, *Danio* (Barman, 1991). L'étude des ADNs mitochondriaux 16S et 12S a confirmé au niveau moléculaire le genre monophylétique *Danio* (Meyer *et al.*, 1995). L'association de critères morphologiques et moléculaires ont permis la distinction de deux sous-clades : les *Danio* à corps large ou élancé. Ces deux sous-clades sont devenus deux genres distincts, *Devario* pour les espèces dites « à corps large » et *Danio* pour les espèces élancées (*Danio rerio* inclus). Les seuls critères morphologiques ne sont pas suffisants pour distinguer les espèces de chaque genre, avec les nombreuses variabilités morphologiques intraspécifiques. Néanmoins, les distinctions sont strictes au niveau écologiques et moléculaires (Spence *et al.*, 2008). Les espèces du genre *Devario* par exemple seront observées principalement au niveau d'eau courante, tandis que les espèces du genre *Danio* préfèrent les eaux calmes (Fang, 2003).

**Table 3 : Classification phylogénétique de *Danio rerio*.** Cette classification s'appuie sur la mise à jour de la classification des poissons osseux de 2017 (Betancur *et al.*, 2017).

Règne	<b>Metazoa (Animalia)</b>
Super-embranchement	<b>Deuterostomia</b>
Embranchement	<b>Chordata</b>
Sous-embranchement	<b>Vertebrata</b>
Infra-embranchement	<b>Gnathostomata</b>
Méga-classe	<b>Osteichthyes (Euteleostome)</b>
Super-Classe	<b>Actinopterygii</b>
Classe	<b>Actinopteri</b>
Sous-classe	<b>Neopterygii</b>
Infra-classe	<b>Teleostei</b>
Super-cohorte	<b>Clupeocephala</b>
Cohorte	<b>Otomorpha</b>
Sous-cohorte	<b>Ostariophysii</b>
Super-ordre	<b>Cypriniphysae</b>
Ordre	<b>Cypriniformes</b>
Sous-ordre	<b>Cyprinoidae</b>
Famille	<b>Danionidae</b>
Genre	<b><i>Danio</i></b>
Espèce	<b><i>Danio rerio</i></b>

### 1.3 L'habitat

L'habitat sauvage du poisson zèbre est composé d'eau douce, dans des zones peu profondes, avec peu ou pas de courants. De ce fait, on retrouve majoritairement les populations de poissons zèbres dans des ruisseaux, des mares et des rizières. Les plans d'eau stagnante représentent un environnement sûr pour les poissons zèbres puisqu'ils sont généralement dénués de prédateurs. Ils sont observés en rivières uniquement lorsque le courant est très faible, et restent proche de la berge (Engeszer *et al.*,

2007). La présence de végétation aquatique est notée dans la grande majorité des lieux de vie du poisson zèbre (Spence *et al.*, 2008). Le climat sud asiatique est concerné par des épisodes de moussons et de fortes sécheresses. Si les moussons peuvent augmenter la surface des biotopes par inondation, les phénomènes climatiques peuvent également faire varier les paramètres physico-chimiques de l'eau. Toutefois, les analyses montrent une préférence des poissons zèbres pour une eau à pH légèrement alcalin et de faible turbidité (Lawrence *et al.*, 2007). Les températures varient énormément, de 6°C en hiver à plus de 38°C l'été.

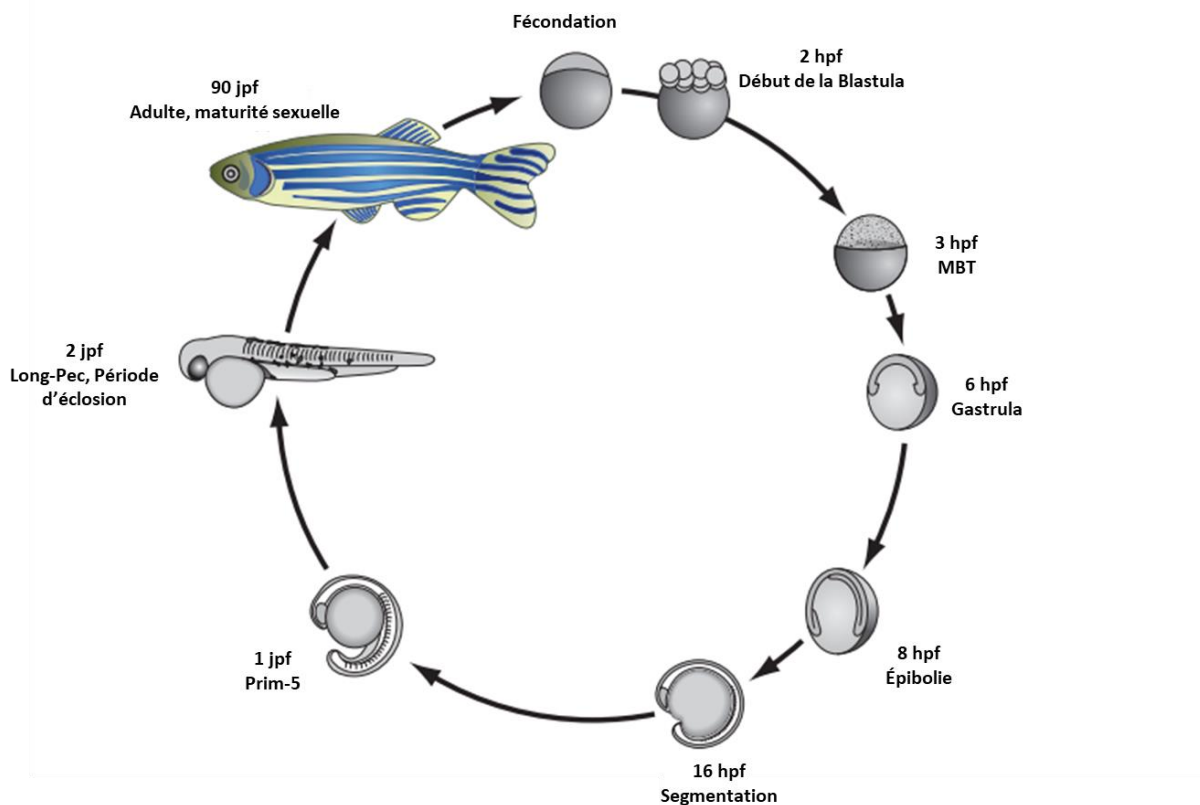
### 1.4 L'alimentation

Le poisson zèbre est un animal omnivore et utilise de nombreuses ressources dans son environnement. Son alimentation est principalement composée de zooplancton. La croissance des zooplanctons est favorisée par l'ajout d'engrais chimiques dans les rizières, ce qui participe à l'attrait du poisson zèbre pour ce milieu (Spence *et al.*, 2007). Avec les zooplanctons, la diète est composée d'insectes, d'algues filamenteuses, de spores, d'œufs d'invertébrés, de petits crustacés, de détritiques animaux et végétaux (Dutta, 1993 ; McClure *et al.*, 2006). L'analyse du contenu du bol alimentaire a révélé une majorité de zooplanctons, ce qui indique que la nourriture est capturée dans l'ensemble de la colonne d'eau. La présence de petits insectes volants et terrestres indique une capture également à la surface de l'eau. La composition du bol alimentaire des poissons est variable en fonction de la période de l'année, l'accès à la nourriture apparaît dépendant de la saisonnalité des apports. Ces apports influencent la croissance des juvéniles (Spencer *et al.*, 2007).

### 1.5 Cycle de vie

La durée de vie moyenne d'un poisson zèbre est située autour de 2 ans. En captivité, la durée de vie semble plus longue, avec un record de longévité de près de 6 ans (Gerhard *et al.*, 2002). Néanmoins, les individus les plus âgés présentent des phénotypes particuliers associés à la sénescence. La dégénération des muscles squelettiques provoque une déformation de la courbure de la colonne vertébrale. Ce phénotype n'a jamais été observé sur les poissons zèbres sauvages (Spence *et al.*, 2007). Le temps de génération, qui sépare un embryon d'un individu fertile, est de 3 mois chez le poisson zèbre (Figure 16). Une femelle peut pondre jusqu'à 200 œufs par semaine en présence d'un mâle. La libération d'hormones sexuelles stéroïdiennes contenues dans les testicules des poissons mâles induit l'ovulation et la ponte des œufs (Eaton & Farley, 1974). Les œufs sont libérés dans le milieu avec un diamètre large en comparaison des autres espèces du genre *Danio*, de 0.7mm. Les poissons zèbres ne présentent aucun comportement parental. Après la ponte, les œufs sont activés au contact de l'eau et entament les premiers processus de développement, même en absence de sperme. Les œufs non

fertilisés ne parviennent pas en revanche à effectuer les premières divisions cellulaires (Lee *et al.*, 1999). Le développement est très rapide. La première division cellulaire a lieu environ 45 minutes après la fécondation. À partir de 3 heures post-fécondation (hpf), la transition midblastula débute avec l'activation de la transcription zygotique (Kane & Kimmel, 1993). Le cœur des embryons commencent à battre dès 24 hpf et à 48 hpf, tous les organes sont mis en place et fonctionnels. L'acquisition de la nage permet l'éclosion des embryons (Kimmel *et al.*, 1995). La croissance des larves puis des juvéniles est très rapide et permet d'identifier visuellement leur stade de développement (Parichy *et al.*, 2009).



**Figure 16 : Représentation schématique du cycle de vie du poisson zèbre.** Les principales divisions cellulaires et mouvements morphogénétiques sont observables en seulement quelques heures. Le temps de génération très court du poisson zèbre en fait un modèle vertébré précieux en laboratoire. hpf : heures post-fécondation ; MBT : Transition midblastula ; Prim-5 : position 5 du primordium de la ligne latérale postérieure ; jpf : jour post-fécondation ; Long-Pec : Long pectoral fins stage. Figure adaptée de Richards, 2011.

## 1.6 Comportements du poisson zèbre

Le poisson zèbre est un animal social, qui de façon innée a tendance à former des bancs d'individus, dès l'éclosion (Engeszer *et al.*, 2007). Les individus de la même espèce et même des individus proches génétiquement sont reconnus par le biais de molécules olfactives (Gerlach & Lysiak, 2006). Les facteurs qui influencent la formation de banc sont multiples : les espèces en présence, les liens de parenté, l'activité du banc, les sexes des poissons, le nombre de poissons, les paramètres physico-chimiques de

l'eau et même l'état nutritionnel (Spence *et al.*, 2008). Ce comportement social à un rôle de protection des individus. Plus le nombre d'individus est grand, plus les chances pour qu'un prédateur soit repéré par au moins un membre du banc sont grandes (Pitcher & Parrish, 1993). Néanmoins au fur et à mesure des générations élevées en captivité, l'absence de prédateurs poussent les poissons à adopter un comportement exploratoire plus prononcé (Spence *et al.*, 2008). Enfin, plusieurs études ont révélé que les poissons zèbres sont capables d'apprentissage et de mémorisation. Des expériences de conditionnements classiques, décrites par Pavlov comme l'association d'un stimulus conditionné à un stimulus renforçant, ont été plusieurs fois réalisées chez les poissons zèbres adultes et larvaires (Roberts *et al.*, 2013). De la même façon, le conditionnement opérant, qui permet l'association d'un comportement à une conséquence sur l'animal, a été réalisé. Les protocoles utilisés sur les poissons sont généralement des protocoles de conditionnement d'évitement face à un stimulus. Grâce à ces études comportementales, des équipes ont mis en évidence que les structures cérébrales et les voies de signalisations impliquées dans l'apprentissage sont identiques chez les poissons et les mammifères (Schumake *et al.*, 2010 ; Roberts *et al.*, 2013). La palette de comportements déployée par le poisson zèbre en font également un modèle de choix pour vérifier les modifications comportementales induites par une mutation ou une molécule.

### 1.7 Élevage du poisson zèbre en laboratoire

Les poissons sont élevés dans les conditions standards en suivant les recommandations européennes. Le pH de l'eau est maintenu entre 7 et 8. Ce rang de pH facilite l'oxydation des déchets azotés par les bactéries présentes dans les biofilms des aquariums (Masser *et al.*, 1999). Le poisson zèbre est un animal eurytherme, qui supporte bien les variations de température. En revanche, la température peut faire varier les processus biologiques et biochimiques. En laboratoire, la température de l'eau est donc maintenue de manière constante autour de 28.5°C, selon les recommandations de Westerfield (1995). Ces paramètres sont contrôlés quotidiennement. Les individus adultes sont nourris deux fois par jour, tandis que les larves sont nourries trois fois par jour. La fréquence de prise alimentaire peut faire varier le métabolisme et le comportement des poissons. Si ce fait peut expliquer les variations de résultats obtenus entre laboratoires, c'est aussi dans cette optique que beaucoup de personnes tentent dorénavant d'uniformiser les protocoles de soin des poissons zèbres (Dametto *et al.*, 2018). Chaque jour, 10% du volume d'eau total des aquariums est remplacé. Ce renouvellement, ajouté au flux continu de chaque bac participe à l'oxygénation de l'eau et l'élimination des déchets, ainsi qu'à l'enrichissement des poissons. Ces dernières années, l'exigence du soin apporté aux animaux a augmenté. L'utilisation d'animaux dans les laboratoires est soumise à des règles strictes et doit être associée à une attention particulière à leur bien-être. Les poissons sont automatiquement anesthésiés

lors de manipulations invasives. Des proies vivantes sont ajoutées à leur alimentation pour stimuler leurs comportements de prédation. Les poissons zèbres sont élevés en groupe afin de respecter le caractère social de cette espèce. Cette forme d'enrichissement permet une meilleure récupération après les prélèvements d'un morceau de nageoire caudale dans le cadre de génotypage (White *et al.*, 2017). Les aquariums sont aménagés pour reproduire leur habitat naturel de la végétation en plastique a été installée. Ces enrichissements permettent de réduire le stress des individus et améliorer leur reproduction (Marcon *et al.*, 2018 ; Wafer *et al.*, 2016). Le fond des aquariums peut également être amélioré. Les résultats expérimentaux sont parfois contradictoires concernant les préférences pour le sable ou le gravier. Néanmoins les images de gravier placées en dessous des aquariums sont suffisantes pour améliorer le bien-être des poissons (Stevens *et al.*, 2021).

## 2. Le développement du poisson zèbre

En seulement 48 hpf, un œuf fécondé devient un embryon prêt à éclore. Les principales grandes étapes du développement précoce, que ce soit la gastrulation, la neurulation ou la segmentation se déroulent en quelques heures. Chez le poisson zèbre, malgré une fécondation *in vitro* simultanée des ovules, les embryons se développent de manière asynchrone (Westerfield, 1994). L'identification horaire du stade de développement n'est de ce fait pas la méthode la plus adaptée. Des efforts de caractérisation des stades de développement ont permis de définir des critères morphologiques visibles et identifiables avec leur période d'apparition moyenne (Kimmel *et al.*, 1995). La connaissance approfondie de ces stades de développement permet d'identifier chez des embryons mutants des altérations développementales. Le développement externe et rapide du poisson zèbre en a fait un modèle de choix dans l'étude du développement. Le développement larvaire dépend uniquement des réserves vitellines jusqu'à 5 jpf, et facilite l'étude des organes digestifs. De nombreux gènes de développement précoces sont équivalents chez les mammifères et les téléostéens. La famille des gènes *Hox* est largement représentée chez le poisson zèbre et dirige l'établissement de l'axe antéro-postérieur. Au cours de l'évolution, plusieurs événements de duplication du génome entier ont eu lieu et sont proposés comme porteurs innovations évolutives (Ohno, 1970). Une duplication entière du génome a eu lieu il y a environ 350 millions d'années, et est à l'origine de la spécification de la branche évolutive des téléostéens (Glasauer & Neuhauss, 2014). La majorité des gènes dupliqués sont perdus au fur et à mesure des générations. Néanmoins, les mutations de certains gènes confèrent un avantage évolutif et sont ainsi conservés. Le poisson zèbre, comme une partie des téléostéens, possède 7 clusters de gènes *Hox* tandis que les tétrapodes n'en ont que 4 (Amores *et al.*, 1998). Néanmoins, le degré d'homologie entre les *Hox* humains et ceux du poisson zèbres est très important (Corredor-Adámez *et*

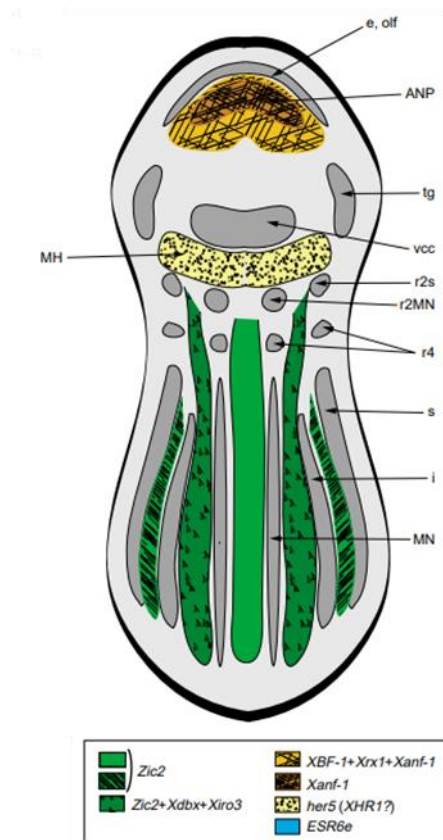
*al.*, 2005). Le développement du poisson zèbre, et notamment le développement d'embryons mutants pour des gènes de développement est intéressant pour la modélisation de pathologie du développement chez l'homme. De plus, la compréhension de chaque étape menant à la formation d'un organe et des facteurs impliqués permet une meilleure compréhension des phénotypes observés plus tardivement.

## 2.1 Le système nerveux central

La formation du système nerveux chez le poisson zèbre est similaire à celle des mammifères et débute très précocement dans le développement. Dès 6 hpf, la gastrulation débute et le système nerveux central est mis en place (Woo & Fraser, 1995). Les cellules du tissu mésodermique sécrètent des facteurs qui permettent d'encadrer la première grande étape du développement neural, l'induction neurale. Les facteurs Bmp, Wnt et la famille des Fgf déterminent les cellules neurales en fonction de leur position (Schmidt *et al.*, 2013). Des programmes géniques sont également activés, et notamment la famille de gènes *soxB1*, déterminent une partie du lignage du neuroectoderme (Sasai *et al.*, 1998). Le neuroectoderme forme un épithélium pseudostratifié appelé plaque neurale. Après 10 hpf, au début de la somitogenèse, la plaque neurale suit un mouvement de convergence et de pluristratification pour former la quille neurale et la tige neurale en position dorsale (Papan & Campos-Ortega, 1994). Les cellules sont polarisées par la présence de protéines d'adhérence en basolatérale et se divisent de manière symétrique, ce qui permet d'augmenter le nombre de progéniteurs neuraux tout en formant la lumière du tube neural (Clarke J, 2009). La formation d'une quille neurale, appelée aussi seconde neurulation, chez le poisson zèbre diffère de la neurulation des autres vertébrés. Chez les mammifères, les cellules de la plaque neurale suivent des mouvements morphogénétiques qui forment directement le tube neural, tandis qu'une étape intermédiaire est ajoutée chez le poisson zèbre. Néanmoins, les facteurs requis pendant la neurulation ainsi que la structure du tube neural sont similaires (Schmidt *et al.*, 2013). Après la formation du tube neural, les progéniteurs de l'épithélium neural débutent des divisions asymétriques. Le devenir des cellules filles dépendra de l'arsenal de domaines membranaires, les cellules présentant un phénotype « apical » deviendront des précurseurs neuraux tandis que les cellules basales renforcent le pool de progéniteurs. La détermination des cellules filles est également assurée par une activation différentielle de la voie de signalisation Notch (Dong *et al.*, 2012). Les cellules neurales sont aussi spécifiées selon l'axe antéro-postérieur par plusieurs signaux de type Wnt, Fgf et l'acide rétinoïque.

La neurogénèse est initiée à la fin de la gastrulation par l'activation de gènes pro-neurax, tels que *neurog1* ou *asc1*. L'expression de ces gènes modulée par les voies Notch et Hedgehog n'est pas activée dans tout le neuroderme et délimite des territoires pro-neurax (Blader *et al.*, 1997). Les territoires

pro-neuraux sont encadrés par des territoires non-neurogéniques exprimant *soxB1*. Parmi les cibles du facteur *soxB1* sont observés les gènes de la famille *her/hes*. Ces gènes répriment l'expression de gènes pro-neuraux tels que *neurog1*. Leur expression définit spatialement les territoires pro- et non-neurogéniques. La plaque neurale, puis le tube neural sont régionalisés et expriment des facteurs différents. Dès la gastrulation, une frontière entre le cerveau moyen et postérieur (MHB) est établie. C'est une région d'intérêt puisqu'elle permet l'organisation et la détermination des cellules neurales dans le cerveau moyen et postérieure (Wurst & Bally-Cuif, 2001). La MHB est formée très précocement et est composée de progéniteurs. Ces progéniteurs expriment fortement les gènes *her3*, *her5*, *her9* et *her11* et participent à l'expansion du réservoir de progéniteurs. Ensuite, l'activation de la voie Notch en concomitance avec l'expression de gènes *her* oriente les progéniteurs dans la voie de différenciation neurale (Stigloher *et al.*, 2008). En avant de la MHB, la plaque neurale antérieure formera le cerveau antérieur et le cerveau moyen. Les facteurs *xanf-1*, *xbf-1* et *xrx1* permettent d'inhiber la neurogenèse dans cette région neurale (Bally-Cuif & Hammerschmidt, 2003). En arrière de la MHB se situe la corde. Cette partie du système nerveux central en développement est caractérisée par des domaines pro-neuraux longitudinaux qui expriment des facteurs de transcription pro-neuraux. Ces domaines sont espacés par des domaines exprimant des facteurs inhibiteurs de la différenciation neurale comme *Zic2*, *Xdbx*, *Xiro3* (Figure 17) (Bally-Cuif & Hammerschmidt, 2003).



**Figure 17 : Représentation schématique du pattern de neurogenèse au stade neural plate.** Les régions de différenciation neurale sont séparées par des régions d'inhibition de la neurogenèse. Les inhibiteurs sont différents en fonction du positionnement sur la plaque neurale. MH : domaine cerveau moyen/postérieur ; MN : motoneurone spinal ; n : notochorde ; r4 : neurone moteur et sensoriel du rhombomère 2 ; s : neurone sensoriel spinal ; tg : placode trigéminal ; vcc : cluster ventro-caudal. Figure de Bally-Cuif & Hammerschmidt, 2003).



Cette configuration en domaines cellulaires permet un contrôle spatio-temporel du développement neural. D'une part, la différenciation des premiers précurseurs neuronaux est initiée avec la mise en place des premiers neurones sensoriels, et des inter- et motoneurons dans les régions dédiées. D'autre part, l'inhibition de la neurogenèse maintient une quantité importante, à proximité, de progéniteurs indifférenciés qui pourront entrer dans une voie de différenciation au cours du développement.

Après s'être répliqués et avoir été spécifiés, les progéniteurs entrent dans une voie de différenciation. Des cellules neuronales sont très vite différenciées dans le développement et peuvent être fonctionnelles à 18 hpf. La mise en place de structures neurales fonctionnelles précocement permet l'apparition des premiers mouvements et réflexes dès 2 jpf et permet l'éclosion (Chapouton, 2010).

Les progéniteurs neuraux pluripotents peuvent se différencier en plusieurs types cellulaires, des neurones jusqu'aux cellules gliales. Là encore, des profils d'expression précis tant au niveau spatial que temporel, vont façonner le destin de chaque cellule. Dès le stade gastrula, Shh induit l'expression de *neurog1* dans quelques progéniteurs, qui seront principalement dirigés vers la voie neuronale. Au fur et à mesure du développement, l'expression de *neurog1* est réduite et d'autres facteurs comme *neuroD* sont alors exprimés. L'expression séquentielle de facteurs spécifiques à la lignée neuronale fait progresser les cellules neurales dans le processus de différenciation (Blader *et al.*, 1997 ; Korzh *et al.*, 1998). La famille des facteurs neurogéniques participe à la régulation dynamique de la différenciation des progéniteurs neuronaux. Les gènes *deltaA-D* sont des ligands du récepteur Notch et participent à l'inhibition latérale. Lorsqu'une cellule entre en différenciation, elle exprime ces ligands à sa surface. Les ligands Delta sont fixés par le récepteur Notch de la cellule voisine. L'activation de la voie Notch active la transcription des facteurs Her/Hes et inhibe l'expression des gènes pro-neuronaux (Takke *et al.*, 1999). Cette inhibition latérale permet le maintien des progéniteurs à proximité de cellules différenciées et promeut un établissement des types cellulaires progressifs. Le miR-9, un micro ARN non-codant, maintient un statut ambivalent des progéniteurs neuraux. Parmi ses cibles de répression sont observés des gènes de la famille *her* ainsi que *elav3*. Lorsque miR-9 cible les gènes *her*, l'expression de *elav3* facilite la sortie de cycle cellulaire et la différenciation cellulaire, en fonction du stade de développement (Coolen *et al.*, 2012). En effet, deux stades particuliers de neurogenèse sont décrits chez le poisson zèbre, un premier avant 24 hpf, et un second post-embryonnaire (Mueller & Wullmann, 2003). Le premier stade de neurogenèse est caractérisé par une augmentation du nombre de progéniteurs avec un nombre de types neuronaux faibles. Après une nouvelle phase de prolifération des progéniteurs, ce nombre augmente fortement lors du second stade de neurogenèse, avec plus de 53 sous-types neuronaux détectés à 5 jpf (Raj *et al.*, 2020). Les facteurs exprimés sont très variables,

et leur combinaison définit des clusters correspondants à des sous-types neuronaux et gliaux avec des particularités morphologiques et fonctionnelles.

Les progéniteurs neuraux sont à l'origine des cellules gliales et neuronales. Les progéniteurs neuraux présents dans la zone ventriculaire du système nerveux central se différencient en motoneurons ou en oligodendrocytes, sous la direction de la signalisation hedgehog. Lors du développement et de la différenciation, certains progéniteurs migrent. Le facteur Shh est fortement exprimé dans les régions ventrales, ce qui établit un gradient de concentration le long de l'axe dorso-ventral. Ce gradient induit des réponses différentes selon la position des progéniteurs (Ericson *et al.*, 1997). Les progéniteurs expriment *olig2* et *nkx2.2*. Sous l'influence du gradient de Shh et de facteurs complémentaires, certains vont exprimer des gènes pro-neurax, réprimer l'expression de *olig2* et induire leur différenciation en motoneurons et interneurons. Lorsque les gènes pro-neurax sont réprimés, les progéniteurs se différencient en précurseurs d'oligodendrocytes. Ces précurseurs migrent le long des faisceaux axonaux, et se différencient en oligodendrocytes (Lyons & Talbot, 2014).

Les astrocytes, présents dans le système nerveux des mammifères, sont absents chez les poissons zèbres. Les cellules gliales radiaires sont des cellules progénitrices des types neuronaux et gliaux. Une sous-population de ces cellules radiaires gliales se spécialisent chez le poisson zèbre et remplissent les fonctions des astrocytes des mammifères, et participent par exemple à la formation de la barrière hémato-encéphalique. Les mécanismes qui régissent le développement de cette population cellulaire sont encore méconnus ((Lyons & Talbot, 2014).

La neurogenèse est dépendante de l'inhibition successive de gènes de pluripotence, pro-neurax ou pro-gliaux, et est contrôlée à la fois par des facteurs intrinsèques et extrinsèques. La répression épigénétique prend alors un rôle clé dans la mise sous silence de programmes géniques de manière dynamique (Seritrakul & Gross, 2019 ; Balasubramanian *et al.*, 2019 ; Lakstygal *et al.*, 2018). La perte de fonction de *hdac1*, une désacétylase qui participe à la répression épigénétique, provoque une perte des oligodendrocytes dans le cerveau des larves mutantes (Cunliffe *et al.*, 2006). Des embryons traités avec de la morphine montrent une surexpression du gène *nestin* corrélée avec une hyperacétylation de la lysine 27. Les modifications épigénétiques inhibent la différenciation des cellules souches neurales (Jimenez-Gonzalez *et al.*, 2017). Les altérations du développement neural sont régulièrement associées avec des troubles du comportement. Une augmentation de la méthylation de *gabbr2* est corrélée à un déficit de compétences sociales chez le poisson, tout comme chez l'homme et la souris (Lakstygal *et al.*, 2019). Le poisson zèbre se révèle être un modèle pertinent dans l'étude du rôle des mécanismes épigénétiques sur le développement neural et comportementale. Néanmoins, tous les

mécanismes de régulation ne sont pas résolus. De plus, peu d'informations sont connues sur le rôle précis de la répression Polycomb chez le poisson zèbre alors que les protéines sont exprimées au cours du développement au niveau du cerveau en formation (Dupret *et al.*, 2016 ; 2017).

## 2.2 Les organes digestifs

Les organes digestifs sont formés à partir du tissu embryonnaire endodermique. La spécificité des organes est associée avec l'origine spatiale dans l'endoderme. Dans la région ventrale de l'endoderme, l'intestin antérieur se forme. Le pancréas et le foie sont générés à partir des régions latérales de l'intestin antérieur. (Zaret & Grompe, 2008). Tous ces organes sont issus d'un même feuillet embryonnaire, mais sous l'influence de nombreux facteurs, les cellules qui les composent se spécialisent au cours du développement. Néanmoins, une seule mutation d'un gène peut affecter le développement de chacun de ces organes (Pack *et al.*, 1996). L'origine commune de ces organes, avec des voies de différenciation partagées peut expliquer l'apparition de ces phénotypes chez les mutants.

### 2.2.1 L'intestin

L'utilisation de modèles poissons zèbres transgéniques a permis une meilleure compréhension du développement intestinal précoce par suivi du marquage GFP (Ng *et al.*, 2005). Trois stades de formation et différenciation ont été déterminés en fonction du stade de développement larvaire.

**Stade I : La formation de la lumière intestinale.** A partir de 26 hpf, les cellules endodermiques préalablement spécifiées forment l'intestin primitif. Cet intestin primitif est composé d'une fine couche de cellules mésodermiques au contact du sac vitellin et qui s'étend de la future bouche au futur anus. Une part de cet endoderme, dans les régions les plus antérieures formera par ailleurs l'endoderme pharyngoesophagéal et plus tard l'œsophage, le pharynx ainsi que la bouche. De 26 hpf à 52 hpf, la morphogenèse des principaux organes digestifs se met en place : le foie et le pancréas se positionnent sous forme de masses cellulaires le long de l'intestin primitif. Dès 48 hpf les cellules sont déjà spécifiées et sont déjà assignées à un futur organe. Elles ne sont pas forcément morphologiquement distinguables, des analyses transcriptomiques ont en revanche permis d'identifier à 2 jpf les différents types cellulaires (Wallace *et al.*, 2005). Au sein de l'endoderme, les cellules s'organisent pour former un tube, et cette formation n'est pas dépendante de phénomènes de mort programmée, mais bien de mouvements cellulaires. A partir de 52 hpf, la lumière est formée dans les régions antérieures, et gagne ensuite les régions postérieures. L'épithélium au contact de la lumière est organisé en monocouche de cellules cuboïdales. Ces cellules sont entourées de mésenchyme, peuplés de progéniteurs de neurones entériques. Des progéniteurs de cellules

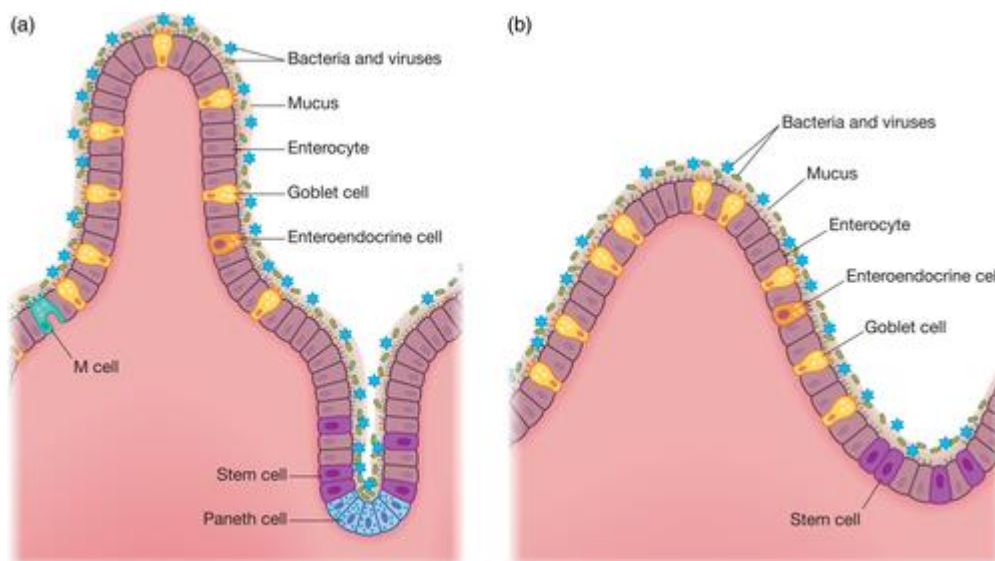
musculaires lisses entourent les régions intestinales antérieures. Ces deux types de progéniteurs prolifèrent et migrent dans les régions postérieures au cours du développement (Wallace *et al.*, 2005).

**Stade II : La polarisation cellulaire.** De 52 à 76 hpf, le deuxième stade de la morphogenèse prend place. Le primordium intestinal est entièrement organisé sous forme de tube. La connexion entre le foie et l'intestin par le conduit hépatique permet la séparation de l'intestin primitif en deux grandes régions. Les régions antérieures sont regroupées au sein de la région pharyngoesophagéal. Dans cette région, l'endoderme rejoint la vessie natatoire par le conduit pneumatique et délimite le futur œsophage. La partie caudale de l'endoderme jusqu'au conduit hépatique forme l'intestin primitif. Chez le poisson zèbre, l'estomac et le cecum sont absents, au contraire des mammifères. A partir de ce stade débute la polarisation des cellules épithéliales. Les PAN-cytokératines sont exprimées au pôle apical des cellules, en contact direct de la lumière intestinale. Le marqueur de jonction serrée ZO-1 et la cadhérine sont exprimés et les desmosomes sont visibles en microscopie électronique. Des jonctions adhérentes, jonctions serrées et desmosomes établis entre les cellules indiquent la polarisation de ces cellules intestinales. Ces jonctions établissent une barrière entre l'espace paracellulaire et la lumière intestinale (Wallace *et al.*, 2005). La polarisation des cellules, de même que l'ouverture de la lumière intestinale est mise en place des régions antérieures jusqu'aux régions postérieures. Les cellules qui composent le tube digestif primitif prolifèrent très activement.

**Stade III : Le remodelage et la polarisation cellulaire.** La troisième phase de la morphogenèse intestinale comprend la structuration des cellules épithéliales intestinales et leur différenciation complète. Le marqueur de différenciation intestinale *fabp2* est exprimé, et son expression augmente avec l'avancement de la différenciation de l'intestin (André *et al.*, 2000) Le tube digestif est spécialisé et compartimenté. Les compartiments intestinaux auront une fonction précise. Dès 4 jpf, la partie proximale de l'intestin s'élargit, et l'épithélium forme des invaginations. L'apparition de ces plis conditionne la prolifération des cellules intestinales : seules les cellules à la base des plis se répliquent. Les cellules épithéliales qui ne prolifèrent plus se différencient activement, et des changements morphologiques sont visibles. Une bordure avec des microvillosités apparaît sur la face apicale des entérocytes. Au niveau de l'intestin moyen, de larges vacuoles sont visibles au pôle apical des cellules pour former les cellules caliciformes qui secréteront le mucus. Les cellules caliciformes sont présentes entre les entérocytes. Enfin, chez les mammifères, les cellules entéroendocrines, présentes principalement dans l'estomac sécrètent des hormones nécessaires au contrôle de la prise alimentaire, de la digestion et la gestion de la glycémie (Montuenga *et al.*, 2003). Chez les téléostéens, les cellules entéroendocrines sont présentes et assurent les mêmes fonctions (Pan *et al.*, 2000). Une partie des cellules épithéliales intestinales assurent ces fonctions chez le poisson zèbre et Ng a montré que ces

cellules peuvent sécréter des hormones telle que le glucagon ou la somatostatine (Ng *et al.*, 2005). Ces cellules sont présentes tout au long du *tractus* intestinal.

A partir de 5 jpf, les réserves vitellines sont totalement consommées, mais l'ensemble du système digestif est fonctionnel. Les villosités intestinales sont restreintes principalement au niveau du bulbe intestinal. Au fur et à mesure de la croissance, les trois compartiments digestifs vont gagner en spécificité. Après 12 jpf, les villosités intestinales sont présentes dans chacun des compartiments intestinaux mais restent plus profondes au niveau du bulbe intestinal. L'étude du taux de prolifération cellulaire montre que les cellules en prolifération sont restreintes à la base de chaque villosité et ressemble à la structure intestinale des mammifères, connue sous le nom de crypte de Lieberkühn (Figure 18). Les cellules souches se répliquent et se différencient pour donner les différents types cellulaires intestinaux. Cette structure permet le maintien de l'épithélium digestif tout au long de la vie des poissons zèbres, de la même manière que chez les mammifères. Les cellules caliciformes sécrètent le mucus intestinal nécessaire à la protection de l'épithélium, la circulation du bol alimentaire ainsi que les échanges avec la flore intestinale.



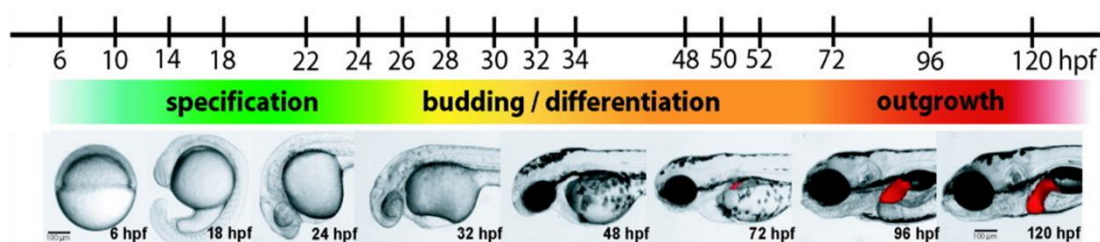
**Figure 18 : Comparaison de l'architecture de l'épithélium intestinale entre les mammifères en (a) et le poisson zèbre en (b).** L'épithélium intestinal des mammifères est composé de villi et de cryptes. Les cryptes contiennent des cellules sécrétrices de Paneth, qui ont un rôle dans le maintien de l'homéostasie de l'intestin et les échanges avec le microbiote. Les cryptes contiennent également des cellules souches qui se divisent de manière asymétrique et génèrent les entérocytes, les cellules caliciformes et les cellules entéroendocrines. Les cellules de Paneth et les cryptes de Lieberkühn sont absentes de l'intestin des poissons zèbres. Néanmoins les cellules souches sont présentes à la base des villosités et suivent le même schéma de différenciation en se rapprochant du sommet des villosités. Les cellules immunitaires M sont également absentes du système intestinal des poissons zèbres. La similarité architecturale et fonctionnelle de l'intestin fait du poisson zèbre un très bon modèle d'étude concernant le développement et des pathologies intestinales. Figure de Flores *et al.*, 2019.

### 2.2.2 Le foie

Chez les mammifères, le développement précoce du foie est lié à l'hématopoïèse. Aussi, des mutations qui impactent le développement hépatique ou hématologique précocement provoquent des phénotypes larges et compliquent l'étude du foie spécifiquement. L'hématopoïèse a lieu dans un premier temps au niveau de la masse cellulaire intermédiaire puis dans les reins chez le poisson zèbre (Tao & Peng, 2009) et est distinct du processus développemental de l'hépatogenèse.

Les hépatoblastes proviennent des cellules intestinales primitives, issues de l'endoderme (Figure 18). Les cellules progénitrices sont identifiables très tôt dans le développement par leur expression de deux marqueurs *hhex* et *prox1* dans le mésoderme. Ces cellules vont participer à l'établissement du bourgeon hépatique (Tao & Peng, 2009).

Le développement du foie est divisé en deux grandes étapes (Figure 19) : le bourgeonnement et la croissance. La première partie du bourgeonnement hépatique débute après 24 hpf. Des cellules progénitrices exprimant des marqueurs d'hépatoblaste se rassemblent en avant du tube intestinal. Au cours de la deuxième partie du bourgeonnement, le bourgeon hépatique se projette vers la partie gauche tout en restant à proximité de l'intestin. La troisième partie du bourgeonnement façonne fonctionnellement l'organe avec l'apparition d'une connexion sous forme d'un épithélium canalaire reliant le foie au bulbe intestinal. La deuxième grande étape, la croissance permet l'expansion de l'organe. Les cellules hépatiques poursuivent leur différenciation et l'organe se spécialise afin de remplir ses fonctions physiologiques (Field *et al.*, 2003).



**Figure 19 : Représentation de l'organogenèse hépatique en fonction du développement.** La phase de croissance est détectée par l'expression de la RFP dans la lignée transgénique Tg(*fabp10 : dsred*). Figure de Chu & Sadler, 2009.

Au niveau moléculaire, l'établissement de gradients de facteurs régulateurs par les tissus mésodermiques adjacents active les hépatoblastes. Les facteurs de transcription Bmp et Fgf, à une concentration spécifique, sont indispensables à l'initiation de la différenciation et induisent l'expression de facteurs hépatiques (Shin *et al.*, 2007). Les facteurs *foxA1-3*, *gata4-6*, *hhex*, *prox1* ou encore *hnf4* sont exprimés dès 28 hpf durant l'initiation du bourgeon hépatique (Field *et al.*, 2003). Les

cellules progénitrices se différencient en hépatocyte, en cellule du conduit biliaire, ou hépatopancréatique. Les cascades transcriptionnelles ont été décryptées. Les facteurs *notch1a*, *1b*, *2*, *5* sont exprimés dans les hépatoblastes, et participent à l'expression de *onecut1-3*. Ces facteurs autorisent l'expression de *vhnf1*, un facteur essentiel du développement biliaire (Tao & Peng, 2009).

Le facteur de transcription Fgf10 participe à la formation du conduit hépatopancréatique (Dong *et al.*, 2007). Des marqueurs de différenciation terminale hépatique tel que *lfabp* sont observés dès 48 hpf (Her *et al.*, 2003). L'expression de ces facteurs est stimulée ou inhibée par d'autres facteurs, mais également par des mécanismes épigénétiques. L'utilisation de morpholinos dirigés contre Dnmt2 inhibe particulièrement ses fonctions cytoplasmiques. Les histones désacétylases ont également des rôles spécifiques. L'enzyme Hdac1 semble avoir un rôle prépondérant au cours de l'organogenèse en orientant les progéniteurs endodermiques vers la différenciation hépatique et pancréatique (Noël *et al.*, 2008). De la même façon, Hdac3 est impliquée dans les étapes allant de la spécification à la différenciation (Farooq *et al.*, 2008).

La détoxification de l'organisme est également assurée au moins en partie par le foie. Cette fonction est assurée par la superfamille des enzymes cytochromes P450. Toutes les enzymes CYP humaines possèdent au moins un orthologue chez le poisson zèbre, avec des degrés d'homologie important (Goldstone *et al.*, 2010). Leur expression a été étudiée au cours du développement. Tous les transcrits n'apparaissent pas aux mêmes stades de développement, certains sont exprimés à 24 hpf, au moment du bourgeonnement hépatique tandis que d'autres ne seront exprimés que lors de la phase de croissance hépatique (Nawaji *et al.*, 2020). Les CYPs avec leur activité métabolique semblent prendre part au développement hépatique. De plus, la présence de l'entièreté du Cytochrome P450 dans l'arsenal métabolique du poisson zèbre en fait un modèle d'étude toxicologique et métabolique relevant.

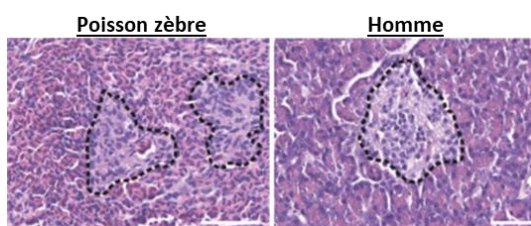
L'anatomie du foie diffère légèrement entre les poissons et les mammifères. L'organe est organisé en deux lobes latéraux et un ventral sans pédicules de séparation entre les lobes, présents chez les mammifères. A l'inverse d'une organisation portale chez les mammifères, les hépatocytes des poissons sont disposés en tubules autour de ductules biliaires. Les cellules immunitaires spécialisées du foie, les cellules Kupffer, n'ont également pas été identifiées chez le poisson. Néanmoins, toutes les autres cellules hépatiques sont très similaires entre mammifères et téléostéens et partagent des morphologies et fonctions communes : sécrétion de bile et glycogène, réponse à l'insuline, stockage des lipides, métabolisme de l'ammoniaque et xénobiotique, sécrétion des protéines sériques (Goessling & Sadler, 2015 ; Katoh & Patial, 2020). La conservation des fonctions hépatiques chez le

poisson zèbre en fait un modèle de choix dans l'étude des maladies affectant le foie. Dès 1965, Stanton établit le modèle poisson zèbre comme modèle du cancer du foie par traitement avec un carcinogène (Stanton, 1965). Des protocoles pour induire les cirrhoses alcooliques ont également été mis en place (Laale, 1971). De nombreuses lignées transgéniques ont été développées pour étudier le développement et la carcinogenèse hépatique (Katoch & Patial, 2020). Le poisson zèbre est donc devenu un modèle de choix dans l'étude des différentes maladies hépatiques.

### 2.2.3 Le pancréas

Les progéniteurs pancréatiques sont spécifiés dès 6 hpf. Le premier marqueur exprimé par les progéniteurs est *pancreatic and duodenal homeobox 1 (pdx1)*. Ce marqueur est exprimé dès le stade gastrula dans l'ensemble du primordium intestinal. Après 20 hpf, les progéniteurs pancréatiques sont rassemblés en 2 bourgeons cellulaires ventralement et dorsalement. Ces bourgeons déterminent le pancréas exocrine et endocrine. Les cellules pancréatiques exocrines proviennent du bourgeon ventral qui expriment le marqueur *ptf1a* tandis que les cellules endocrines sont issues du bourgeon dorsal, avec des marqueurs comme l'insuline ou le glucagon (Ward *et al.*, 2007 ; Tiso *et al.*, 2009). Ces deux bourgeons fusionnent autour de 48 hpf. Ce mécanisme de formation du pancréas est très conservé parmi les vertébrés (Slack, 1995).

Les cellules exocrines se rassemblent sous forme d'acini et sécrètent les enzymes digestives trypsine, amylase et carboxypeptidase dans l'intestin qui permettront la digestion (Biemar *et al.*, 2001). Le pancréas exocrine entoure les cellules endocrines. C'est pourquoi les marqueurs de cellules exocrines comme la trypsine sont souvent choisis pour évaluer le développement général du pancréas. Les cellules endocrines sont rassemblées en un îlot unique au stade larvaire. Au centre de cet îlot se trouvent les cellules  $\beta$  qui produisent l'insuline. Elles sont entourées des cellules  $\alpha$ ,  $\epsilon$ ,  $\delta$ , qui produisent respectivement le glucagon, la ghréline et la somatostatine (Kinkel & Prince, 2009). L'îlot de cellules endocrine est similaire, tant d'un point de vue structure que fonctionnel, des îlots de Langerhans des mammifères (Figure 20).



**Figure 20 : Coupe histologique de pancréas de poisson zèbre et humain.** La coloration hématoxyline et éosine révèle un cytoplasme foncé riche en mucus des cellules pancréatiques exocrines. Les cellules claires rassemblées en îlots sont les cellules endocrines. L'organisation du pancréas est similaire chez le poisson zèbre et l'homme. La barre d'échelle représente 1 mm. Figure de Bordeira-Carriço *et al.*, 2022.



De l'induction primaire des progéniteurs à la différenciation complète des cellules pancréatiques, le processus d'organogenèse du pancréas est soumis à une régulation fine, complexe et multifactorielle. Les grandes de voie de signalisation telle que Notch, l'acide rétinoïque, les FGFs, et les BMPs sont impliqués (Figure 21). Des acteurs épigénétiques permettent également la spécification des progéniteurs et leur différenciation. L'inhibition de la voie de différenciation intestinale par Hdac1 promeut la spécification et la différenciation pancréatique ainsi que la formation de l'îlot endocrine (Noel *et al.*, 2008). L'inhibition de Hdac3 par le biais de morpholinos conduit à la dérégulation de *gdf11* (*growth differentiation factor 11*) et à l'inhibition de la formation du pancréas exocrine. De façon surprenante, le pancréas endocrine n'est pas impacté par la perte de *hdac3* (Farooq *et al.*, 2008).

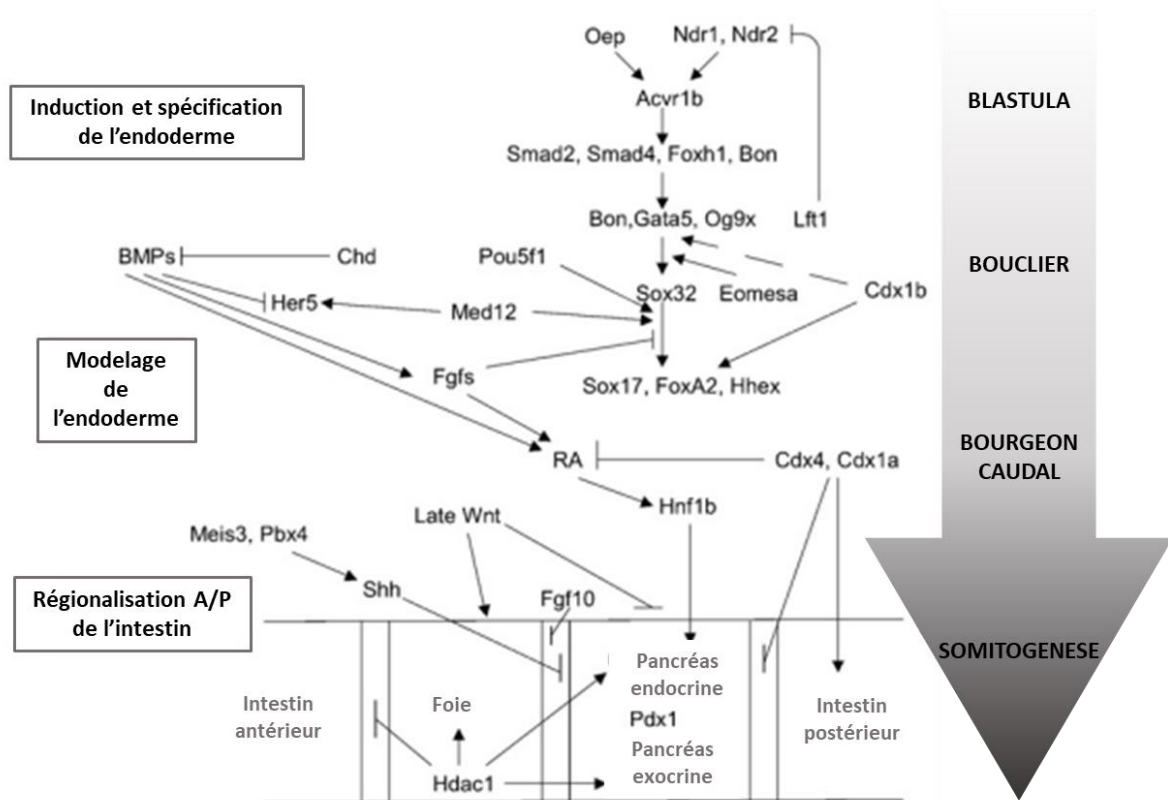


Figure 21 : Représentation schématique de la cascade de signalisation à l'origine de l'induction, la spécification et la régionalisation des organes du système digestif. Les lignes fléchées et barrées indiquent respectivement les relations négatives et positives. Le recrutement de facteurs spécifiques évolue en fonction du développement. Figure de Tiso *et al.*, 2009

La perte de fonction de la méthyltransférase Dnmt1 n'altère pas le développement pancréatique mais le maintien de l'organe, et en particulier la survie des cellules acineuses. De plus, le renouvellement des cellules  $\beta$  n'est plus assurée (Anderson *et al.*, 2009). La répression épigénétique semble donc avoir

un rôle clé dans la spécification des progéniteurs mais également le maintien des structures mises en place au cours du développement. Et la perte de fonction de régulateurs épigénétiques n'affectent que très rarement uniquement un organe, mais semble affecter l'ensemble du système digestif.

### 3. Les complexes PRC1 et PRC2 chez le poisson zèbre

Les régulations épigénétiques gouvernant le destin et la différenciation des cellules sont des mécanismes extrêmement conservés au cours de l'évolution chez les vertébrés (Sun *et al.*, 2008). La répression Polycomb est elle aussi conservée chez le poisson zèbre. Les protéines du groupe polycomb *cbx2* et *ph2* ont été mises en évidence pour leur rôle dans le développement des embryons de poissons zèbres dès le début des années 2000 (Kawamura *et al.*, 2002a ; 2002b). Après la duplication du génome survenue dans l'embranchement des téléostéens, plusieurs paralogues sont apparus et ont été maintenus au cours de l'évolution (Le Faou *et al.*, 2011), tandis que d'autres, comme *RING1*, *RBBP7* ou *PCGF2* ont été perdus (Table 4). L'expression de plusieurs gènes de la famille Polycomb a été observée par hybridation *in situ*. Les gènes sont fortement exprimés au cours du développement des embryons ainsi que dans les organes adultes, notamment dans les gonades, ce qui montre une importance de ces gènes dans le contrôle du développement et du statut de différenciation cellulaire (Chrispijn *et al.*, 2018). Le séquençage du génome entier du poisson zèbre ainsi que l'amélioration des technologies de biologies moléculaires ont permis l'édition précise des gènes chez le poisson zèbre et en particulier des gènes de la famille Polycomb. Ces mutants ont permis l'étude fonctionnelle des gènes chez le poisson zèbre.

Table 4 : Orthologues des gènes de la famille Polycomb humain chez le poisson zèbre

		Gène humain	Gène de poisson zèbre
PRC2	Core	<i>EZH1</i>	<i>ezh1</i>
		<i>EZH2</i>	<i>ezh2</i>
		<i>EED</i>	<i>eed</i>
		<i>SUZ12</i>	<i>suz12a/b</i>
		<i>RBBP4</i>	<i>rbbp4</i>
		<i>RBBP7</i>	
	PRC2.1	<i>PHF1</i>	<i>phf1</i>
		<i>MTF2</i>	<i>mtf2</i>
		<i>PHF19</i>	<i>phf19</i>
	PRC2.2	<i>AEBP2</i>	<i>aebp2</i>
		<i>JARID2</i>	<i>jarid2a/b</i>
	PRC1	PRC1c	<i>RING1</i>
<i>RNF2</i>			<i>rnf2</i>
<i>CBX2</i>			<i>cbx2</i>
<i>CBX4</i>			<i>cbx4</i>
<i>CBX6</i>			<i>cbx6a/b</i>
<i>CBX7</i>			<i>cbx7a/b</i>
<i>CBX8</i>			<i>cbx8a/b</i>
<i>PCGF2</i>			

	<i>BMI1</i>	<i>bmi1a/b</i>
	<i>PHC1</i>	<i>phc1</i>
	<i>PHC2</i>	<i>phc2a/b</i>
	<i>PHC3</i>	<i>phc3</i>
PRC1nc	<i>RYBP</i>	<i>rybpa/b</i>
	<i>YAF2</i>	<i>yaf2</i>
PRC1.1	<i>PCGF1</i>	<i>pcgf1</i>
	<i>BCOR</i>	<i>bcor</i>
	<i>BCORL</i>	<i>bcorl</i>
	<i>KDM2B</i>	<i>kdm2ba/b</i>
	<i>SKP1</i>	<i>skp1</i>
PRC1.3	<i>PCGF3</i>	
PRC1.5	<i>PCGF5</i>	<i>pcgf5a/b</i>
	<i>CSNK2A1-3</i>	<i>csnk2a, csnk2a2a/b</i>
	<i>AUTS2</i>	<i>auts2a/b</i>
	<i>FBRSL1</i>	<i>fbrsl1</i>
	<i>FBR5</i>	<i>fbr5</i>
PRC1.6	<i>PCGF6</i>	<i>pcgf6</i>
	<i>MGA</i>	<i>mgaa/b</i>
	<i>MAX</i>	<i>max</i>
	<i>CBX3</i>	<i>cbx3a/b</i>
	<i>L3MBTL2</i>	<i>l3mbtl2a/b</i>
	<i>E2F6</i>	<i>e2f6a/b</i>
	<i>TDP1</i>	<i>tdp1a/b</i>
	<i>WDR5</i>	<i>wdr5</i>

### 3.1 Le complexe PRC1

La sous-unité catalytique du complexe PRC1 a été largement étudiée chez le poisson zèbre. Alors que les mammifères possèdent deux orthologues *RING1* et *RNF2*, seul *rnf2* est conservé chez le poisson zèbre et porte l'activité catalytique du complexe PRC2. De la même façon que chez les mammifères, la perte de fonction de *rnf2* a des effets délétères très tôt dans le développement. La mise en place de la répression Polycomb via l'application de la marque H2AK119ub1 a lieu avant l'activation du génome zygotique, et surtout préalablement à la déposition de la marque H3K27me3. L'établissement de cette répression permet la régulation des programmes développementaux dans les cellules souches embryonnaires (Hickey *et al.*, 2022). La perte de fonction de *Rnf2* altère la prolifération et la différenciation de plusieurs types cellulaires (van der Velden *et al.*, 2013 ; Feng & Sun, 2022). Les mutants *rnf2* ne développent pas de nageoires pectorales en raison d'un défaut d'activation des voies de signalisation Fgf (van der Velden *et al.*, 2012). Ils présentent également des malformations crâniofaciales sévères avec une réduction drastique des cartilages. Les cartilages crâniofaciaux sont issus de progéniteurs de la crête neurale crâniale. Le complexe PRC1 a donc un rôle prépondérant dans

l'orientation de ces progéniteurs en chondrocytes (van der Velden *et al.*, 2013). Le développement neural est également organisé par le PRC1. Le profil d'expression de plusieurs gènes impliqués dans l'initiation et la maturation des cellules neurales a été observé par hybridations *in situ*. Très tôt dans le développement, l'expression de *nestin* ou *sox9b* est réduite tandis qu'une expression ectopique de *sox1b* le long de l'axe AP est observée. L'activité de PRC1 médiée par *rnf2* est indispensable à la différenciation des cellules neurales (Feng *et al.*, 2022). La protéine Rnf2 est également impliquée dans la mise en place et le développement cardiaque. La répression médiée par Rnf2 sur les facteurs de transcription Tbx permet la différenciation des cellules cardiaques et leur spécialisation en fonction de leur position dans l'organe (Chrispijn *et al.*, 2019). Par ailleurs, la protéine Rnf2 est impliquée dans l'activité propre du cœur. La perte de fonction de Rnf2 obtenue après ciblage par CRISPR/Cas9 n'altère pas le développement cardiaque précoce, mais la perte de répression transcriptionnelle de gènes impliqués dans la conduction cardiaque contribue à des problèmes de contractions cardiaques (Peng *et al.*, 2021).

Plusieurs des protéines associées à *rnf2* au sein du complexe PRC1 ont également fait l'objet d'investigations. La protéine Bmi1a est impliquée dans le maintien des cellules souches et progéniteurs hématopoïétiques chez les mammifères et chez le poisson zèbre. Chez les poissons zèbres, la régulation de ce statut est sous le contrôle du facteur C/ebpα et médiée par Bmi1a (Zhou *et al.*, 2011). La protéine Pcgf1 est exprimée précocement dans le développement des poissons zèbres. Les mutants *pcgf1* se développent correctement mais présentent un retard de développement. En revanche, la perte de *pcgf1* provoque une sénescence prématurée des poissons (Dupret *et al.*, 2016). Le KD de *pcgf1* inhibe le développement neural précoce des embryons et semblent bloquer les cellules neurales au stade progéniteurs (Li *et al.*, 2021). L'utilisation de morpholinos dirigés contre *ph2α* a révélé son action au cours du développement précoce. Le recrutement de *ph2α* par les FGF est nécessaire aux mouvements morphogénétiques durant l'épibolie (Komoike *et al.*, 2005). Chez l'homme, des mutations de *BCOR* sont associées avec des anomalies de développement. Les phénotypes observés chez les embryons mutants sont similaires à ce qu'il est noté en clinique (Ng *et al.*, 2004). L'activité du PRC1nc médiée par *bcor* est indispensable à la répression de *tp53* et au contrôle de l'apoptose, et autorise le développement normal de l'œil (Lee *et al.*, 2013). L'inhibition de *mga* par des morpholinos provoquent des malformations cérébrales chez les embryons dès 24 hpf. De la même façon que chez les mutants *rnf2*, la survie des cellules neurales est compromise en l'absence de Mga. Les organes sont mis en place précocement mais des malformations cardiaques et intestinales sont observées (Rikin & Evans, 2010). De nombreuses zones d'ombre demeurent néanmoins sur l'expression et les activités spécifiques des autres acteurs du PRC1 chez le poisson zèbre.

### 3.2 Le complexe PRC2

Le rôle du complexe PRC2 a également été étudié. Chez le poisson zèbre, deux orthologues de *SUZ12* sont exprimés : *suz12a* et *suz12b*. La mutation de ces deux sous-unités est létale autour de 15 jpf (Oppel *et al.*, 2020). La délétion de *rbbp4* est létale avant 10jpf. Les larves mutantes présentent des phénotypes neurogéniques sévères. La perte de *Rrbp4* altère la différenciation des cellules neurales et compromet la survie des précurseurs neuraux (Schultz *et al.*, 2018). Mais les fonctionnalités du complexe PRC2 ont surtout été analysées par le biais des deux sous-unités catalytiques *ezh1* et *ezh2*. La délétion de *Ezh1* favorise la prolifération des cellules souches et progéniteurs hématopoïétiques (Soto *et al.*, 2016). Le gène *ezh2* est fortement exprimé dans les stades précoces du développement, et son expression est restreinte aux zones de prolifération cellulaire dans les stades plus tardifs. L'activité de *Ezh2* est nécessaire au maintien de l'intégrité intestinale. La perte de fonction de *Ezh2* conduit à la mort des larves autour de 12 jpf (Dupret *et al.*, 2017). Un transcrite *ezh2* est déposé dans les ovocytes. Afin d'étudier le rôle de ce transcrite d'origine maternelle, un mutant zygotique a été généré par l'équipe de San. Contrairement aux mutants murins, la gastrulation se déroule normalement chez les embryons mutants. Néanmoins les mutants *MZezh2* meurent après 2 jpf en raison de phénotypes multiples : des organes comme le cœur ou le foie sont mis en place, mais ne sont pas maintenus au cours du développement (San *et al.*, 2016). Une approche multi-omique révèle par ailleurs que *Ezh2* est requis pour le maintien d'une répression génique spatiale nécessaire à la formation du plan du corps (Rougeot *et al.*, 2019). L'inhibition chimique de *Ezh2* a également révélé un rôle dans le métabolisme lipidique (den Broeder *et al.*, 2020). La sous-unité catalytique est exprimée de manière cyclique au cours des journées et est impliquée dans le contrôle du rythme circadien chez le poisson zèbre. Le développement des cellules hématopoïétiques chez les mutants *ezh2* est altéré, le profil d'expression des marqueurs hématopoïétiques précoces est réduit chez les embryons *ezh2*<sup>-/-</sup> (Zhong *et al.*, 2018). L'activité des protéines accessoires de PRC2 a également été questionnée. Des questions demeurent néanmoins sur l'activité spécifique de chacune des sous-unités catalytiques.

# OBJECTIFS DE LA THESE

---

Malgré les dernières avancées sur la caractérisation de la répression Polycomb, de nombreuses inconnues demeurent. Les orthologues des sous-unités du PRC1 et du PRC2 sont conservées chez le poisson zèbre (Sun *et al.*, 2008 ; Le Faou *et al.*, 2011). Si ces orthologues ont été identifiés dans le génome des poissons, très peu de choses sont connues quant à leur expression et leur spécificité d'action. Chez le modèle murin, il existe parfois une redondance d'activité mais également des spécificités d'action au cours du développement et de la carcinogénèse, et notamment des protéines CBX (Schuettengruber *et al.*, 2017). Une caractérisation de ces gènes est nécessaire afin d'améliorer nos connaissances sur la répression Polycomb, le recrutement des complexes et leurs activités chez le poisson zèbre. L'activité du complexe PRC1 dans le développement et au stade adulte sera abordée au travers de l'étude de l'expression des gènes *cbx*. Le laboratoire possède une expertise dans l'étude fonctionnelle de sous-unités des complexes Polycomb chez le poisson zèbre. Plusieurs mutants ont été générés par l'utilisation des TALENs, et ont permis l'étude des protéines Pcgf1 et Ezh2 chez le poisson zèbre (Dupret *et al.*, 2016 ; 2017). Concernant les mutants *ezh2*, la perte de fonction chez la souris montre un arrêt du développement au moment de la gastrulation des embryons (O'Caroll *et al.*, 2001). Chez le poisson zèbre, l'ARNm d'origine maternelle semble permettre le développement très précoce des embryons. Les mutants meurent autour de 12 jpf (Dupret *et al.*, 2017). Ainsi, le poisson zèbre est un modèle relevant pour étudier le rôle de la répression Polycomb dans le développement tardif des embryons et observer des phénotypes qu'il n'est pas possible de constater chez la souris sans une mutation conditionnelle ou tissu-spécifique.

Les objectifs de cette thèse comprennent l'utilisation du modèle poisson zèbre pour étudier le rôle de la répression Polycomb dans le développement des embryons ainsi que dans le développement tumoral. Le développement embryonnaire et la tumorigénèse partagent des mécanismes, des voies de signalisation et des transcriptomes proches. Dans certaines tumeurs du système nerveux central, une mutation du gène *ACVR1*, un récepteur à tyrosine kinase, est observée. Cette mutation conduit à une activation constitutive de la voie BMP (Duchatel *et al.*, 2019). Cette voie est particulièrement importante dans le développement précoce des vertébrés, et du système nerveux central. Ainsi, une meilleure compréhension du développement embryonnaire peut révéler des processus impliqués dans chaque étape de la carcinogénèse.

Ainsi les finalités de ce travail sont multiples : (i) étudier les gènes de la famille Polycomb chez le poisson zèbre, (ii) étudier les fonctions de ces gènes dans le développement du poisson zèbre et (iii)

montrer la pertinence du poisson zèbre comme modèle d'étude de cancers qui impliquent la répression Polycomb.



# RESULTATS

---

Dans un premier temps, l'étude des gènes de la famille Polycomb sera abordée. Les gènes *cbx* du complexe PRC1 seront décrits dans un premier article. Le complexe PRC2 sera traité par deux articles. Le premier article porte sur l'évolution conjointe de *ezh1* et *ezh2* et leur fonction dans le développement. Le deuxième article portera sur l'activité du complexe PRC2 au travers de l'étude fonctionnelle de *eed*.

Dans un second temps, le poisson zèbre sera décrit comme modèle de cancérologie. Une revue de littérature et un article de recherche sur le cancer du sein décriront la transplantation de cellules cancéreuses chez le poisson zèbre. Une deuxième revue exposera l'utilisation du poisson zèbre dans la modélisation génétique de cancers humains.



# 1. Les gènes *cbx*: The Polycomb Orthologues in Teleost Fishes and Their Expression in the Zebrafish Model

## 1.1 Contexte scientifique

La duplication entière du génome à l'origine de la spéciation de la branche des téléostéens chez les vertébrés est à l'initiation d'une diversité d'espèce exceptionnelle. Après une duplication de la sorte, dans la majorité des cas l'un des deux gènes dupliqués est aléatoirement perdu. L'accumulation d'événements mutationnels sur ces ohnologues est délétère et forme des pseudogènes qui ne seront pas fonctionnels, tandis que parfois, ces mutations confèrent un avantage sélectif et autorisent la conservation des deux ohnologues (Glasauer & Neuhauss, 2014). La conservation des ohnologues peut également être expliquée par la mise sous silence d'un des deux gènes par des mécanismes épigénétiques comme la méthylation de l'ADN. Ces phénomènes peuvent favoriser le maintien de la copie du gène (Rodin & Riggs, 2003). La conservation d'un ohnologue peut être également associée avec l'acquisition d'une nouvelle fonction. Dans ces cas-là, les rôles spécifiques des ohnologues ne sont pas toujours définis et connus. Les premières études sur ces gènes montrent qu'ils sont principalement impliqués dans les grands processus biologiques tels que le développement, la signalisation ou encore la régulation fonctionnelle (Brunet *et al.*, 2006).

D'abord identifiés et caractérisés chez la drosophile, les gènes de la famille Polycomb sont conservés chez les vertébrés. Le nombre d'orthologues a augmenté au cours de l'évolution et les deux complexes se sont diversifiés au cours de l'évolution. Alors que plusieurs événements de duplication génomique sont connus pour avoir eu lieu au cours de l'évolution des vertébrés, la duplication à l'origine de l'apparition des téléostéens a augmenté le nombre d'homologues Polycomb. La conservation des copies n'est pas uniforme entre toutes les espèces. Tandis que *cbx8* est dupliqué chez le poisson zèbre, ce gène n'est présent qu'en une seule unité chez le Fugu (Senthilkumar & Mishra, 2009). Les protéines de la famille Cbx relèvent d'une importance capitale dans le recrutement du complexe PRC1, par la reconnaissance de la marque H3K27me3 (Fischle *et al.*, 2003). Les protéines CBX chez les mammifères sont intégrées et définissent la composition des complexes PRC1, mais elles définissent également les modalités de recrutement et la fonctionnalité des complexes (Vandamme *et al.*, 2011 ; Morey *et al.*, 2012 ; Klauke *et al.*, 2013). Afin d'approfondir nos connaissances sur les protéines de la famille CBX, et sur l'histoire évolutive du PRC1 chez les téléostéens, une analyse *in silico* a été conduite sur 68 espèces. Le poisson zèbre est couramment utilisé comme modèle de recherche. Aussi, nous avons examiné le profil d'expression des gènes *cbx* dans le développement et à l'âge adulte de l'espèce *Danio rerio*.

## 1.2 Article

## The *Polycomb* Orthologues in Teleost Fishes and Their Expression in the Zebrafish Model

Ludivine Raby, Pamela Völkel, Xuefen Le Bourhis and Pierre-Olivier Angrand \*

University Lille, CNRS, Inserm, CHU Lille, Centre Oscar Lambret, UMR 9020 - UMR 1277 - Canther - Cancer Heterogeneity, Plasticity and Resistance to Therapies, F-59000 Lille, France; ludivine.raby.etu@univ-lille.fr (L.R.); pamelavoelkel@univ-lille.fr (P.V.); xuefen.le-bourhis@univ-lille.fr (X.L.B.)

\* Correspondence: pierre-olivier.angrand@univ-lille.fr; Tel.: + 33-320-336-222

Received: 21 February 2020; Accepted: 26 March 2020; Published: 27 March 2020

**Abstract:** The Polycomb Repressive Complex 1 (PRC1) is a chromatin-associated protein complex involved in transcriptional repression of hundreds of genes controlling development and differentiation processes, but also involved in cancer and stem cell biology. Within the canonical PRC1, members of Pc/CBX protein family are responsible for the targeting of the complex to specific gene loci. In mammals, the Pc/CBX protein family is composed of five members generating, through mutual exclusion, different PRC1 complexes with potentially distinct cellular functions. Here, we performed a global analysis of the *cbx* gene family in 68 teleost species and traced the distribution of the *cbx* genes through teleost evolution in six fish super-orders. We showed that after the teleost-specific

whole genome duplication, *cbx4*, *cbx7* and *cbx8* are retained as pairs of ohnologues. In contrast, *cbx2* and *cbx6* are present as pairs of ohnologues in the genome of several teleost clades but as singletons in others. Furthermore, since zebrafish is a widely used vertebrate model for studying development, we report on the expression of the *cbx* family members during zebrafish development and in adult tissues. We showed that all *cbx* genes are ubiquitously expressed with some variations during early development.

### Introduction

Teleosts, with over 29,000 species, represent the largest and most diverse infraclass (*Teleostei*) of vertebrates accounting for about half of the living vertebrates and more than 96% of all fish species [1–3]. Gene and gene family early sequencing revealed the existence of a whole-genome duplication event that occurred in the teleost lineage about 320 Myr ago, before its diversification [4–6], and known as the teleost genome duplication (TGD). Ohnologous genes produced by whole genome duplication are presumably redundant and one copy is subsequently lost randomly [7]. Less frequently, both ohnologues are maintained due to accumulation of mutations leading to neofunctionalization, subfunctionalization or subneofunctionalization of one or two of the ohnologues [8]. Following the TGD in the teleost ancestor, about 15% to 20% of the ohnologues were retained as pairs [9]. Interestingly, an asymmetric acceleration of evolutionary rate is observed for one of the ohnologues allowing a possible neofunctionalization and/or subfunctionalization. Furthermore, the teleost evolution is associated with major interchromosomal rearrangements having occurred after the TGD [10,11] and to the fact that teleost genes evolved faster compared to their tetrapode orthologues [12]. These genomic characteristics, sub/neofunctionalization of the retained ohnologues, large interchromosomal rearrangements and rapid evolutionary rates, are thus believed to be, in part, responsible for the remarkable diversity of the teleost fishes in terms of morphology, physiology, behaviors and adaptations.

Here, we report on the study of the Polycomb orthologues in teleosts. Initially discovered in *Drosophila*, Polycomb repression is a conserved epigenetic mechanism involved in the repression of hundreds of genes implicated in various biological processes such as the regulation of tissue homeostasis, maintenance and lineage specification of stem cells, cell differentiation and development [13–16]. Polycomb repression is ensured by at least two chromatin-modifying protein complexes, named Polycomb Repressive Complex 1 and 2 (PRC1 and PRC2). According to the

canonical pathway, the PRC2 complex is recruited at defined chromatin locations and initiates gene silencing by catalyzing trimethylation of histone H3 at lysine 27 (H3K27me3). The PRC1 protein complex is then recruited to the target regions by binding to H3K27me3 epigenetic marks, catalyzes the monoubiquitinylation of histone H2A at lysine 119 (H2AK119ub1) and induces local chromatin compaction to maintain gene silencing [17–21]. In *Drosophila*, the PRC2 core complex is composed of Suppressor of zeste 12 [Su(z)12], Extra sex combs (Esc) and Enhancer of zeste [E(z)], which is the catalytic subunit of the complex [22,23].

The canonical PRC1 core complex comprises the Polycomb protein itself (Pc), Polyhomeotic (Ph), Posterior sex combs (Psc) and Sex combs extra (Sce). Sce deposits the H2AK119ub1 epigenetic mark, whereas Pc binds to H3K27me3 through its chromodomain [24,25]. In mammals, PRC1 organization appears to be more complex than in *Drosophila* because there are several orthologues for each of its components [26,27]. Actually, mammalian genomes code for five orthologues for Pc (CBX2, 4, 6, 7 and 8), six Psc orthologues (BMI1, PCGF1, 2, 3, 5, and 6), three Ph orthologues (PHC1, 2 and 3) and two Sce orthologues (RING1 and RNF2). Due to combinatorial permutations, distinct PRC1 complexes, different in their subunit composition, are present in mammalian cells [28–33]. Functionally, PRC1 diversity depends largely on the Pc orthologue present in the complex since it has been shown that the different CBX proteins could bind in part, to different targets [33–36]. Indeed, differentiation of embryonic or hematopoietic stem cells relies on a dynamic switch in the expression of the *CBX* genes and on the identity of the Polycomb paralogs present in the PRC1 [33,35,36].

In a previous study, we showed that some PRC1 subunit members are not similarly retained in all teleost species [37]. In zebrafish (*Danio rerio*), the Psc orthologues *pcgf1* and *pcgf6* exist as singletons, *bmi1a/b* and *pcgf5a/b* as pairs, whereas *pcgf2* and *pcgf3* are absent. In contrast, in medaka (*Oryzias latipes*, strain HdrR), *bmi1*, *pcgf2*, *pcgf3*, *pcgf5* and *pcgf6* are all present as singletons but *pcgf1* is absent. Similarly, the medaka genome contains two Sce orthologues (*ring1* and *rnf2*), whereas only *rnf2* is present in zebrafish [37]. However, it is not known whether the Pc orthologues are all maintained nor whether the *CBX* genes are present as singletons or pairs in teleosts. Thanks to recent advances in teleost genome sequencing, we conducted the analysis of the Pc/CBX gene family in 68 teleost species that belong to 21 different orders. Furthermore, we report here on the expression of the Pc/CBX family members during zebrafish development and in adult tissues since this teleost is a widely used vertebrate model for studying development and gene functions.

## Materials and Methods

### *In Silico Identification and Analyses of Pc/CBX Sequences in Databases*

TBLASTN searches in the NCBI (National Center for Biotechnology, Bethesda, Maryland, USA) and Ensembl databases were used to retrieve the teleost *cbx* genes. The human (CBX2\_HUMAN, Q14781; CBX4\_HUMAN, O00257; CBX6\_HUMAN, O95503; CBX7\_HUMAN, O95931 and CBX8\_HUMAN, Q9HC52) and zebrafish (Cbx2, F1Q982\_DANRE; Cbx4, B7ZUU4\_DANRE; Cbx6a, XP\_021325056.1; Cbx6b, X1WBY5\_DANRE; Cbx7a, F1RBK7\_DANRE; Cbx7b, E7F8F2\_DANRE; Cbx8a, Q6TH37\_DANRE and Cbx8b, A8KC45\_DANRE) reference sequences were used as queries. Teleost Cbx3 sequences were identified through TBLASTN searches using zebrafish Cbx3a, F1QNF2\_DANRE and Cbx3b, E9QDB7\_DANRE reference protein sequences as queries. Sequence alignments were performed using the Clustal Omega [38] and ClustalW [39] algorithms. Phylogeny was determined by aligning the teleost Cbx sequences with the MUSCLE algorithm [40] included in MEGA X (Pennsylvania State University, USA) [41]. Phylogenetic trees were produced by the Maximum Likelihood (ML) method and JTT matrix-based model [42], the Neighbor-Joining (NJ) method [43] and the Maximum Parsimony (MP) using the Subtree-Pruning-Regrafting (SPR) algorithm [44] in MEGA X [41]. SMART and Pfam protein domains were identified using the Simple Modular Architecture Research Tool [45]. Consensus patterns were drawn in logo format using WebLogo [46].

Syntenic group identification at the *cbx* gene loci and vicinal genes in teleost genomes were assessed by manual chromosome walking and reciprocal BLAST searches using the NCBI Gene database (National Center for Biotechnology, Bethesda, MD, USA).

### *Zebrafish Maintenance, Embryo Preparation and Animal Ethics Statements*

Zebrafish from the TU strain were kept at 27.5 °C in a 14/10 h light/dark cycle. The evening before spawning, males and females were separated into individual breeding tanks (Tecniplast). Spontaneous spawning occurred the following morning when the plastic divider is removed. Embryos were collected and staged according to Kimmel *et al.* [47]. The chorions were removed from embryos by action of 1% pronase (Sigma, St. Louis, MO, USA) for 1 min. Zebrafish embryos were fixed overnight in 4% paraformaldehyde in PBS (phosphate-buffered saline, Euromedex, Souffelweyersheim, France), dehydrated gradually to 100% methanol and kept at -20 °C.

Zebrafish were maintained in compliance with the French and European Union guidelines for the handling of laboratory animals (Directive 2010/63/EU of the European Parliament and of the Council of 22 September 2010 on the protection of animals used for scientific purposes). The experimental procedures carried out on zebrafish were reviewed and approved by the local Ethics Committee, CEEA 75 Nord Pas-de-Calais and the French Ministry of Higher Education and Research (APAFiS approval number 13527-2018011722529804\_v3).

### *Whole-Mount In Situ Hybridization*

Antisense-RNA probes were synthesized using the DIG RNA Labeling Kit (SP6/T7) (Roche, 11175025910) according to the manufacturer's instruction. The following IMAGE cDNA clones purchased at imaGenes GmbH (Berlin, Germany), were used as templates: *cbx2*, cDNA clone MGC:103563 IMAGE:7241894; *cbx7a*, cDNA clone MGC:110152 IMAGE:7289412; *cbx8a*, cDNA clone MGC:153854 IMAGE:8001741; *cbx8b*, cDNA clone MGC:111978 IMAGE:7399500. *cbx4*, *cbx6a*, *cbx6b* and *cbx7b* antisense probes were generated using RT-PCR from total mRNA extracted from zebrafish larvae at 5 dpf using the RNeasy Mini Kit (Qiagen, Hilden, Germany). After Reverse Transcription, cDNAs were amplified by PCR using the probe-specific primers, coupled to the T7 sequence for forward primers and the SP6 sequence for reverse primers. The primers used for probe generation were:

ISH\_cbx4a1\_508F: 5'-TAATACGACTCACTATAGGGGTCGAGCAGCATATGAGCGG-3'  
ISH\_cbx4a1\_1921R: 5'-GATTTAGGTGACACTATAGAATTTGCCAGGGGCATGAAT-3'  
ISH\_cbx6a2\_1210F: 5'-TAATACGACTCACTATAGGGGTGGGTCGAGCTTCGGTCTGCC-3'  
ISH\_cbx6a2\_1460R: 5'-GATTTAGGTGACACTATAGGCAGTTGTGGTGGGAGAGTTTGGGT-3'  
ISH\_cbx6b\_1037F: 5'-TAATACGACTCACTATAGGGCCCTTCGACTAAGGCCCTGA-3'  
ISH\_cbx6b\_1633R: 5'-GATTTAGGTGACACTATAGGCTTGACCACATCCAAGCCCAACC-3'  
ISH\_cbx7b\_304F: 5'-TAATACGACTCACTATAGGGGCTCATCTGCCACTCTCGCTGGA-3'  
ISH\_cbx7b\_692R: 5'-GATTTAGGTGACACTATAGAAGCCTCTCGCAACCAAAGCCTCT-3'

In situ hybridization was performed as described by Thisse and Thisse [48]. Briefly, the fixed embryos were rehydrated and permeabilized with 10 µg/mL proteinase K for 30 s (1-cell stage embryos or 10 min (24 and 48 hpf embryos) at room temperature. Ten to 15 embryos from each time point were hybridized with digoxigenin-labeled antisense RNA probes at 70 °C. After extensive washing, the probes were detected with anti-digoxigenin-AP Fab fragment (Roche Diagnostics, 1093274, diluted at 1:10,000), followed by staining with BCIP/NBT (5-bromo-4-chloro-3-indolyl-phosphate/nitro blue tetrazolium) alkaline phosphate substrate. The embryos were imaged using a Leica MZ10F stereomicroscope equipped with a Leica DFC295 digital camera.

**RNA Extraction and RT-PCR**

Total RNAs were purified from 1 hpf, 3 hpf, 6 hpf, 1 dpf, 2 dpf, 3 dpf, 4 dpf and 5 dpf wild-type embryos or larvae, as well as from adult tissues using Trizol, as previously described [49]. cDNA was synthesized using Superscript III (18080-044, Invitrogen, Carlsbad, CA, USA) following the manufacturer's instructions. A quantity of 1 µg of total RNAs were used to perform the reverse transcription experiments. Primers used were:

cbx2_984F: 5'-TCAAAGGTCAGCATCTCTCCCGA-3'	cbx2_1111R: 5'-
TCAAATCAGTCCCCTTCCCTGC-3'	cbx4a1_656F: 5'-
GGAACCTCCCTCCAGCCCTAC-3'	cbx4a1_785R: 5'-
AGTGTCGCTCTGCTTTGTTAGC-3'	cbx6a2_1295F: 5'-
CCGATTGGCACCCCTGAAATGGC-3'	cbx6a2_1461R: 5'-
TGCAGTTGTGGTGGGAGAGTTTGG-3'	cbx6b_1275F: 5'-
ACCAGAGGTGCCCCGTGTCAGA-3'	cbx6b_1437R: 5'-
AGAAGAAGCAGAAGGCGGATGGC-3'	cbx7_1691F: 5'-
CCTCACCTGCTTGTCTGTAAACCA-3'	cbx7_1933R: 5'-
AGGCTCGGCTTTGATTTCTGCC-3'	cbx7b_304F: 5'-
GCTCATCTGCCACTCTCGCTGG-3'	cbx7b_428R: 5'-
TCCCTTCTGTCCCCTCTTCCTCA-3'	cbx8a_831F: 5'-
GGTTGGCTCTGGTGGTGCCATAA-3'	cbx8a_977R: 5'-
TTCAAACAAGGCAGCCAGGACTG-3'	cbx8b_787F: 5'-
AGCAACGTCATTCCAAGCAGGAGT -3'	cbx8b_1045R: 5'-
AGGGTCTCCAAGCATCTTCCACC -3'	Dr_cDNA_ube2a_F: 5'-
TGACTGTTGACCCACCTTACAG-3'	
Dr_cDNA_ube2a_R: 5'-CAAATAAAAGCAAGTAACCCCG-3'	
Dr_beta-actin_cDNA_5: 5'-CGTGACATCAAGGAGAAGCT-3'	
Dr_beta-actin_cDNA_3: 5'-ATCCACATCTGCTGGAAGGT-3'	

PCR reactions were performed as follow: 95 °C 4 min, (95 °C 45 s, 65 °C 45 s, 72 °C 1 min) 35 cycles, 72 °C 10 min.

RT-PCR experiments were performed from at least two independent RNA extraction samples.

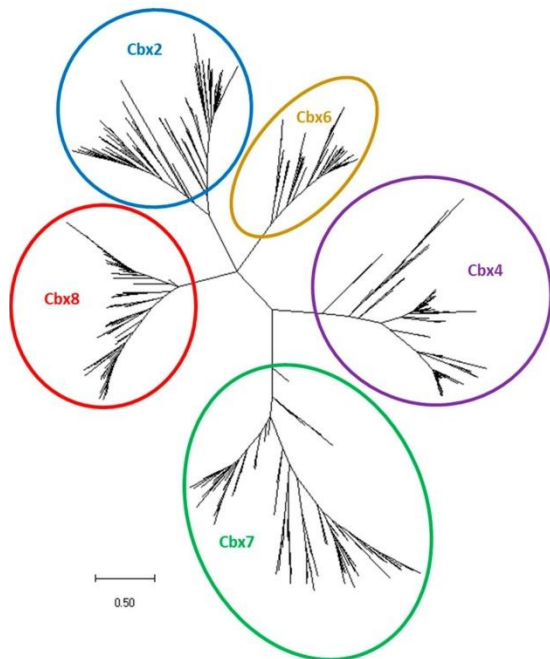
**Results**

***In Silico Identification of Teleost Pc Orthologues***

Polycomb (Pc) and its CBX mammalian orthologues interact with H3K27me3 epigenetic marks via the chromodomain (SMART, SM00298; Pfam, PF00385) [34], whereas the Pc box at the C-terminus (Pfam, PF17218) is involved in transcriptional silencing and binding to other PRC1 components such as RNF2 [50]. Adjacent to the chromodomain, all CBX protein have a DNA binding motif, the AT-hook (SMART, SM000384; Pfam, PF02178) (in CBX2) or an AT-hook-like motif (in the other four CBX proteins, CBX4, CBX6, CBX7 and CBX8) [51]. Less conserved sequences in the middle of the CBX proteins could play a role in specifically directing each CBX family member to distinct regions of the chromatin [51,52].

We applied TBLASTN searches in the NCBI and Ensembl databases using the human and zebrafish Pc orthologues as queries, to identify the corresponding genes in 67 teleost species (Table A1). A total of 689 genes from 68 teleosts including zebrafish, were identified. Based on sequence alignments and phylogenetic analyses (Figure 1, Supplementary File S1, Supplementary File S2), these Pc orthologues were associated to one of the paralogous class CBX2, CBX4, CBX6, CBX7 or CBX8 (Supplementary Table S1). In contrast to Psc/PCGF and Sce/RING1 orthologues that are not all present in fishes including zebrafish and medaka [37], the five *cbx* paralogues, *cbx2*, *cbx4*, *cbx6*, *cbx7*

and *cbx8*, are maintained in all sequenced teleost genomes. In many cases, but not systematically, the *cbx* genes are even retained as pairs of ohnologues (Supplementary Table S1).



**Figure 1.** Unrooted phylogenetic tree of the evolutionary relationships between teleost Pc orthologues. The tree generated using the Maximum Likelihood methods and JTT matrix-based model, with the highest log likelihood (-1307026.38) is shown. The tree is drawn to scale, with branch lengths measured in the number of substitutions per site. This analysis involves 558 amino acid sequences (116 for Cbx2, 129 for Cbx4, 72 for Cbx6, 117 for Cbx7 and 124 for Cbx8) with a total of 1529 positions in the final dataset. Evolutionary analyses were conducted in MEGA X.

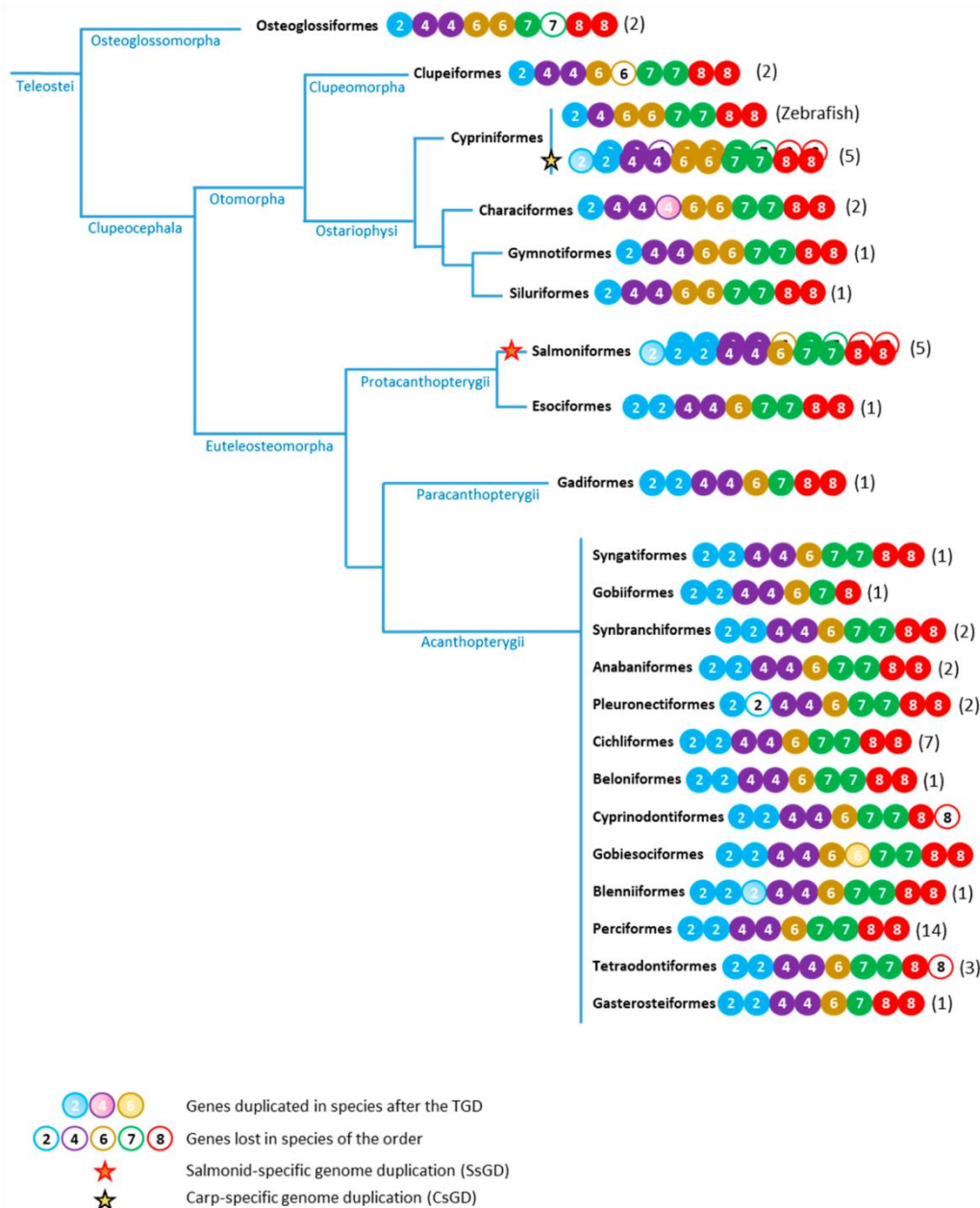
To investigate the evolution of the *cbx* gene content in teleosts, a cladogram of the teleost fishes [53–56] showing the *cbx* gene content in different fish orders, has been drawn (Figure 2). From this representation, it appears that *cbx7* and *cbx8* are present as pairs of ohnologues in teleost. In contrast, *cbx2* and *cbx6* are retained as singleton in the genome of particular clades of teleost fishes. Following the TGD, one of the two *cbx2* ohnologues has been lost in Osteoglossomorpha and Otomorpha, whereas one *cbx6* ohnologue has been lost in Euteleostei. Finally, the loss of a *cbx4* ohnologue occurred later in teleost evolution since it just concerns fishes from the order of the Cypriniformes and zebrafish is so far the only teleost containing *cbx4* as a singleton in its genome.

#### *The cbx6 Parologue in Teleost Fishes*

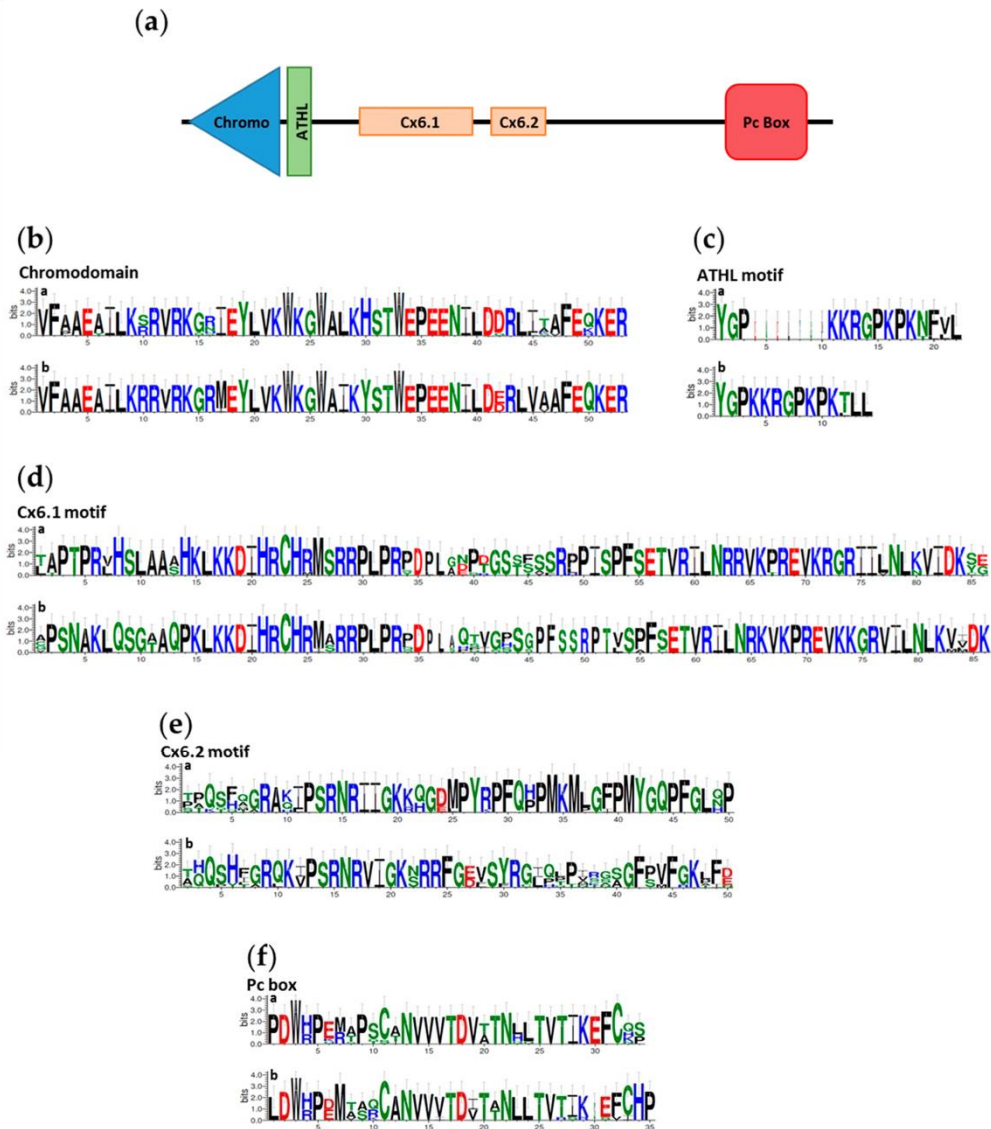
Our phylogenetic analysis shows that in Clupeomorpha and Ostariophysii fishes, *cbx6* is found as a pair of ohnologues hereafter called *cbx6a* or *cbx6b* (Supplementary Figure S1). While Cbx6a and Cbx6b belong to distinct branches of the phylogenetic Cbx6 tree, protein sequence alignments showed that Cbx6a and Cbx6b share high levels of similarity within conserved motifs such as the chromodomain, the ATHL motif, the Cx6.1 motif and the Pc box (Figure 3, Supplementary Figure S2). In contrast, the Cx6.2 motif, as well as sequences outside of the conserved motifs, present lower levels of similarity between Cbx6a and Cbx6b proteins.



## PARTIE I : LA REPRESSION POLYCOMB



**Figure 2.** Simplified cladogram of the teleost fishes showing the content of the Pc/cbx genes in different orders. *Cbx2*, *cbx4*, *cbx6*, *cbx7* and *cbx8* genes are shown as solid circle blue, violet, yellow, green and red, respectively. Empty circles indicate that a given *cbx* gene is absent in at least one species in the order. Circles filled with clear colors show duplicated genes in at least one species in the order. Stars indicate a recent whole genome duplication in the salmoniformes (red star—salmonid-specific 4th round of genome duplication, SsGD) and in the carp lineage of cypriniformes (black star—carp-specific genome duplication, CsGD). The number of species studied in each order is in brackets.

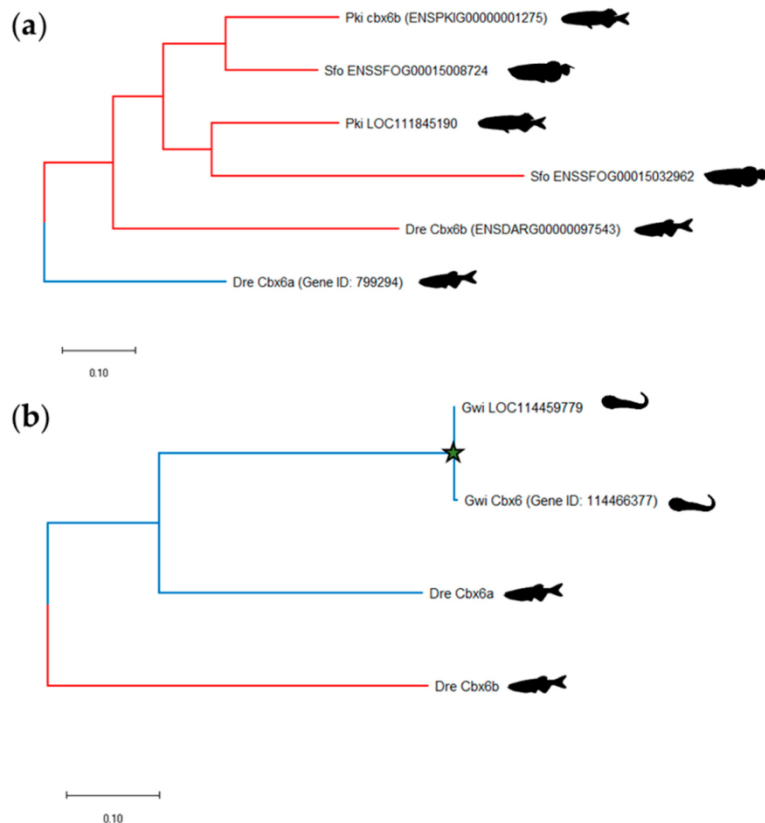


**Figure 3.** The Cbx6 conserved protein motifs. (a) Schematic representation of Cbx6 and its conserved regions. The sequences of the chromodomain (b), the ATHL (c), Cx6.1 (d), Cx6.2 (e) motifs and the Pc box (f) is shown for the Cbx6a (a: up) and Cbx6b (b: down) proteins. The height of amino acids is proportional to the frequency of occurrence (degree conservation) in each position and drawn in a logo format. Basic amino acids are in blue, hydrophobic in black, polar in green and acidic in red.

Previous synteny analyses at genomic levels in zebrafish and tetraodon showed that the chromosomal correspondence of duplicated gene pairs arising from the TGD has been extensively preserved in doubly conserved synteny blocks, while local gene order has been largely scrambled [57,58]. Examination of the genes surrounding the *cbx6* ohnologues reveals that *cbx6a* is flanked at its 5<sup>0</sup> by *pla2g6* coding for the phospholipase A2 group VI in Ostariophysi fishes from the orders Cypriniformes (zebrafish, goldfish), Characiformes (Mexican cavefish), Siluriformes (channel catfish) and Gymnotiformes (electric eel), as well as in Clupeomorpha fishes (Atlantic herring) (Supplementary Figure S3). In contrast, the *baiap212* gene encoding the BAR/IMD domain containing adaptor protein 2 like 2 is located at the 5<sup>0</sup> position of *cbx6b*, thus defining distinct doubly conserved synteny blocks for *cbx6a* and *cbx6b*.

Among the Osteoglossomorpha, the elephant fish (*Paramormyrops kingsleyae*) and the Asian bonytongue (*Scleropages formosus*) are so far the only two species that have fully sequenced genomes [59,60]. Like Ostariophysi and Clupeomorpha, the *cbx6* gene exists as pairs of ohnologues in the elephant fish and in the Asian bonytongue

genomes (Figure 2, Supplementary Table S1). Phylogenetic analyses revealed that the Osteoglossomorpha Cbx6 ohnologues are more closely related to each other and to the Ostariphysi and Clupeomorpha Cbx6b rather than to Cbx6a (Figure 4A, Supplementary Figure S1). However, syntenic studies identify that the Osteoglossomorpha *cbx6* ohnologues are located within the same distinct doubly conserved synteny blocks as those harboring *cbx6a* and *cbx6b* in Ostariphysi and Clupeomorpha (Supplementary Figure S3). Thus, consistent with the teleost cladogram, the two Osteoglossomorpha *cbx6* ohnologues originate from the TGD but are less divergent to each other than *cbx6a* to *cbx6b*.



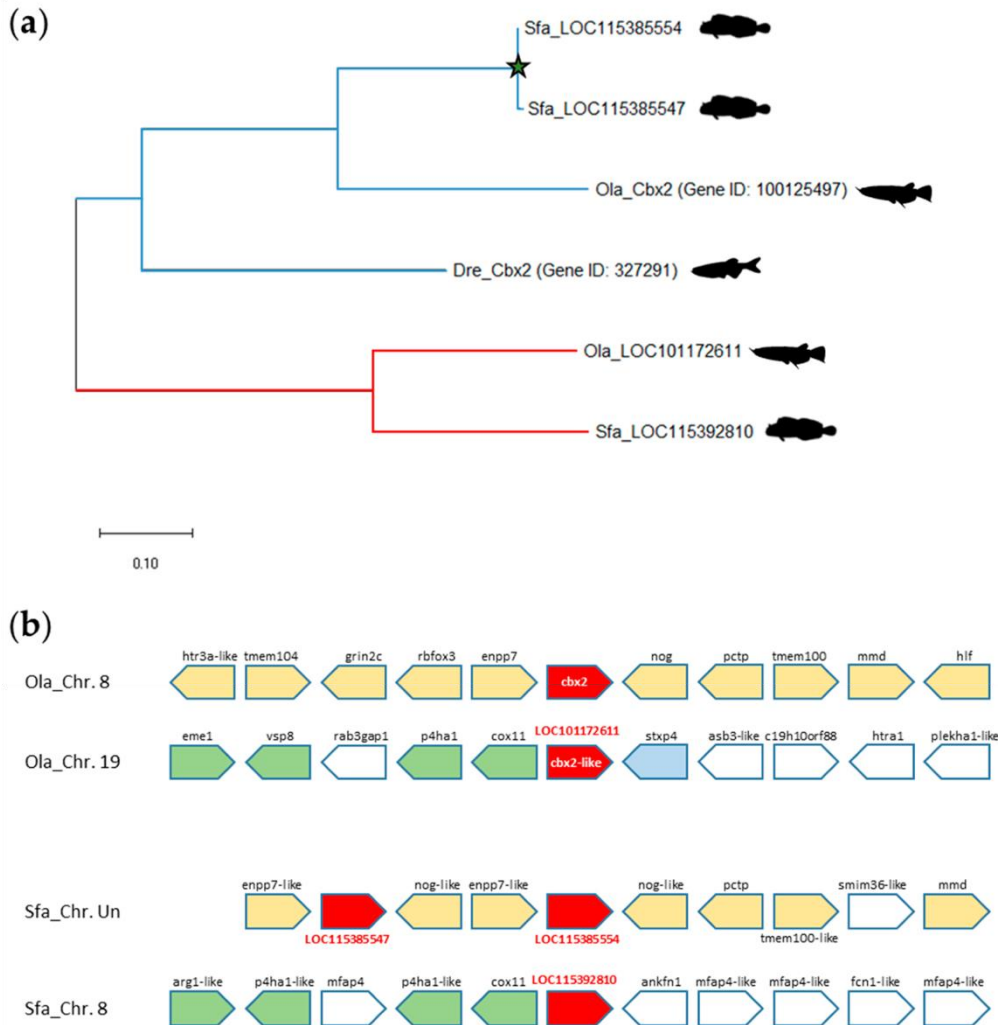
**Figure 4.** Phylogenetic trees of the evolutionary relationships between Cbx6 proteins from teleost fishes. (a) Phylogenetic tree of the evolutionary relationships between zebrafish (*Dre*, *Danio rerio*) Cbx6a and Cbx6b proteins with their orthologues in elephant fish (*Pki*, *Paramormyrops kingsleyae*) and Asian bonytongue (*Sfo*, *Scleropages formosus*). (b) Phylogenetic tree of the evolutionary relationships between zebrafish Cbx6 proteins and their orthologues in blunt-snouted clingfish (*Gwi*, *Gouania willdenowi*). The evolutionary history was inferred using the Neighbor-Joining method conducted in MEGA X. The star illustrates the recent duplication of the *cbx6* gene in *Gouania willdenowi*.

The study of the *cbx6* paralogue content in teleost also revealed that *cbx6* is present as a singleton in Euteleostomorpha fishes including Protacanthopterygii, Paracanthopterygii and Acanthopterygii (Figure 2, Supplementary Table S1), indicating that one *cbx6* ohnologue arising from the TGD has been lost in this fish clade. The only Acanthopterygii that has its genome sequenced and possesses a pair of *cbx6* genes is the blunt-snouted clingfish (*Gouania willdenowi*, order Gobiesociformes). However, both the phylogenetic analyses (Figure 4B) and the syntenic studies indicating that one *cbx6* is located in a highly rearranged region of Chromosome 1 (Supplementary Figure S3), suggesting that the second copy of *cbx6* arose from a recent duplication event.

#### The *cbx2* Paralogue in Teleost Fishes

A search for the *cbx2* paralogues in the sequenced teleost genomes revealed that *cbx2* is present as a singleton in Osteoglossomorpha and Otomorpha fishes, but exists as a pair of ohnologues in Euteleostomorpha

(Figure 2, Supplementary Table S1). This suggests that after the TGD leading to the duplication of *cbx2*, one of the ohnologues was lost at least twice independently in the Osteoglossomorpha and in the Otomorpha fish lineages. Within the Acanthopterygii, the jewelled blenny (*Salarias fasciatus*, order Blenniformes) possesses three copies of *cbx2* in its genome (Supplementary Table S1). However, phylogenetic analyses revealed that two of these copies (LOC115385554 and LOC115385547) are highly similar and might arise from a recent gene duplication event (Figure 5A).



**Figure 5.** The *cbx2* paralogues in the jewelled blenny (*Salarias fasciatus*). **(a)** Phylogenetic tree of the evolutionary relationships between zebrafish (*Dre*, *Danio rerio*) Cbx2 protein and its orthologues in medaka (*Ola*, *Oryzias latipes*) and jewelled blenny (*Sfa*, *Salarias fasciatus*). The evolutionary history was built using the Neighbor-Joining method conducted in MEGA X. **(b)** Blocks of synteny at the *cbx2* locus in medaka and jewelled blenny. The *cbx2* genes are shown in red, while orthologous gene loci are highlighted with different colors. Chromosomal locations of *cbx2* and the neighboring gene loci were drawn based on information from the NCBI Gene server. Chromosome (Chr.) numbers are indicated. Un, unplaced scaffold.

The synteny studies showed that the *cbx2* ohnologues are associated to distinct doubly conserved synteny blocks in Acanthopterygii including in the medaka fish model (Figure 5B, Supplementary Figure S4). In jeweled blenny, one of these blocks was being subjected to rearrangements generating a second *cbx2* gene within the synteny block. Thus, it is likely that the additional *cbx2* copy in the jeweled blenny genome results from a recent intrachromosomal rearrangement. Similar rearrangements are also found in species of the Cypriniformes and Salmoniformes orders.

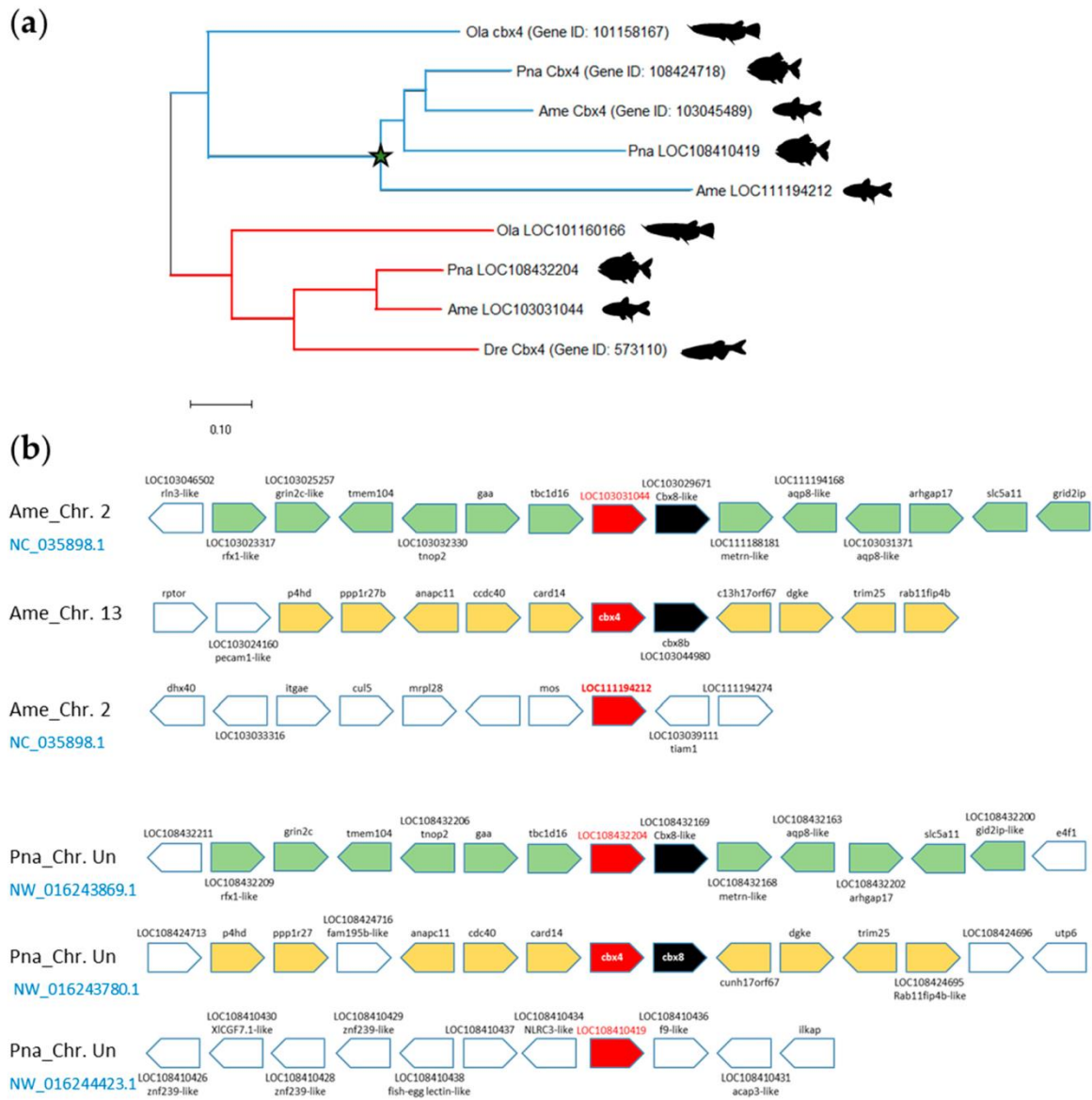
Cypriniformes is the largest group of freshwater fishes comprising about 4300 described species including zebrafish, goldfish, carps, barbels, minnows, loaches and suckers [61]. Among them, several species were subject to an additional whole-genome duplication event called the carp-specific genome duplication (CsGD). Indeed, while the zebrafish diploid genome is composed of 50 chromosomes, the diploid status of common carp (*Cyprinus carpio*), goldfish (*Carassius auratus*) and Chinese barbels (*Sinocyclocheilus anshuiensis*, *S. grahami* and *S. rhinocerosus*) is supported by a double-size karyotype consisting of approximately 100–104 chromosomes [62–64]. In zebrafish, *cbx2* is present as a singleton and consistent with the CsGD event, there are two *cbx2* gene copies in common carp and Chinese barbels (Supplementary Table S1). However, in goldfish, an additional *cbx2* copy is identified and genomic analyses showed that two goldfish *cbx2* genes (LOC113067828 and LOC113067392) are located in the same linkage group (LG28B, position NC\_039293.1) (Supplementary Figure S4) as the result of an intrachromosomal rearrangement. Similarly, within the Protacanthopterygii, the northern pike (*Esox lucius*, order Esociformes) has a karyotype composed of about 50 chromosomes [65]. An additional whole-genome duplication event, called the salmonid-specific genome duplication (SsGD), occurred in the common ancestor of Salmoniformes after their divergence from Esociformes [66,67]. Then, while the Northern pike genome harbors a pair of *cbx2* ohnologues, most of the Salmonidae including rainbow trout (*Oncorhynchus mykiss*), Coho salmon (*Oncorhynchus kisutch*), Arctic char (*Salvelinus alpinus*) and huchen (*Hucho hucho*) host four *cbx2* genes in their genome because of the SsGD (Supplementary Table S1). However, the Atlantic salmon (*Salmo salar*) possesses a fifth *cbx2* copy in its genome. Remarkably, two of these Atlantic salmon *cbx2* genes (LOC106606958 and LOC106606946) are located on the same chromosome (ssa06, position NC\_027305.1) (Supplementary Figure S4), suggesting that the additional *cbx2* copy appeared due to a recent chromosomal rearrangement.

### *The cbx4, cbx7 and cbx8 Paralogues in Teleost Fishes*

Our TBLASTN searches in the sequenced teleost genomes showed that the three Pc orthologues *cbx4*, *cbx7* and *cbx8* are present as pairs of ohnologues in all teleost clades (Supplementary Table S1, Figure 2). One possible exception could concern the *cbx7* paralogue in Paracanthopterygii. Within the Paracanthopterygii super-order, Atlantic cod (*Gadus morhua*, order Gadiformes) is the only specie for which whole-genome sequencing data are available (Ensembl, assembly gadMor1) and in its genome, *cbx7* is found as a singleton. However, a cod-specific loss of one *cbx7* ohnologue due to a particular chromosomal rearrangement in this specie cannot be ruled out. Alternatively, the failure in the identification of the second *cbx7* homologue could be the result of a technical artefact such as sequencing or gene prediction information missing. Then, a definitive conclusion about the presence of *cbx7* as a singleton or a pair of ohnologues in Paracanthopterygii could be made when genomic information will be available for other species of the clade.

The *cbx4* paralogue is even present as three copies in the Mexican cavefish (*Astyanax mexicanus*) and red-bellied piranha (*Pygocentrus nattereri*) genomes, two species of the teleost order Characiformes. Genomic analysis reveals a remarkable conservation of *cbx4* chromosomal organization (Supplementary Figure S5). First, *cbx4* is located immediately 5' to *cbx8*. This feature is not specific to teleost, since it is also the case in mammals including in human [37]. Second, the *cbx4-cbx8* ohnologous loci are associated to two recognizable — with distinct signatures — but conserved syntenic blocks. The first *cbx4-cbx8* locus is flanked by the genes *card14* (caspase recruitment domain family member 14) and *dgke* (diacylglycerol kinase epsilon), whereas the other is flanked by the genes *tbc1d16* (TBC1 domain family member 1) and *arhgap17* (Rho GTPase activating protein 17). The phylogenetic analysis of the Mexican cavefish and red-bellied piranha *cbx4* paralogues suggest that two copies of *cbx4* arise from the duplication of the same *cbx4* orthologue in the two Characiformes species (Figure 6A). In the Mexican cavefish, *cbx4* (Gene ID: 10304589) and LOC111194212 derive from the same ancestor gene, while in red-bellied piranha, *cbx4* (Gene ID: 108424718) and LOC108410419 also originate from the same ancestor.





**Figure 6.** The *cbx4* paralogs in Characiformes. (a) Phylogenetic tree of the evolutionary relationships between zebrafish (*Dre*, *Danio rerio*) *Cbx4* protein and its orthologues in medaka (*Ola*, *Oryzia latipes*), Mexican cavefish (*Ame*, *Astyanax mexicanus*) and red-bellied piranha (*Pna*, *Pygocentrus nattereri*). The evolutionary history was built using the Neighbor-Joining method conducted in MEGA X. The star indicated a possible *cbx4* gene duplication having occurred in Characiformes. (b) Blocks of synteny at the *cbx4-cbx8* locus in Mexican cavefish and red-bellied piranha. The *cbx4* genes are shown in red, *cbx8* in black, while orthologous gene loci are highlighted with different colors. Chromosomal locations of *cbx4*, *cbx8* and the neighboring gene loci were drawn based on information from the NCBI Gene server. Chromosome (Chr.) numbers are indicated. Un, unplaced scaffold.

Remarkably, in Mexican cavefish, *cbx4* is associated to the *card14-dgke* synteny block, whereas LOC111194212 local chromosomal organization totally lack the characteristic of the conserved *cbx4-cbx8* synteny blocks (Figure 6B). Similarly, red-bellied piranha, *cbx4* is associated to the *card14-dgke* synteny block whereas LOC108410419 is in a gene region without homology with the conserved *cbx4-cbx8* synteny blocks. The similar properties of the three *cbx4* copies in Mexican cavefish and red-bellied piranha suggest that the additional *cbx4* copy in these species originate from a single *cbx4* duplication event having occurred in the Characiformes before their speciation.

In contrast to Characiformes having three *cbx4* copies, zebrafish contains a single *cbx4* copy in its haploid genome. The loss of one *cbx4* ohnologue in zebrafish is specific and restricted since it does not affect the neighboring *cbx8* gene nor the other genes of the syntenic block (Supplementary Figure S5). Furthermore, this *cbx4* ohnologue loss occurred in the zebrafish lineage after it diverged from the carp lineage. Indeed, the goldfish and the three Chinese barbels from the genus *Sinocyclocheilus* all contain four *cbx4* gene copies consequential to a pair of ohnologues subjected to the CsGD (Supplementary Table S1, Supplementary Figure S5).

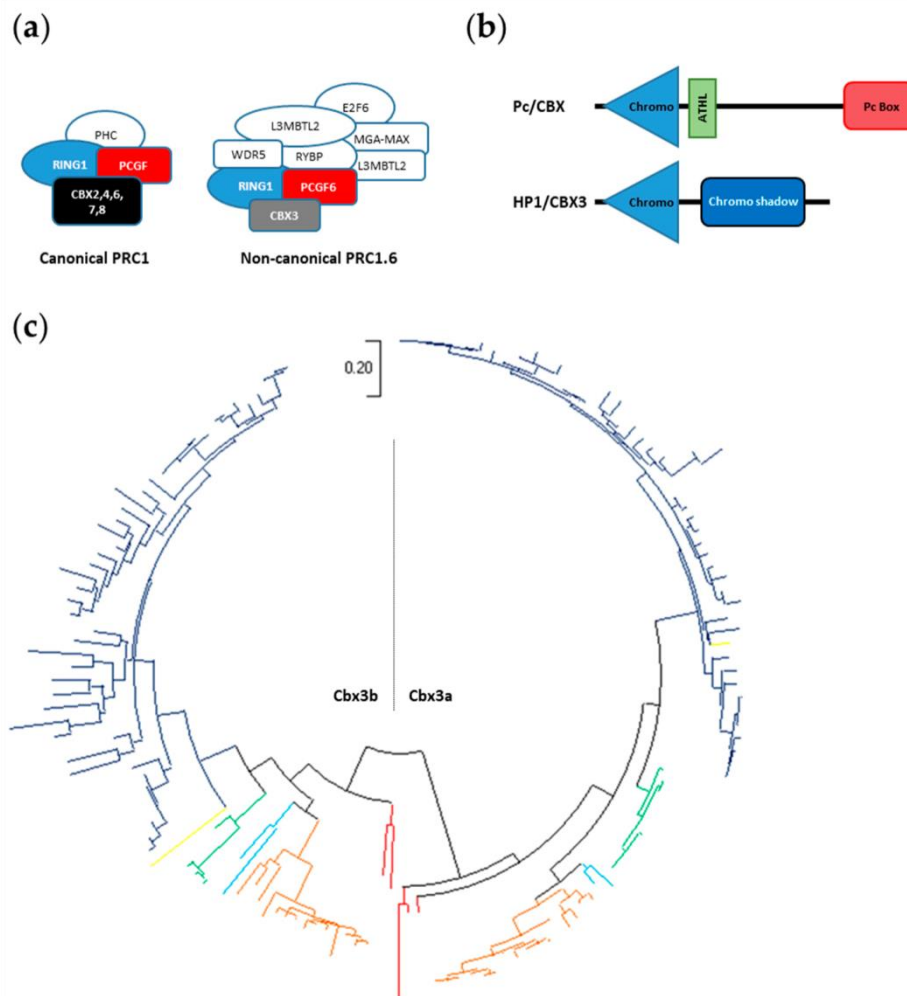
### *The cbx3 Genes in Teleost Fishes*

The canonical PRC1 (cPRC1) complex is the functional homologues of *Drosophila* PRC1 composed of Pc, Psc, Ph and Sce. However, in mammal, a heterogeneous group of non-canonical PRC1 (ncPRC1) complexes have also been described [68–70]. These ncPRC1 are characterized by the absence of the Pc orthologues CBX2, CBX4, CBX6, CBX7 and CBX8, but the presence of YY1-binding protein (RYBP), or its homolog YAF2 associated to RING1/RNF2, one of the PCGF proteins and various other subunits. Among the different ncPRC1 complexes, PRC1.6 (also named E2F6.com, [71,72]) (Figure 7A) is composed of the transcriptional repressor E2F6 in association with RNF2-PCGF6-RYPB/YAF2. In addition, the complex contains WDR5, the oncoprotein L3MBTL2, the transcription factors MAX and MGA and the chromodomain-containing protein CBX3. In contrast to CBX2, CBX4, CBX6, CBX7 and CBX8 which are orthologous to the *Drosophila* Pc protein, CBX3 is an orthologue of the *Drosophila* protein HP1 / Su(var)205. In addition, Pc orthologues contain a chromodomain, an AT hook/ATHL motif and a Pc box, whereas CBX3 is composed of a chromodomain associated to a chromo shadow domain (Pfam ID: PF00385) (Figure 7B). Finally, while the chromodomain of Pc and its orthologues recognize H3K27me3, the CBX3 chromodomain preferentially binds to the H3K9me3 epigenetic marks.

Although PRC1.6 function is not known and there is no evidence showing that PRC1.6 exists in teleost fishes, we conducted a TBLASTN search to identify CBX3 orthologues in fishes. From the NCBI and Ensembl databases, we identified 151 *cbx3* genes in the 68 teleost species covered by this study (Supplementary Table S2, Supplementary File S3). Phylogenetic analyses show that *cbx3* is present as a pair of ohnologues *cbx3a* and *cbx3b* in all the teleost clades studied, Osteoglomorpha, Clupeomorpha, Ostariophysii, Protacanthopterygii, Paracanthopterygii and Acanthopterygii (Figure 7C, Supplementary File S3, Supplementary File S4). Then, like *cbx4*, *cbx7* and *cbx8*, *cbx3* has been retained as two gene copies in the teleost genome after the TGD.

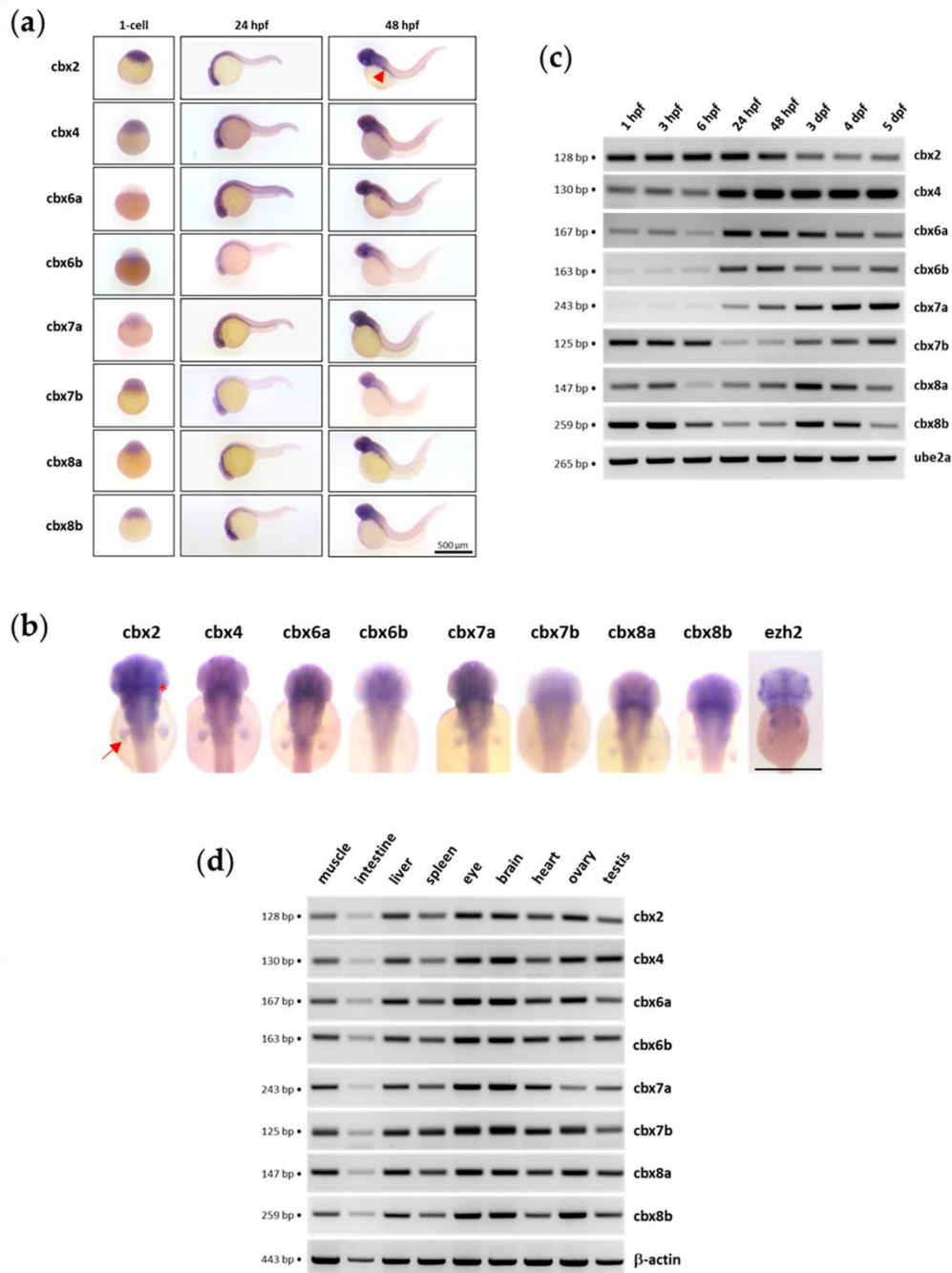
### *Pc Gene Expression in the Zebrafish Model*

The zebrafish genome encodes eight Pc orthologues. These *cbx* paralogues are *cbx2*, *cbx4*, *cbx6a*, *cbx6b*, *cbx7a*, *cbx7b*, *cbx8a* and *cbx8b*. Since zebrafish serves as a powerful vertebrate model for studying gene expression during early development and modeling human diseases, we have investigated gene expression profiles for the eight *cbx* paralogues in this organism. In zebrafish, zygotic transcription starts at about cell cycle 10–13 (around 3.5 hours post fertilization (hpf)). Before this midblastula transition (MBT) stage, all developmental events rely on maternally deposited gene products [73,74]. The study of *cbx* expression patterns before the MBT at the 1-cell stage, as well as after the MBT at 24 and 48 hours post-fertilization (hpf), using whole-mount in situ hybridization revealed that all the *cbx* family members are globally ubiquitously expressed (Figure 8A, Supplementary File S5). In situ hybridization showed that the *cbx* transcripts are maternally loaded into the embryos since a signal could be detected before MBT at the 1-cell stage, even if the labelling remains quite low for *cbx6a*, *cbx6b* and *cbx7a*. At 24 hpf, *cbx* mRNAs are ubiquitously present in the embryo. The expression becomes more restricted in the developing brain, the gut and in the pectoral fin buds at 48 hpf. However, some differences in the *cbx* expression could be observed. At 48 hpf, the expression pattern of *cbx2* is similar to the *ezh2* expression profile with a marked signal at the midbrain–hindbrain boundary and in the pectoral fin buds, whereas *cbx6b* or *cbx7b* show a more diffuse labelling in the brain and a weaker signal in the pectoral fin buds (Figure 8B).



**Figure 7.** Cbx3 in teleost fish. (a) Schematic representation of the canonical PRC1 complex (left) and of the non-canonical PRC1.6 complex (right). (b) Schematic representation of the Pc/CBX (CBX2, CBX4, CBX6, CBX7 and CBX8) proteins with the conserved chromodomain, ATHL motif and Pc box (up) and of the HP1/CBX3 protein with the conserved chromodomain and chromo shadow domain (down) (c) Phylogenetic tree of the evolutionary relationships between the teleost Cbx3 proteins. The tree generated using the Maximum Likelihood methods and JTT matrix-based model, with the highest log likelihood (-12504.92) is shown. The tree is drawn to scale, with branch lengths measured in the number of substitutions per site. This analysis involves 150 amino acid sequences with a total of 316 positions in the final dataset. Evolutionary analyses were conducted in MEGA X. Both Cbx3a and Cbx3b are present in Osteoglossomorpha (red branches), Clupeomorpha (light-blue branches), Ostariophysii (brown branches), Protacanthopterygii (green branches), Paracanthopterygii (yellow branches) and Acanthopterygii (dark-blue branches).





**Figure 8.** Expression of the *cbx* family members in zebrafish. (a) In situ hybridization at the 1-cell stage, 24 hpf and 48 hpf showing maternally provided transcripts (1-cell stage) and zygotic mRNA distribution at 24 and 48 hpf. The red arrowhead shows the gut. (b) Dorsal view of the anterior region of embryos at 48 hpf after in situ hybridization. The arrow shows the pectoral fin buds and the asterisk the midbrain-hindbrain boundary. Scale bar is 500  $\mu$ M. (c) RT-PCR experiment showing the detection of *cbx2* (n = 3), *cbx4* (n = 3), *cbx6a* (n = 2), *cbx6b* (n = 2), *cbx7a* (n = 3), *cbx7b* (n = 4), *cbx8a* (n = 3) and *cbx8b* (n = 3) transcripts at 1 hpf, 3 hpf, 6 hpf, 24 hpf, 48 hpf, 3 dpf, 4 dpf and 5 dpf. Ube2a is used as a control. The size of the amplicons is indicated. (d) RT-PCR experiment showing *cbx* transcripts in adult zebrafish tissues (n = 2). Beta-actin is used as a control. The size of the amplicons is indicated. All RT-PCR agarose gels are shown in Supplementary File S6.

Analyses of mRNA abundance measured by RT-PCR showed that *cbx* mRNA levels vary during zebrafish development from 1 hpf to 5 days post-fertilization (dpf) (Figure 8C). In particular, *cbx4*, *cbx6a*, *cbx6b* and *cbx7a* expression is reduced at early stages and increases after 6 hpf, whereas for *cbx7b*, *cbx8a* and *cbx8b*, a decrease in

mRNA levels is found between 6 and 48 hpf. A delay between the degradation maternal mRNAs occurring at MBT and the start of zygotic expression of these genes might account for the reduction in mRNA abundance between 6 and 48 hpf. It is worth noting that these variations in *cbx* mRNA abundance parallel those reported using large-scale RNA-Seq experiments during zebrafish development [75]. Finally, in adult zebrafish, our RT-PCR experiments showed that all *cbx* family members are expressed ubiquitously (Figure 8D).

### Discussion

The Pc/CBX family member proteins are components of the canonical PRC1 protein complex that maintain transcriptional repression of hundreds of genes involved in development, differentiation, signaling or cancer. Since these proteins directly bind to the epigenetic mark H3K27me<sub>3</sub>, they are key elements targeting the PRC1 complex to its chromatin sites. While *Drosophila melanogaster* contains a single Pc protein, mammalian genomes code for five CBX paralogues, CBX2, CBX4, CBX6, CBX7 and CBX8. This diversity might be crucial for the control of gene expression programs since it is believed that changes in the CBX protein content within the PRC1 complexes could relocalize the PRC1 to different target genes during the differentiation processes.

Here, we performed a global analysis of the *cbx* gene family in teleost fishes and traced the distribution of the *cbx* genes through teleost evolution. Teleost fish experienced at least three rounds of whole-genome duplication; the first two before the divergence of lamprey from the jawed vertebrates, and a third teleost-specific TGD at the base of the teleosts [58,76,77]. The TGD was followed by a rediploidization process associated to a massive loss of duplicated genes. However, a number of genes are still maintained as pairs of ohnologues in teleost genomes. Consequently, it is expected that the number of *cbx* family members will be higher in teleost fishes than in mammals. Database searches identified 689 *cbx* genes in 68 teleost species that belong to 21 different orders and 6 different super-orders (Osteoglossomorpha, Clupeomorpha, Ostariophysi, Protacanthopterygii, Paracanthopterygii and Acanthopterygii). The five *cbx* paralogues, *cbx2*, *cbx4*, *cbx6*, *cbx7* and *cbx8* were found in all fish teleost species. This is in total contrast with the situation observed for other components of the PRC1 protein complex [37]. Indeed, mammalian genomes contain six Psc/PCGF paralogues, *PCGF1*, *PCGF2*, *PCGF3*, *BM11*, *PCGF5* and *PCGF6*, whereas both *pcgf2* and *pcgf3* are absent in zebrafish and *pcgf1* is absent in medaka. Similarly, there are two Sce/RING1 paralogues in mammals, *RING1* and *RNF2*, while *rnf2* is the only paralogue present in zebrafish. The maintenance of the five *cbx* paralogues in teleost could suggest the absence of functional redundancy between the Cbx family members, whereas Pcgf members could be redundant. The hypothesis of a Pcgf redundancy in teleost fishes is in agreement with the absence of certain *pcgf* gene members in the genome of several fish species including zebrafish and the medaka [37] and with the fact that *pcgf1* zebrafish mutants are viable and fertile [78]. However, the situation might be different in mammals where PCGF functions are definitively not redundant [79,80].

Surprisingly, although about 15% to 20% of the ohnologues were retained as pairs after the TGD in teleost [9], the maintenance of two ohnologues is globally a general characteristic of all *cbx* family members. *Cbx8* is present as a pair of ohnologues (*cbx8a* and *cbx8b*) in all the teleost clades examined, as it is the case for *cbx4* and probably also for *cbx7*. *Cbx7* is identified as a pair of ohnologues in all clades except in Paracanthopterygii. However, the Atlantic cod is the only specie from this super-order having its genome sequence available. It is thus difficult to conclude whether the loss of one of the two *cbx7* ohnologues is specific to the Atlantic cod or whether it reflects a feature common to all Paracanthopterygii. In contrast, *cbx2* and *cbx6* are present as pairs of ohnologues in the genome of several teleost clades but as singletons in others. Following the TGD, one *cbx6* ohnologue has been lost in the common ancestor of Euteleostomorpha. Concerning *cbx2*, one of the two ohnologues has been lost in Osteoglossomorpha and Otomorpha, but retained as two copies in Euteleostomorpha. This suggests that the loss of *cbx2* ohnologues occurred at least twice, in Osteoglossomorpha and in Otomorpha, during teleost evolution. The reason why *cbx4*, *cbx7*, *cbx8*, but also *cbx3* are retained as pairs of ohnologues in the teleost genomes, whereas *cbx2* and *cbx6* remain as singletons in different teleost clades is not clear. One possibility could be linked to a possible neofunctionalization and/or subfunctionalization of one of the *cbx3*, *cbx4*, *cbx7*, *cbx8* ohnologues, but it could also be due to constraints applied by the presence/absence of other genes in the blocks of synteny.

Our analysis of the teleost *cbx* gene family also shed light on other rearrangements having occurred later in particular teleost lineages. For instance, a third copy of *cbx2*, probably arising from an intrachromosomal rearrangement, is identified in jewelled blenny (*Salarias fasciatus*, order Blennioformes). Notably, an additional *cbx4* copy also originate from a single *cbx4* duplication event having occurred in the Characiformes before their speciation since these three *cbx4* copies are found in the genomes of the two Characiformes, the Mexican cavefish (*Astyanax mexicanus*) and the red-bellied piranha (*Pygocentrus nattereri*), for which genomic sequences are available.

In this landscape, the genomic *cbx* gene content in zebrafish, with *cbx2* and *cbx4* present as singletons but *cbx6*, *cbx7* and *cbx8* retained as pairs of ohnologues, appears as an original combination, unique among the teleost fishes having their genome sequenced.

Zebrafish has proven being a unique vertebrate model for studying Polycomb group (PcG) genes during early development [81–84]. In this context, the description of PcG gene expression during development is of particular interest. The expression of the Psc/PCGF family as well as several other PcG genes has already been reported [78,85], but very little is known about Pc/CBX gene expression during zebrafish development. Whole-mount in situ hybridization and RT-PCR experiments showed that zebrafish *cbx* genes are maternally expressed in the embryo at diverse levels. In particular, *cbx4*, *cbx6a* and *cbx7a* mRNAs appear less abundant at the 1-cell stage than at later developmental stages, while it is not the case for *cbx2*, *cbx7b*, *cbx8a* or *cbx8b*. At 24 hpf, *cbx* mRNAs are present in all the embryos but the expression becomes enriched in anterior regions such as the brain, the pectoral fin buds and the gut at 48 hpf. If all the *cbx* genes are expressed in the brain, differences in their expression patterns could be observed using in situ hybridization. Finally, all the *cbx* family members are expressed in adult zebrafish tissues.

In conclusion, our observations contribute to the understanding of Polycomb orthologues evolution in fish and the characterization of *cbx* expression during zebrafish development will be useful to future studies aiming at understanding the functional role of each *cbx* family member.

**Supplementary Materials:** The following are available online at <http://www.mdpi.com/2073-4425/11/4/362/s1>, Table S1: The teleost Pc orthologues, Table S2: The teleost *cbx3* genes, File S1: Rooted version including gene names of the phylogenetic tree of Figure 1, File S2: List of the protein sequences encoded by the teleost Pc orthologues identified in this study, File S3: List of Cbx3 protein sequences encoded by the teleost genomes and identified in this study, File S4: Version of the Cbx3 phylogenetic tree of Figure 7 including the gene symbols, File S5: *in situ* hybridization on groups of zebrafish embryos at the 1-cell stage, 24 hpf and 48 hpf for *cbx2*, *cbx4*, *cbx6a*, *cbx6b*, *cbx7a*, *cbx7b*, *cbx8a* and *cbx8b*, File S6: Original RT-PCR gels, Figure S1: Phylogenetic tree of the evolutionary relationships between teleost Cbx6 paralogues, Figure S2: Sequence alignment for Cbx6 proteins from Ostariophysi and Clupeomorpha fishes, Figure S3: Blocks of synteny at the *cbx6* locus in teleost fishes, Figure S4: Blocks of synteny at the *cbx2* locus in teleost fishes, Figure S5: Blocks of synteny at the *cbx4-cbx8* locus in teleost fishes, Figure S6: Average quantification of RT-PCR amplicons.

**Author Contributions:** Conceptualization, P.-O.A.; Methodology, L.R., P.V. and P.-O.A.; Validation, L.R., P.V. and P.-O.A.; Formal Analysis, L.R., P.V. and P.-O.A.; Investigation, L.R., P.V. and P.-O.A.; Resources, X.L.B. and P.-O.A.; Data Curation, L.R., P.V. and P.-O.A.; Writing – Original Draft Preparation, P.-O.A.; Writing – Review & Editing, L.R., P.V., X.L.B. and P.-O.A.; Supervision, P.-O.A.; Project Administration, P.-O.A.; Funding Acquisition, X.L.B. and P.-O.A. All authors read and approved the final manuscript.

**Funding:** This work was supported by the CNRS, Inserm, the University of Lille, the Ligue Contre le Cancer – Comité du Pas-de-Calais and the GIP Cancéropôle Nord-Ouest. **Conflicts of Interest:** The authors declare no conflict of interest.

## PARTIE I : LA REPRESSION POLYCOMB

### Appendix A

**Table A1.** List of the teleost genomes.

		Family	NCBI Genome	Ensembl ID	Assembly
<b>Aca</b>	Eastern happy ( <i>Astatotilapia calliptera</i> )	Cichlidae	ID: 15018	ENSACAG	fAstCal1.2
<b>Aci</b>	Midas cichlid ( <i>Amphilophus citrinellus</i> )	Cichlidae		ENSACIG	Midas_v5
<b>Ame</b>	Mexican cavefish ( <i>Astyanax mexicanus</i> )	Characidae	ID: 13073	ENSAMEG	Astyanax_mexicanus-2.0
<b>Aoc</b>	Clown anemonefish ( <i>Amphiprion ocellaris</i> )	Pomacentridae	ID: 8181	ENSAOOG	AmpOce1.0
<b>Ape</b>	Orange clownfish ( <i>Amphiprion percula</i> )	Pomacentridae		ENSAPEG	Nemo_v1
<b>Apo</b>	Spiny chromis ( <i>Acanthochromis polyacanthus</i> )	Pomacentridae	ID: 54316	ENSAPOG	ASM210954v1
<b>Ate</b>	Climbing perch ( <i>Anabas testudineus</i> )	Anabantidae	ID: 33761	ENSATEG	fAnaTes1.1
<b>Bsp</b>	Siamese fighting fish ( <i>Betta splendens</i> )	Osphronemidae	ID: 36335		fBetSpl5.2
<b>Cau</b>	Goldfish ( <i>Carassius auratus</i> )	Cyprinidae	ID: 10773		ASM336829v1
<b>Cca</b>	Common carp ( <i>Cyprinus carpio</i> )	Cyprinidae	ID: 10839		common carp genome
<b>Cgo</b>	Channel bull blenny ( <i>Cottoperca gobio</i> )	Bovichtidae	ID: 76087		CotGob3.1
<b>Cha</b>	Atlantic herring ( <i>Clupea harengus</i> )	Clupeidae	ID: 15477		Ch_v2.0.2
<b>Cse</b>	Tongue sole ( <i>Cynoglossus semilaevis</i> )	Cynoglossidae	ID: 11788	ENSSCEG	Cse_v1.0
<b>Cva</b>	Sheepshead minnow ( <i>Cyprinodon variegatus</i> )	Cyprinodontidae	ID: 13078	ENSCVAG	C_variegatus-1.0
<b>Dcl</b>	Denticle herring ( <i>Denticeps clupeioides</i> )	Denticipitidae	ID: 7889		fDenClu1.1
<b>Dre</b>	Zebrafish ( <i>Danio rerio</i> )	Cyprinidae	ID: 50	ENSDARG	GRCz11
<b>Eel</b>	Electric eel ( <i>Electrophorus electricus</i> )	Gymnotidae	ID: 72955	ENSEELG	Ee_SOAP_WITH_SSPACE

## PARTIE I : LA REPRESSION POLYCOMB

<b>Elu</b>	Northern pike ( <i>Esox lucius</i> )	Esocidae	ID: 22932	ENSELUG	Eluc_v4 / Eluc_V3
<b>Ena</b>	Live sharksucker ( <i>Echeneis naucrates</i> )	Echeneidae	ID: 22910		fEcheNa1.1
<b>Fhe</b>	Mummichog ( <i>Fundulus heteroclitus</i> )	Fundulidae	ID: 743	ENSFHEG	Fundulus_heteroclitus-3.0.2
<b>Gac</b>	Stickleback ( <i>Gasterosteus aculeatus</i> )	Gasterosteidae		ENSGACG	BROAD S1
<b>Gaf</b>	Western mosquitofish ( <i>Gambusia affinis</i> )	Poeciliidae		ENSGAFG	ASM309773v1
<b>Gmo</b>	Cod ( <i>Gadus morhua</i> )	Gadidae		ENSGMOG	gadMor1

**Table A1.** *Cont.*

		<b>Family</b>	<b>NCBI Genome</b>	<b>Ensembl ID</b>	<b>Assembly</b>
<b>Gwi</b>	Blunt-snouted clingfish ( <i>Gouania willdenowi</i> )	Gobiesocidae	ID: 76090	ENSGWIG	fGouWil2.1
<b>Hbu</b>	Burton's mouthbrooder ( <i>Haplochromis burtoni</i> )	Cichlidae	ID: 3328	ENSHBUG	AstBur1.0
<b>Hco</b>	Tiger tail seahorse ( <i>Hippocampus comes</i> )	Syngnathidae	ID: 16754	ENSHCOG	H_comes_QL1_v1
<b>Hhu</b>	Huchen ( <i>Hucho hucho</i> )	Salmonidae		ENSHHUG	ASM331708v1
<b>Ipu</b>	Channel catfish ( <i>Ictalurus punctatus</i> )	Ictaluridae	ID: 198	ENSIPUG	IpCoco_1.2
<b>Kma</b>	Mangrove rivulus ( <i>Kryptolebias marmoratus</i> )	Rivulinae	ID: 23212	ENSKMAG	ASM164957v1
<b>Lbe</b>	Ballan wrasse ( <i>Labrus bergylta</i> )	Labridae	ID: 14767	ENSLBEG	BallGen_V1
<b>Lca</b>	Barramundi perch ( <i>Lates calcarifer</i> )	Centropomidae	ID: 14180	ENSLCAG	ASM164080v1 ASB_HGAPassembly_v1
<b>Lcr</b>	Large yellow croaker ( <i>Larimichthys crocea</i> )	Sciaenidae	ID: 12197		L_crocea_2.0
<b>Mal</b>	Swamp eel ( <i>Monopterus albus</i> )	Synbranchidae	ID: 24053	ENSMALG	M_albus_1.0
<b>Mar</b>	Zig-zag eel ( <i>Mastacembelus armatus</i> )	Mastacembelidae	ID: 31700	ENSMAMG	fMasArm1.1
<b>Mmo</b>	Ocean sunfish ( <i>Mola mola</i> )	Molidae		ENSMMOG	ASM169857v1
<b>Mze</b>	Zebra mbuna ( <i>Maylandia zebra</i> )	Cichlidae	ID: 2640	ENSMZEG	M_zebra_UMD2a

## PARTIE I : LA REPRESSION POLYCOMB

<b>Nbr</b>	Lyretail cichlid ( <i>Neolamprologus brichardi</i> )	Cichlidae	ID: 3329	ENSNBRG	NeoBri1.0
<b>Nco</b>	Black rockcod ( <i>Notothenia coriiceps</i> )	Nototheniidae	ID: 10753		NC01
<b>Nfu</b>	Turquoise killifish ( <i>Nothobranchius furzeri</i> )	Nothobranchiidae	ID: 2642		Nfu_20140520
<b>Oki</b>	Coho salmon ( <i>Oncorhynchus kisutch</i> )	Salmonidae	ID: 13127		Okis_V1
<b>Ola</b>	Medaka ( <i>Oryzias latipes</i> )	Adrianichthyidae	ID: 542	ENSORLG	ASM223467v1
<b>Omy</b>	Rainbow trout ( <i>Oncorhynchus mykiss</i> )	Salmonidae	ID: 196		Omyk_1.0
<b>Oni</b>	Tilapia ( <i>Oreochromis niloticus</i> )	Cichlidae	ID: 197	ENSONIG	O_niloticus_UMD_NMBU Orenil1.0
<b>Pfl</b>	Yellow perch ( <i>Perca flavescens</i> )	Percidae	ID: 15707		PFLA_1.0
<b>Pfo</b>	Amazon molly ( <i>Poecilia formosa</i> )	Poeciliidae	ID: 13072	ENSPFOG	Poecilia_formosa-5.1.2
<b>Pki</b>	Elephantfish ( <i>Paramormyrops kingsleyae</i> )	Mormyridae	ID: 66590	ENSPKIG	PKINGS_0.1
<b>Pla</b>	Sailfin molly ( <i>Poecilia latipinna</i> )	Poeciliidae	ID: 17477	ENSPLAG	P_latipinna-1.0
<b>Pma</b>	Mudskipper ( <i>Periophthalmus magnuspinnatus</i> )	Oxudercidae		ENSPMG	PM.fa

**Table A1.** *Cont.*

		<b>Family</b>	<b>NCBI Genome</b>	<b>Ensembl ID</b>	<b>Assembly</b>
<b>Pme</b>	Shortfin molly ( <i>Poecilia mexicana</i> )	Poeciliidae	ID: 14658	ENSPMEG	P_mexicana-1.0
<b>Pna</b>	Red-bellied piranha ( <i>Pygocentrus nattereri</i> )	Serrasalminidae	ID: 10452	ENSPNAG	Pygocentrus_nattereri-1.0.2
<b>Pny</b>	Makobe Island cichlid ( <i>Pundamilia nyererei</i> )	Cichlidae	ID: 3330	ENSPNYG	PunNye1.0
<b>Pra</b>	Indian glassy fish ( <i>Parambassis ranga</i> )	Ambassidae	ID: 76088		fParRan2.1
<b>Pre</b>	Guppy ( <i>Poecilia reticulata</i> )	Poeciliidae	ID: 23338	ENSPREG	Guppy_female_1.0_MT
<b>Sal</b>	Arctic char ( <i>Salvelinus alpinus</i> )	Salmonidae	ID: 12179		ASM291031v2

## PARTIE I : LA REPRESSION POLYCOMB

<b>San</b>	安水金线鲃 ( <i>Sinocyclocheilus anshuiensis</i> )	Cyprinidae	ID: 38362		SAMN03320099.WGS_v1.1
<b>Sdu</b>	Greater amberjack ( <i>Seriola dumerili</i> )	Carangidae	ID: 12614	ENSSDUG	Sdu_1.0
<b>Sfa</b>	Jewelled blenny ( <i>Salarias fasciatus</i> )	Blenniidae	ID: 7248	ENSSFAG	fSalaFa1.1
<b>Sfo</b>	Asian bonytongue ( <i>Scleropages formosus</i> )	Osteoglossidae	ID: 7635	ENSSFOG	fScFor1.1 / ASM162426v1
<b>Sgr</b>	Golden-line barbel ( <i>Sinocyclocheilus grahami</i> )	Cyprinidae	ID: 9142		SAMN03320097.WGS_v1.1
<b>Sla</b>	Yellowtail amberjack ( <i>Seriola lalandi dorsalis</i> )	Carangidae	ID: 12613	ENSSLDG	Sedor1
<b>Sma</b>	Turbot ( <i>Scophthalmus maximus</i> )	Scophthalmidae		ENSSMAG	ASM318616v1
<b>Spa</b>	Bicolor damselfish ( <i>Stegastes partitus</i> )	Pomacentridae	ID: 13077	ENSSPAG	Stegastes_partitus-1.0.2
<b>Srh</b>	犀角金线鲃 ( <i>Sinocyclocheilus rhinocerosus</i> )	Cyprinidae	ID: 38394		SAMN03320098_v1.1
<b>Ssa</b>	Atlantic salmon ( <i>Salmo salar</i> )	Salmonidae	ID: 369		ICSASG_v2
<b>Tni</b>	Tetraodon ( <i>Tetraodon nigroviridis</i> )	Tetraodontidae		ENSTNIG	TETRAODON 8.0
<b>Tru</b>	Japanese puffer ( <i>Takifugu rubripes</i> )	Tetraodontidae	ID: 63	ENSTRUG	fTakRub1.2 / FUGU5
<b>Xco</b>	Monterrey platyfish ( <i>Xiphophorus couchianus</i> )	Poeciliidae	ID: 16825	ENSXCOG	X_couchianus-1.0 Xiphophorus_couchianus-4.0.1
<b>Xma</b>	Platyfish ( <i>Xiphophorus maculatus</i> )	Poeciliidae	ID: 10764	ENSXMAG	X_maculatus-5.0-male

## References

1. Nelson, J.S. *Fishes of the World*, 4th ed.; John Wiley & Sons: Hoboken, NJ, USA, 2006.
2. Faircloth, B.C.; Sorenson, L.; Santini, F.; Alfaro, M.E. A Phylogenomic Perspective on the Radiation of Ray-Finned Fishes Based upon Targeted Sequencing of Ultraconserved Elements (UCEs). *PLoS ONE* **2013**, *8*, e65923. [[CrossRef](#)]
3. Sallan, L. Major issues in the origins of ray-finned fish (Actinopterygii) biodiversity. *Biol. Rev. Camb. Philos. Soc.* **2014**, *89*, 950–971. [[CrossRef](#)] [[PubMed](#)]
4. Amores, A.; Force, A.; Yan, Y.-L.; Joly, L.; Amemiya, C.; Fritz, A.; Ho, R.K.; Langeland, J.; Prince, V.; Wang, Y.-L.; *et al.* Zebrafish hox Clusters and Vertebrate Genome Evolution. *Science* **1998**, *282*, 1711–1714. [[CrossRef](#)] [[PubMed](#)]
5. Taylor, J.S.; Van De Peer, Y.; Braasch, I.; Meyer, A. Comparative genomics provides evidence for an ancient genome duplication event in fish. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **2001**, *356*, 1661–1679. [[CrossRef](#)] [[PubMed](#)]
6. Hoegg, S.; Brinkmann, H.; Taylor, J.S.; Meyer, A. Phylogenetic Timing of the Fish-Specific Genome Duplication Correlates with the Diversification of Teleost Fish. *J. Mol. Evol.* **2004**, *59*, 190–203. [[CrossRef](#)] [[PubMed](#)]
7. Sémon, M.; Wolfe, K.H. Reciprocal gene loss between Tetraodon and zebrafish after whole genome duplication in their ancestor. *Trends Genet.* **2007**, *23*, 108–112. [[CrossRef](#)] [[PubMed](#)]
8. He, X.; Zhang, J. Rapid Subfunctionalization Accompanied by Prolonged and Substantial Neofunctionalization in Duplicate Gene Evolution. *Genetics* **2005**, *169*, 1157–1164. [[CrossRef](#)]
9. Brunet, F.; Crollius, H.R.; Paris, M.; Aury, J.-M.; Gibert, P.; Jaillon, O.; Laudet, V.; Robinson-Rechavi, M. Gene Loss and Evolutionary Rates Following Whole-Genome Duplication in Teleost Fishes. *Mol. Biol. Evol.* **2006**, *23*, 1808–1816. [[CrossRef](#)]
10. Sémon, M.; Wolfe, K.H. Rearrangement Rate following the Whole-Genome Duplication in Teleosts. *Mol. Biol. Evol.* **2006**, *24*, 860–867.
11. Hufton, A.; Groth, D.; Vingron, M.; Lehrach, H.; Poustka, A.J.; Panopoulou, G. Early vertebrate whole genome duplications were predated by a period of intense genome rearrangement. *Genome Res.* **2008**, *18*, 1582–1591. [[CrossRef](#)]
12. Robinson-Rechavi, M.; Laudet, V. Evolutionary rates of duplicate genes in fish and mammals. *Mol. Biol. Evol.* **2001**, *18*, 681–683. [[CrossRef](#)] [[PubMed](#)]
13. Di Croce, L.; Helin, K. Transcriptional regulation by Polycomb group proteins. *Nat. Struct. Mol. Biol.* **2013**, *20*, 1147–1155. [[CrossRef](#)] [[PubMed](#)]
14. Margueron, R.; Reinberg, D. The Polycomb complex PRC2 and its mark in life. *Nature* **2011**, *469*, 343–349. [[CrossRef](#)] [[PubMed](#)]
15. Sauvageau, M.; Sauvageau, G. Polycomb group proteins: Multi-faceted regulators of somatic stem cells and cancer. *Cell Stem Cell* **2010**, *7*, 299–313. [[CrossRef](#)]
16. Surface, L.E.; Thornton, S.R.; Boyer, L.A. Polycomb Group Proteins Set the Stage for Early Lineage Commitment. *Cell Stem Cell* **2010**, *7*, 288–298. [[CrossRef](#)]
17. Shao, Z.; Raible, F.; Mollaaghababa, R.; Guyon, J.R.; Wu, C.-T.; Bender, W.; Kingston, R.E. Stabilization of Chromatin Structure by PRC1, a Polycomb Complex. *Cell* **1999**, *98*, 37–46. [[CrossRef](#)]
18. Kuzmichev, A.; Nishioka, K.; Erdjument-Bromage, H.; Tempst, P.; Reinberg, D. Histone methyltransferase activity associated with a human multiprotein complex containing the Enhancer of Zeste protein. *Genome Res.* **2002**, *16*, 2893–2905. [[CrossRef](#)]
19. Cao, R.; Wang, L.; Wang, H.; Xia, L.; Erdjument-Bromage, H.; Tempst, P.; Jones, R.S.; Zhang, Y. Role of Histone H3 Lysine 27 Methylation in Polycomb-Group Silencing. *Science* **2002**, *298*, 1039–1043. [[CrossRef](#)]
20. Ringrose, L.; Paro, R. Epigenetic Regulation of Cellular Memory by the Polycomb and Trithorax Group Proteins. *Annu. Rev. Genet.* **2004**, *38*, 413–443. [[CrossRef](#)]
21. Völkel, P.; Angrand, P.-O. The control of histone lysine methylation in epigenetic regulation. *Biochimie* **2007**, *89*, 1–20. [[CrossRef](#)]
22. Müller, J.; Hart, C.M.; Francis, N.J.; Vargas, M.L.; Sengupta, A.; Wild, B.; Miller, E.L.; O'Connor, M.; Kingston, R.E.; A Simon, J. Histone Methyltransferase Activity of a Drosophila Polycomb Group Repressor Complex. *Cell* **2002**, *111*, 197–208. [[CrossRef](#)]



23. Nekrasov, M.; Klymenko, T.; Fraterman, S.; Papp, B.; Oktaba, K.; Köcher, T.; Cohen, A.; Stunnenberg, H.G.; Wilm, M.; Müller, J. Pcl-PRC2 is needed to generate high levels of H3-K27 trimethylation at Polycomb target genes. *EMBO J.* **2007**, *26*, 4078–4088. [[CrossRef](#)] [[PubMed](#)]
24. Fischle, W.; Wang, Y.; Jacobs, S.A.; Kim, Y.; Allis, C.D.; Khorasanizadeh, S. Molecular basis for the discrimination of repressive methyl-lysine marks in histone H3 by Polycomb and HP1 chromodomains. *Genes Dev.* **2003**, *17*, 1870–1881. [[CrossRef](#)] [[PubMed](#)]
25. Min, J.; Zhang, Y.; Xu, R.-M. Structural basis for specific binding of Polycomb chromodomain to histone H3 methylated at Lys 27. *Genome Res.* **2003**, *17*, 1823–1828. [[CrossRef](#)] [[PubMed](#)]
26. Levine, S.S.; Weiss, A.; Erdjument-Bromage, H.; Shao, Z.; Tempst, P.; Kingston, R.E. The Core of the Polycomb Repressive Complex Is Compositionally and Functionally Conserved in Flies and Humans. *Mol. Cell. Biol.* **2002**, *22*, 6070–6078. [[CrossRef](#)]
27. Connelly, K.E.; Dykhuizen, E.C. Compositional and functional diversity of canonical PRC1 complexes in mammals. *Biochim. Biophys. Acta (BBA) Gene Regul. Mech.* **2017**, *1860*, 233–245. [[CrossRef](#)]
28. Maertens, G.; El Messaoudi-Aubert, S.; Račák, T.; Stock, J.K.; Nicholls, J.; Rodríguez-Niedenfuhr, M.; Gil, J.; Peters, G. Several Distinct Polycomb Complexes Regulate and Co-Localize on the INK4a Tumor Suppressor Locus. *PLOS ONE* **2009**, *4*, e6380. [[CrossRef](#)]
29. Vandamme, J.; Völkel, P.; Rosnoblet, C.; Le Faou, P.; Angrand, P.-O. Interaction proteomics analysis of polycomb proteins defines distinct PRC1 complexes in mammalian cells. *Mol. Cell. Proteomics* **2011**, *10*. [[CrossRef](#)]
30. Ren, X.; Kerppola, T. REST Interacts with Cbx Proteins and Regulates Polycomb Repressive Complex 1 Occupancy at RE1 Elements. *Mol. Cell. Biol.* **2011**, *31*, 2100–2110. [[CrossRef](#)]
31. Trojer, P.; Cao, A.R.; Gao, Z.; Li, Y.; Zhang, J.; Xu, X.; Li, G.; Losson, R.; Erdjument-Bromage, H.; Tempst, P.; *et al.* L3MBTL2 Protein Acts in Concert with PcG Protein-Mediated Monoubiquitination of H2A to Establish a Repressive Chromatin Structure. *Mol. Cell* **2011**, *42*, 438–450. [[CrossRef](#)]
32. Gao, Z.; Zhang, J.; Bonasio, R.; Strino, F.; Sawai, A.; Parisi, F.; Kluger, Y.; Reinberg, D. PCGF Homologs, CBX Proteins, and RYBP Define Functionally Distinct PRC1 Family Complexes. *Mol. Cell* **2012**, *45*, 344–356. [[CrossRef](#)]
33. O’Loughlen, A.; Muñoz-Cabello, A.M.; Gaspar-Maia, A.; Wu, H.-A.; Banito, A.; Kunowska, N.; Racek, T.; Pemberton, H.N.; Beolchi, P.; Laval, F.; *et al.* MicroRNA Regulation of Cbx7 Mediates a Switch of Polycomb Orthologs during ESC Differentiation. *Cell Stem Cell* **2012**, *10*, 33–46. [[CrossRef](#)]
34. Bernstein, E.; Duncan, E.M.; Masui, O.; Gil, J.; Heard, E.; Allis, C.D. Mouse Polycomb Proteins Bind Differentially to Methylated Histone H3 and RNA and Are Enriched in Facultative Heterochromatin. *Mol. Cell. Biol.* **2006**, *26*, 2560–2569. [[CrossRef](#)] [[PubMed](#)]
35. Morey, L.; Pascual, G.; Cozzuto, L.; Roma, G.; Wutz, A.; Benitah, S.A.; Di Croce, L. Nonoverlapping Functions of the Polycomb Group Cbx Family of Proteins in Embryonic Stem Cells. *Cell Stem Cell* **2012**, *10*, 47–62. [[CrossRef](#)] [[PubMed](#)]
36. Klauke, K.; Radulovic, V.; Broekhuis, M.; Weersing, E.; Zwart, E.; Olthof, S.; Ritsema, M.; Bruggeman, S.; Wu, X.; Helin, K.; *et al.* Polycomb Cbx family members mediate the balance between haematopoietic stem cell self-renewal and differentiation. *Nature* **2013**, *15*, 353–362. [[CrossRef](#)] [[PubMed](#)]
37. Le Faou, P.; Völkel, P.; Angrand, P.-O. The zebrafish genes encoding the Polycomb repressive complex (PRC) 1. *Gene* **2011**, *475*, 10–21. [[CrossRef](#)]
38. Sievers, F.; Wilm, A.; Dineen, D.; Gibson, T.J.; Karplus, K.; Li, W.; López, R.; McWilliam, H.; Remmert, M.; Soeding, J.; *et al.* Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol. Syst. Biol.* **2011**, *7*, 539. [[CrossRef](#)]
39. Thompson, J.D.; Higgins, D.; Gibson, T.J. CLUSTAL W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* **1994**, *22*, 4673–4680. [[CrossRef](#)]
40. Edgar, R.C. MUSCLE: Multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* **2004**, *32*, 1792–1797. [[CrossRef](#)]
41. Kumar, S.; Stecher, G.; Li, M.; Nnyaz, C.; Tamura, K. MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms. *Mol. Biol. Evol.* **2018**, *35*, 1547–1549. [[CrossRef](#)]
42. Jones, D.T.; Taylor, W.R.; Thornton, J.M. The rapid generation of mutation data matrices from protein sequences. *Comput. Appl. Biosci.* **1992**, *8*, 275–282. [[CrossRef](#)] [[PubMed](#)]

43. Saitou, N.; Nei, M. The neighbor-joining method: A new method for reconstructing phylogenetic trees. *Mol. Boil. Evol.* **1987**, *4*, 406–425.
44. Nei, M.; Kumar, S. *Molecular Evolution and Phylogenetics*; Oxford University Press: New York, USA, 2000.
45. Letunic, I.; Bork, P. 20 years of the SMART protein domain annotation resource. *Nucleic Acids Res.* **2018**, *46*, D493–D496. [[CrossRef](#)] [[PubMed](#)]
46. Crooks, G.E.; Hon, G.; Chandonia, J.-M.; Brenner, S.E. WebLogo: A Sequence Logo Generator. *Genome Res.* **2004**, *14*, 1188–1190. [[CrossRef](#)]
47. Kimmel, C.B.; Ballard, W.W.; Kimmel, S.R.; Ullmann, B.; Schilling, T.F. Stages of embryonic development of the zebrafish. *Dev. Dyn.* **1995**, *203*, 253–310. [[CrossRef](#)]
48. Thisse, C.; Thisse, B. High-resolution in situ hybridization to whole-mount zebrafish embryos. *Nat. Protoc.* **2007**, *3*, 59–69. [[CrossRef](#)]
49. Dupret, B.; Völkel, P.; Follet, P.; Le Bourhis, X.; Angrand, P.-O. Combining genotypic and phenotypic analyses on single mutant zebrafish larvae. *MethodsX* **2018**, *5*, 244–256. [[CrossRef](#)]
50. Wang, R.; Taylor, A.B.; Leal, B.Z.; Chadwell, L.V.; Ilangoan, U.; Robinson, A.K.; Schirf, V.; Hart, P.J.; Lafer, E.; Demeler, B.; *et al.* Polycomb Group Targeting through Different Binding Partners of RING1B C-Terminal Domain. *Structure* **2010**, *18*, 966–975. [[CrossRef](#)]
51. Senthilkumar, R.; Mishra, R.K. Novel motifs distinguish multiple homologues of Polycomb in vertebrates: Expansion and diversification of the epigenetic toolkit. *BMC Genomics* **2009**, *10*, 549. [[CrossRef](#)]
52. Vincenz, C.; Kerppola, T. Different polycomb group CBX family proteins associate with distinct regions of chromatin using nonhomologous protein sequences. *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 16572–16577. [[CrossRef](#)]
53. Near, T.J.; Eytan, R.I.; Dornburg, A.; Kuhn, K.L.; Moore, J.; Davis, M.P.; Wainwright, P.C.; Friedman, M.; Smith, W.L. Resolution of ray-finned fish phylogeny and timing of diversification. *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 13698–13703. [[CrossRef](#)] [[PubMed](#)]
54. Betancur-R, R.; Broughton, R.E.; Wiley, E.O.; Carpenter, K.; López, J.A.; Li, C.; Holcroft, N.I.; Arcila, D.; Sanciangco, M.; Cureton, J.C., II; *et al.* The Tree of Life and a New Classification of Bony Fishes. *PLoS Curr.* **2013**, *5*, 5.
55. Betancur-R, R.; Wiley, E.O.; Arratia, G.; Acero, A.; Bailly, N.; Miya, M.; Lecointre, G.; Ortí, G. Phylogenetic classification of bony fishes. *BMC Evol. Biol.* **2017**, *17*, 162. [[CrossRef](#)] [[PubMed](#)]
56. Hughes, L.C.; Ortí, G.; Huang, Y.; Sun, Y.; Baldwin, C.C.; Thompson, A.W.; Arcila, D.; Betancur-R, R.; Li, C.; Becker, L.; *et al.* Comprehensive phylogeny of ray-finned fishes (Actinopterygii) based on transcriptomic and genomic data. *Proc. Natl. Acad. Sci. USA* **2018**, *115*, 6249–6254. [[CrossRef](#)]
57. Woods, I.G.; Kelly, P.D.; Chu, F.; Ngo-Hazelett, P.; Yan, Y.-L.; Huang, H.; Postlethwait, J.H.; Talbot, W.S. A Comparative Map of the Zebrafish Genome. *Genome Res.* **2000**, *10*, 1903–1914. [[CrossRef](#)]
58. Jaillon, O.; Aury, J.-M.; Brunet, F.; Petit, J.-L.; Stange-Thomann, N.; Mauceli, E.; Bouneau, L.; Fischer, C.; Ozouf-Costaz, C.; Bernot, A.; *et al.* Genome duplication in the teleost fish Tetraodon nigroviridis reveals the early vertebrate proto-karyotype. *Nature* **2004**, *431*, 946–957. [[CrossRef](#)]
59. Gallant, J.R.; Losilla, M.; Tomlinson, C.; Warren, W.C. The Genome and Adult Somatic Transcriptome of the Mormyrid Electric Fish *Paramormyrops kingsleyae*. *Genome Biol. Evol.* **2017**, *9*, 3525–3530. [[CrossRef](#)]
60. Bian, C.; Hu, Y.; Ravi, V.; Kuznetsova, I.S.; Shen, X.; Mu, X.; Sun, Y.; You, X.; Li, J.; Li, X.; *et al.* The Asian arowana (*Scleropages formosus*) genome provides new insights into the evolution of an early lineage of teleosts. *Sci. Rep.* **2016**, *6*, 24501. [[CrossRef](#)]
61. Stout, C.C.; Tan, M.; Lemmon, A.R.; Lemmon, E.M.; Armbruster, J.W. Resolving Cypriniformes relationships using an anchored enrichment approach. *BMC Evol. Biol.* **2016**, *16*, 244. [[CrossRef](#)]
62. Xu, P.; Zhang, X.; Wang, X.; Li, J.; Liu, G.; Kuang, Y.-Y.; Xu, J.; Zheng, X.; Ren, L.; Wang, G.; *et al.* Genome sequence and genetic diversity of the common carp, *Cyprinus carpio*. *Nat. Genet.* **2014**, *46*, 1212–1219. [[CrossRef](#)]
63. Chen, Z.; Omori, Y.; Koren, S.; Shirokiya, T.; Kuroda, T.; Miyamoto, A.; Wada, H.; Fujiyama, A.; Toyoda, A.; Zhang, S.; *et al.* De novo assembly of the goldfish (*Carassius auratus*) genome and the evolution of genes after whole-genome duplication. *Sci. Adv.* **2019**, *5*, eaav0547. [[CrossRef](#)] [[PubMed](#)]
64. Yang, J.; Chen, X.; Bai, J.; Fang, D.; Qiu, Y.; Jiang, W.; Yuan, H.; Bian, C.; Lu, J.; He, S.; *et al.* The Sinocyclocheilus cavefish genome provides insights into cave adaptation. *BMC Biol.* **2016**, *14*, 1. [[CrossRef](#)] [[PubMed](#)]

65. Rondeau, E.B.; Minkley, D.R.; Leong, J.S.; Messmer, A.M.; Jantzen, J.R.; Von Schalburg, K.R.; Lemon, C.; Bird, N.H.; Koop, B.F. The Genome and Linkage Map of the Northern Pike (*Esox lucius*): Conserved Synteny Revealed between the Salmonid Sister Group and the Neoteleostei. *PLoS ONE* **2014**, *9*, e102089. [[CrossRef](#)] [[PubMed](#)]
66. Berthelot, C.; Brunet, F.; Chalopin, M.; Juanchich, A.; Bernard, M.; Noel, B.; Bento, P.; Da Silva, C.; Labadie, K.; Alberti, A.; *et al.* The rainbow trout genome provides novel insights into evolution after whole-genome duplication in vertebrates. *Nat. Commun.* **2014**, *5*, 3657. [[CrossRef](#)]
67. Lien, S.; Koop, B.F.; Sandve, S.R.; Miller, J.; Kent, M.P.; Nome, T.; Hvidsten, T.R.; Leong, J.S.; Minkley, D.R.; Zimin, A.; *et al.* The Atlantic salmon genome provides insights into rediploidization. *Nature* **2016**, *533*, 200–205. [[CrossRef](#)]
68. Aranda, S.; Mas, G.; Di Croce, L. Regulation of gene transcription by Polycomb proteins. *Sci. Adv.* **2015**, *1*, e1500737. [[CrossRef](#)]
69. Vidal, M.; Starowicz, K.; Starowicz, K. Polycomb complexes PRC1 and their function in hematopoiesis. *Exp. Hematol.* **2017**, *48*, 12–31. [[CrossRef](#)]
70. Hauri, S.; Comoglio, F.; Seimiya, M.; Gerstung, M.; Glatter, T.; Hansen, K.; Aebersold, R.; Paro, R.; Gstaiger, M.; Beisel, C. A High-Density Map for Navigating the Human Polycomb Complexome. *Cell Rep.* **2016**, *17*, 583–595. [[CrossRef](#)]
71. Ogawa, H.; Ishiguro, K.-I.; Gaubatz, S.; Livingston, D.M.; Nakatani, Y. A Complex with Chromatin Modifiers That Occupies E2F- and Myc-Responsive Genes in G0 Cells. *Science* **2002**, *296*, 1132–1136. [[CrossRef](#)]
72. Attwooll, C.; Oddi, S.; Cartwright, P.; Prosperini, E.; Agger, K.; Steensgaard, P.; Wagener, C.; Sardet, C.; Moroni, M.C.; Helin, K. A Novel Repressive E2F6 Complex Containing the Polycomb Group Protein, EPC1, That Interacts with EZH2 in a Proliferation-specific Manner. *J. Biol. Chem.* **2004**, *280*, 1199–1208. [[CrossRef](#)]
73. Pelegri, F. Maternal factors in zebrafish development. *Dev. Dyn.* **2003**, *228*, 535–554. [[CrossRef](#)] [[PubMed](#)]
74. Dosch, R.; Wagner, D.S.; A Mintzer, K.; Runke, G.; Wiemelt, A.P.; Mullins, M.C. Maternal Control of Vertebrate Development before the Midblastula Transition. *Dev. Cell* **2004**, *6*, 771–780. [[CrossRef](#)] [[PubMed](#)]
75. White, R.J.; E Collins, J.; Sealy, I.M.; Wali, N.; Dooley, C.M.; Digby, Z.; Stemple, D.L.; Murphy, D.N.; Billis, K.; Hourlier, T.; *et al.* A high-resolution mRNA expression time course of embryonic development in zebrafish. *eLife* **2017**, *6*. [[CrossRef](#)] [[PubMed](#)]
76. Smith, J.J.; Kuraku, S.; Holt, C.; Sauka-Spengler, T.; Jiang, N.; Campbell, M.S.; Yandell, M.D.; Manousaki, T.; Meyer, A.; Bloom, O.; *et al.* Sequencing of the sea lamprey (*Petromyzon marinus*) genome provides insights into vertebrate evolution. *Nat. Genet.* **2013**, *45*, 415–421. [[CrossRef](#)] [[PubMed](#)]
77. Nakatani, Y.; Takeda, H.; Kohara, Y.; Morishita, S. Reconstruction of the vertebrate ancestral genome reveals dynamic genome reorganization in early vertebrates. *Genome Res.* **2007**, *17*, 1254–1265. [[CrossRef](#)] [[PubMed](#)]
78. Dupret, B.; Völkel, P.; Le Bourhis, X.; Angrand, P.-O. The Polycomb Group Protein Pcgf1 Is Dispensable in Zebrafish but Involved in Early Growth and Aging. *PLoS ONE* **2016**, *11*, e0158700. [[CrossRef](#)] [[PubMed](#)]
79. Bajusz, I.; Henry, S.; Sutus, E.; Kovács, G.; Purity, M. Evolving Role of RING1 and YY1 Binding Protein in the Regulation of Germ-Cell-Specific Transcription. *Genes* **2019**, *10*, 941. [[CrossRef](#)]
80. Fursova, N.A.; Blackledge, N.P.; Nakayama, M.; Ito, S.; Koseki, Y.; Farcas, A.M.; King, H.W.; Koseki, H.; Klose, R.J. Synergy between Variant PRC1 Complexes Defines Polycomb-Mediated Gene Repression. *Mol. Cell* **2019**, *74*, 1020–1036.e8. [[CrossRef](#)]
81. Van Der Velden, Y.U.; Wang, L.; Van Lohuizen, M.; Haramis, A.-P.G. The Polycomb group protein Ring1b is essential for pectoral fin development. *Development* **2012**, *139*, 2210–2220. [[CrossRef](#)]
82. San, B.; Chrispijn, N.; Wittkopp, N.; Van Heeringen, S.J.; Lagendijk, A.K.; Aben, M.; Bakkers, J.; Ketting, R.; Kamminga, L.M. Normal formation of a vertebrate body plan and loss of tissue maintenance in the absence of ezh2. *Sci. Rep.* **2016**, *6*, 24658. [[CrossRef](#)]
83. Dupret, B.; Völkel, P.; Vennin, C.; Toillon, R.-A.; Le Bourhis, X.; Angrand, P.-O. The histone lysine methyltransferase Ezh2 is required for maintenance of the intestine integrity and for caudal fin regeneration in zebrafish. *Biochim. Biophys. Acta (BBA) Gene Regul. Mech.* **2017**, *1860*, 1079–1093. [[CrossRef](#)] [[PubMed](#)]
84. Völkel, P.; Bary, A.; Raby, L.; Chapart, A.; Dupret, B.; Le Bourhis, X.; Angrand, P.-O. Ezh1 arises from Ezh2 gene duplication but its function is not required for zebrafish development. *Sci. Rep.* **2019**, *9*, 4319. [[CrossRef](#)] [[PubMed](#)]
85. Chrispijn, N.; Andralojc, K.M.; Castenmiller, C.; Kamminga, L.M. Gene expression profile of a selection of Polycomb Group genes during zebrafish embryonic and germ line development. *PLoS ONE* **2018**, *13*, e0200316. [[CrossRef](#)] [[PubMed](#)]

### 1.3 Conclusion

L'analyse globale des gènes *cbx* confirme les résultats préliminaires sur l'évolution des vertébrés et des téléostéens. Deux événements de duplication ont eu lieu proche de l'origine des vertébrés et une duplication supplémentaire est intervenue à l'origine des téléostéens (Holland *et al.*, 1994 ; Nakatani *et al.*, 2007). Les mammifères possèdent 5 orthologues du gène *pc* de la drosophile : *CBX2*, *CBX4*, *CBX6*, *CBX7* et *CBX8*. Ces orthologues sont tous conservés chez les téléostéens, au contraire de certains gènes de la famille Polycomb comme *RING1* ou *PCGF3*. Après les grands événements de duplication, la majorité des gènes dupliqués sont perdus. La conservation de chacune des protéines Cbx suggèrent que chacune d'entre elles possèdent une fonction propre. La différence majeure entre les mammifères et les téléostéens est la conservation d'orthologues. Les gènes de la famille *cbx* ne sont néanmoins pas conservés en pair d'orthologues dans toutes les familles de téléostéens et semblent spécifiques des clades. Le poisson zèbre possède un répertoire unique : les gènes *cbx2* et *cbx4* présents en copie unique tandis que les gènes *cbx6*, *cbx7* et *cbx8* sont présents en pair d'orthologues. Le profil d'expression de ces gènes a été observé par hybridation *in situ*. Tous les gènes de la famille *cbx* sont exprimés chez le poisson zèbre au cours du développement et au stade adulte. Les transcrits d'origine maternelle sont déposés dans les œufs et indiquent un rôle dans les premières étapes du développement. Au cours du développement les profils d'expressions sont principalement retrouvés dans les régions antérieures de l'embryon, au niveau de zones de prolifération cellulaire actives. Si les profils sont généralement similaires, quelques nuances suggèrent là encore des spécificités de gènes *cbx*.

## 2. Ezh1 arises from Ezh2 gene duplication but its function is not required for zebrafish development

### 2.1 Contexte scientifique

L'activité enzymatique du complexe PRC2 est portée par deux sous-unités catalytiques EZH1 et EZH2. Chacune de ces protéines possède un domaine SET avec une activité histone méthyltransférase et permet le dépôt de la marque épigénétique H3K27me3. Néanmoins, l'efficacité de EZH1 et EZH2 est différente. La perte de fonction de Ezh2 chez le modèle murin inhibe le déroulement de la gastrulation (O'Carroll *et al.*, 2001). Le KO de Ezh1 au contraire provoque l'apparition de phénotypes plus restreints (Ezhkova *et al.*, 2011). Une certaine redondance fonctionnelle a été observée et une faible proportion de gènes sont réprimés par les complexes PRC2 portant EZH1 ou EZH2 (Margueron *et al.*, 2008). Toutefois, la perte de fonction de Ezh2 dans les ESCs réduit drastiquement la méthylation de la lysine 27 de l'histone H3 sans l'inhiber totalement, suggérant une complémentarité fonctionnelle (Shen *et al.*, 2008). De la même façon, l'action des deux protéines semblent dépendante du stade de développement et du statut de différenciation des cellules. La protéine EZH2 fortement exprimée dans les cellules souches et progéniteurs et EZH1 est observée dans les tissus adultes (Margueron *et al.*, 2008). Dans une étude préalable, le laboratoire a généré par la technologie des TALENs un mutant perte de fonction de *ezh2* chez le poisson zèbre. Bien que létale autour de 12 jpf, les mutants *ezh2<sup>ul2/ul2</sup>* le développement précoce se déroule correctement. Le transcrit *ezh2* est maternellement déposés dans les œufs, l'équipe de San a donc généré un mutant zygotique pour étudier le rôle de *ezh2* dans les phases les plus précoces du développement. De façon surprenante, et au contraire du modèle murin, les embryons se développent jusqu'à 2 jpf (San *et al.*, 2016). Ces résultats suggèrent que la protéine Ezh1 pourrait avoir un rôle non négligeable dans le développement précoce du poisson zèbre. Dans cette étude, des analyses comparatives de Ezh1 et Ezh2 *in silico* révèlent l'origine évolutive de Ezh1. La génération d'un nouveau mutant *ezh1<sup>ul3/ul3</sup>* montre que Ezh1 n'est pas nécessaire pour le développement précoce mais compense une partie de la perte d'activité de Ezh2.

## 2.2 Article

**Ezh1 arises from Ezh2 gene duplication but its function is not required for zebrafish development**Pamela Völkel<sup>1,2,3</sup>, Aurélie Bary<sup>1,2,4</sup>, Ludivine Raby<sup>1,2</sup>, Anaïs Chapart<sup>1,2</sup>, Barbara Dupret<sup>1,2</sup>, Xuefen Le Bourhis<sup>1,2</sup> & Pierre-Olivier Angrand<sup>1,2</sup>

Received: 21 August 2018, Accepted: 18 February 201

**Abstract**

Trimethylation on H3K27 mediated by Polycomb Repressive Complex 2 (PRC2) is required to control gene repression programs involved in development, regulation of tissue homeostasis or maintenance and lineage specification of stem cells. In *Drosophila*, the PRC2 catalytic subunit is the single protein E(z), while in mammals this function is fulfilled by two proteins, Ezh1 and Ezh2. Based on database searches, we propose that *Ezh1* arose from an *Ezh2* gene duplication that has occurred in the common ancestor to elasmobranchs and bony vertebrates. Expression studies in zebrafish using *in situ* hybridization and RT-PCR followed by the sequencing of the amplicon revealed that *ezh1* mRNAs are maternally deposited. Then, *ezh1* transcripts are ubiquitously distributed in the entire embryo at 24 hpf and become more restricted to anterior part of the embryo at later developmental stages. To unveil the function of *ezh1* in zebrafish, a mutant line was generated using the TALEN technology. *Ezh1*-deficient mutant fish are viable and fertile, but the loss of *ezh1* function is responsible for the earlier death of *ezh2* mutant larvae indicating that *ezh1* contributes to zebrafish development in absence of zygotic *ezh2* gene function. Furthermore, we show that presence of *ezh1* transcripts from the maternal origin accounts for the delayed lethality of *ezh2*-deficient larvae.

**Introduction**

Polycomb group (PcG) proteins are epigenetic regulators conserved from fruit flies to humans. They are involved in various biological processes including regulation of tissue homeostasis, maintenance and lineage specification of stem cells, and promote cancer progression when skewed<sup>1-4</sup>. PcG proteins assemble in two main histone-modifying protein complexes named Polycomb Repressive Complexes 1 and 2 (PRC1 and PRC2). PRC2 catalyzes the methylation of Lysine 27 of histone H3, generating the H3K27me2/3 epigenetic mark which in turn acts as a platform to recruit the PRC1 complex that ubiquitylates histone H2A at Lysine 119 (H2AK119ub1)<sup>5-10</sup>. These post-translational modifications are then responsible for local chromatin compaction and gene silencing. The PRC2 protein complex is composed of several subunits but its core comprises Enhancer of zeste homolog (Ezh), Embryonic ectoderm development (Eed) and Suppressor of zeste 12 (Suz12)<sup>2</sup>. *Drosophila* have a single Ezh gene [E(z)], whereas mammalian genomes encode two orthologs<sup>11</sup> defining two alternate PRC2 complexes, PRC2-Ezh1 and PRC2-Ezh2. Ezh1 and Ezh2 are SET domain-containing proteins harboring the histone methyltransferase activity, while Eed and Suz12 are involved in PRC2 stability and are required for Ezh1/2 catalytic activities<sup>12-16</sup>. Although Ezh2 knockout in embryonic stem cells (ESCs) strongly reduces H3K27me2/3, H3K27 methylation is not fully abolished suggesting that PRC2-Ezh1 complexes contribute to this epigenetic mark formation<sup>17</sup>. In contrast to PRC2-Ezh2, PRC2-Ezh1 exhibits a low histone methyltransferase activity and knockdown of Ezh1 does not result in global reduction of H3K27me2/3 levels<sup>13</sup>. However, PRC2-Ezh1 represses transcription *in vivo* and is able to compact chromatin *in vitro*<sup>13</sup>. Also, depletion of Ezh1 in cells lacking Ezh2 abolishes residual methylation on H3K27<sup>17</sup>. In addition, *Ezh1* is ubiquitously expressed, whereas *Ezh2* expression is associated with proliferating tissues<sup>13</sup>. Finally, *EZH2* has been largely documented to be involved in tumorigenesis<sup>18</sup>, which it is also the case for *EZH1*. In particular, mutations in *EZH1* on Glutamine 571 (Q571R) were found to occur in more than 25% of adult autonomous thyroid adenomas<sup>19</sup>. This *EZH1*<sup>Q571R</sup> mutation lies within the SET domain of *EZH1* and is responsible

---



for an increase of H3K27me3 methylation. The zebrafish (*Danio rerio*) provides a unique tool to investigate gene function during development and provides important models for human diseases. Owing to external fertilization and optical transparency of the embryos, zebrafish early development can be easily monitored. Furthermore, the recent emergence of powerful genome-editing technologies, such as the Transcription Activator-Like Effector Nuclease (TALEN) and Clustered Regularly Interspaced Short Palindromic Repeats/CRISPR-associated System (CRISPR/Cas9) applied to zebrafish allows rapid gene function studies in this organism<sup>20-25</sup>. We previously used the TALEN-mediated gene editing methodology to generate an *ezh2*<sup>ul2</sup> null allele in zebrafish<sup>26</sup>. Mouse embryos lacking *Ezh2* function fail to complete gastrulation<sup>27</sup>, but *ezh2*<sup>ul2/ul2</sup> zebrafish mutants present a normal body plan and die at around 12 days post fertilization (dpf) with intestinal defects. In contrast, maternal-zygotic (MZ) *ezh2* zebrafish mutants generated through germ cell transplantation (*MZezh2*<sup>hu5670/hu5670</sup><sup>28</sup>) also gastrulate and form a normal body organization, but die at around 2 dpf. The difference in the lethality timing between zygotic and maternal-zygotic *ezh2* mutants outlines the key role played by *ezh2* maternal products in zebrafish development. However, the fact that *MZezh2*<sup>hu5670/hu5670</sup> mutants develop a normal body plan raises the possibility that *ezh1* could contribute to early embryogenesis in absence of *ezh2* function.

Using searches in vertebrate genomic and transcriptomic databases, we first show here that *ezh1* arises from the duplication of the *ezh2* gene in the common ancestor to elasmobranchs and bony vertebrates. On the other hand, we conducted a study of *ezh1* expression and function during zebrafish development. Using the TALEN technology, we generated an *ezh1* loss-of-function zebrafish line. This line is viable and fertile indicating that *Ezh1* is dispensable in zebrafish. However, we show that *ezh1* contributes to zebrafish development in absence of zygotic *ezh2* gene function.

### Results

#### *Phylogenetic analysis of Ezh1.*

Amongst the about 50 histone lysine methyltransferases encoded by the mammalian genomes, *Ezh1* and *Ezh2* are the only enzymes able to perform the H3K27me2/3 methylation. Furthermore, *Ezh1* and *Ezh2* present a unique protein architecture composed of 2 SANT domains (SMART ID: SM00717), a catalytic SET domain (SMART ID: SM00317) and a pre-SET domain rich in Cysteines, N-terminal to the SET domain and comprising a CXC domain (SMART ID: SM01114) (Fig. 1A)<sup>11,29</sup>. The 3D-structure of the catalytic domain of human EZH2<sup>30,31</sup> reveals that the pre-SET domain is organized as two three-atom clusters of bound zinc coordinated by two distinct nine-residues. The first three zinc atoms are coordinated by Cysteines 528, 535, 539, 541, 548, 552, 554, and 558 together with Histidine 530. The second group of three zinc atoms is coordinated by Cysteines 565, 567, 571, 576, 578, 585, 590, 593, and 606. Each zinc binding domain contains a short helical structure ( $\alpha 1$  and  $\alpha 2$ ) which is formed just after the fourth zinc binding cysteine found in the cluster. The core of the SET domain is formed by 2 three-stranded anti-parallel  $\beta$ -sheets ( $\beta 3$ ,  $\beta 7$ ,  $\beta 8$ ) and ( $\beta 4$ ,  $\beta 5$ ,  $\beta 6$ ) diagonally pressed across each other and flanked by 2 short  $\alpha$ -helices  $\alpha 4$  and  $\alpha 5$  (Fig. 1B)<sup>30,31</sup>.

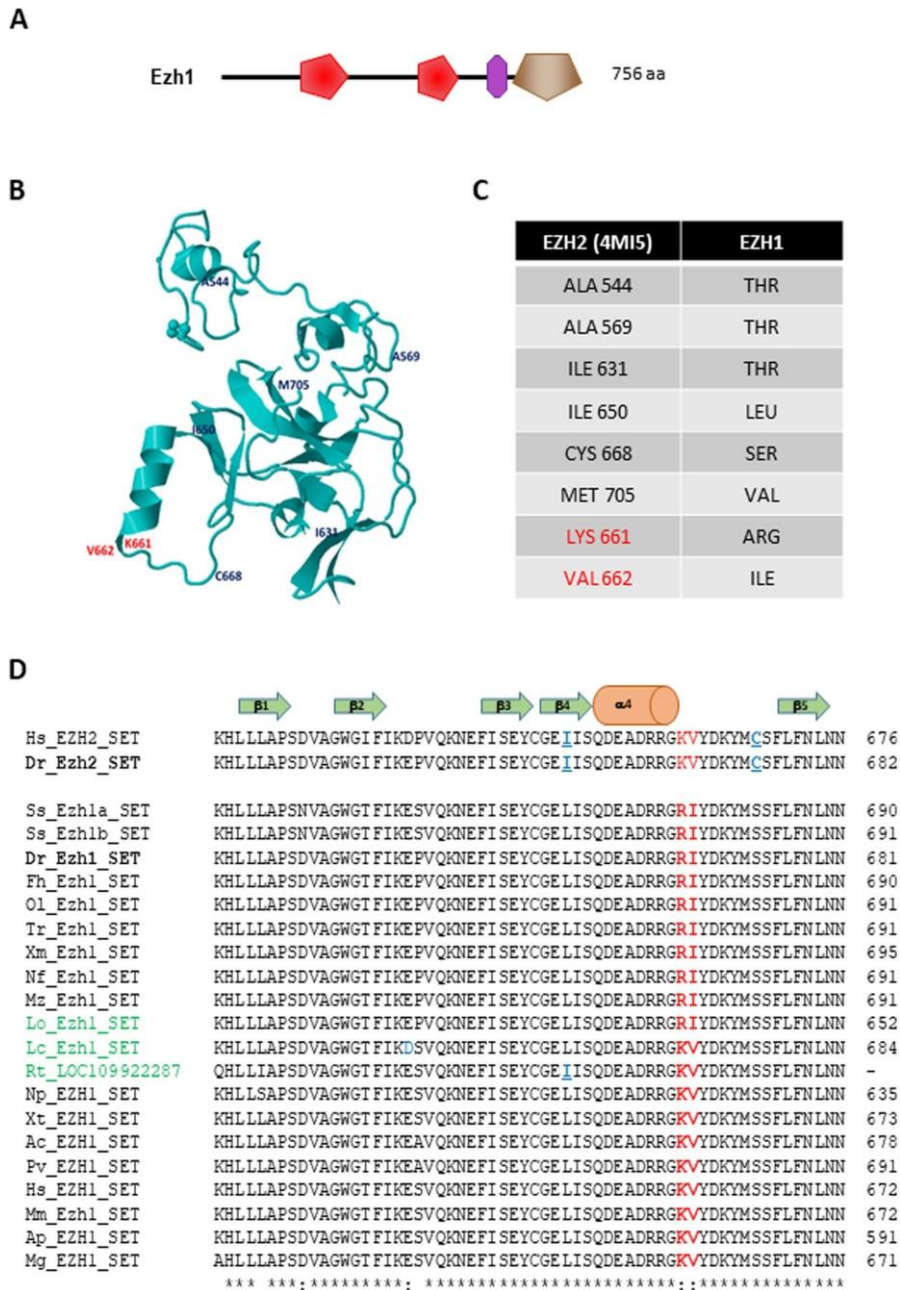
Sequence comparison between the SET domains of *Ezh1* and *Ezh2* from two mammals, the human (*Homo sapiens*) and the mouse (*Mus musculus*), two birds, the mallard (*Anas platyrhynchos*) and the turkey (*Meleagris gallopavo*), two reptiles, the green anole (*Anolis corolinensis*) and the bearded dragon (*Pogona vitticeps*), two amphibians, the western clawed frog (*Xenopus tropicalis*) and the high Himalaya frog (*Nanorana parkeri*) and two fishes, the zebrafish (*Danio rerio*) and the medaka (*Oryzias latipes*) reveals a limited number of amino-acid substitutions specifying the two different catalytic domains (Supplementary Fig. S1). These substitutions, A544T, A569T, I631T, I650L, C668S and M705V are located in different regions of the SET domain (Fig. 1B,C) and will be used to distinguish *Ezh1* from *Ezh2* in Blast searches in databases. Furthermore, we identify two additional amino-acid substitutions specific to zebrafish and medaka *Ezh1* SET domains after  $\alpha 4$ , the changes K661R and V662I (Fig. 1B,C and Supplementary Fig. S1).

To identify *Ezh1* from teleost genomes, the sequences of the SET domains from human and zebrafish *Ezh1* proteins were used in independent TBLASTN searches against genome databases (NCBI, <https://blast.ncbi.nlm.nih.gov/Blast.cgi>; Ensembl, <http://www.ensembl.org/Multi/Tools/Blast?db = core>; EFish Genomics, <https://>

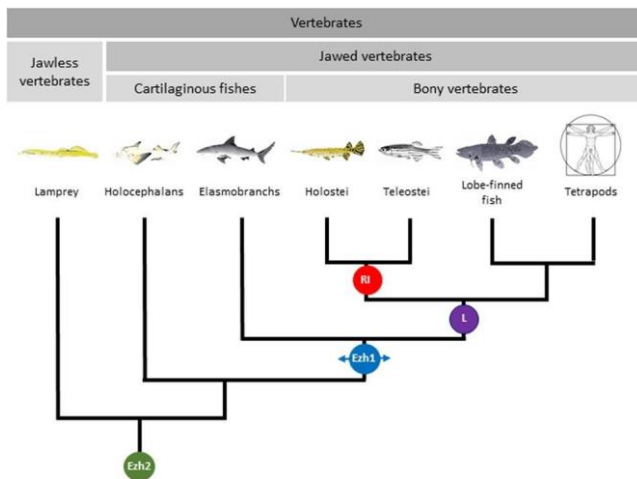
[efishgenomics.integrativebiology.msu.edu/blast\\_search/](http://efishgenomics.integrativebiology.msu.edu/blast_search/)) for 25 teleost species (Supplementary Table S1). At least one gene encoding Ezh1 could be identified in each teleostean checked (Supplementary Table S1). Two *ezh1* genes (LOC108928200 and LOC108941744) are found in the Asian arowana (*Scleropages formosus*) genome. This is due to the fact that a whole-genome duplication, termed the teleost-specific third whole-genome duplication (Ts3R) has occurred in the teleost lineage after it splits from the tetrapods<sup>32,33</sup>. Following whole-genome duplication, the resulting duplicated genomes eventually retain only a small part of the duplicated genes (onhologs). Thus, while the Asian arowana possesses two duplicated *ezh1* genes, most of the other teleosts retain only one gene. The finding that the Asian arowana has two *ezh1* onhologs parallels the fact that this fish genome contains a complete set of post-Ts3R complement for different genes including eight Hox clusters, while other teleost genomes retain less Hox clusters [7 in the zebrafish and even 5 in the African butterfly fish (*Pantodon buchholzi*)]<sup>34</sup>. The common carp (*Cyprinus carpio*) also contains two *ezh1* onhologs (LOC109102790 and LOC109057529), but arising from a more recent whole-genome duplication since the carp is a tetraploid fish containing about 100 chromosomes, approximately twice the number of most other cyprinidae, including zebrafish<sup>35,36</sup>. Finally, in agreement with a relatively recent whole-genome duplication event, the salmonid-specific fourth whole-genome duplication (Ss4R)<sup>37</sup> having occurred in the salmonidae family of teleosts fishes, the Atlantic salmon (*Salmo salar*), the rainbow trout (*Oncorhynchus mykiss*) and the Coho salmon (*Oncorhynchus kisutch*) possess two *ezh1* copies in their haploid genomes (Supplementary Table S1). The analysis of the Ezh1 SET domain sequences in teleost fishes reveals that the K661R and V662I substitutions are present in all teleostean Ezh1 proteins suggesting that these changes occurred before the Ts3R genomic event (Fig. 1D). Holostei are composed of eight living species of ray-finned fish (Actinopterygii) whose lineage diverged from teleosts before the Ts3R genome duplication. Among them, the genome of the spotted gar (*Lepisosteus oculatus*) has been sequenced<sup>38</sup>. The genome of the spotted gar contains one *ezh1* gene (NCBI Gene ID: 102693111, Supplementary Table S1) coding for a protein having a SET domain also harboring the K661R and V662I substitutions (Fig. 1D). In contrast, these two amino-acid changes were not found in the Ezh1 SET domain of the African coelacanth (*Latimeria chalumnae*) which belongs to the class of lobe-finned fish (Sarcopterygii) (Fig. 1D)<sup>39</sup>. Then, the two amino-acids changes at K661R and V662I are not specific to the teleosts. They are found in other ray-finned fish including the holosteian spotted gar, but not in the lobe-finned fish nor in the other tetrapods.



## PARTIE I : LA REPRESSION POLYCOMB



**Figure 1.** Structure of Ezh1 and its SET domain. (A) Schematic representation of the zebrafish Ezh1 protein. Red, violet and brown motifs correspond to SANT (SMART: SM00717), CXC (SMART: SM01114) and SET (SMART: SM00317) domains, respectively. The size of the protein is indicated. (B) 3D-structure of the human EZH2 SET domain (PDB ID: 4MI5) represented as a ribbon model<sup>30</sup>. The position of the substitutions identified in Ezh1 proteins are shown. Substitutions specific to ray-finned fish are in red. (C) List and nature of the substitutions identified in Ezh1 proteins are shown. Substitutions specific to ray-finned fish are in red. (D) Multiple protein alignment of the region of the SET domains between the  $\beta$ -sheets  $\beta$ 1 and  $\beta$ 5 from human (Hs) and mouse (Mm) Ezh2 and Ezh1 from several vertebrates. Hs: *Homo sapiens*; Mm: *Mus musculus*; Ss: *Salmo salar*; Dr: *Danio rerio*; Fh: *Fundulus heteroclitus*; Ol: *Oryzias latipes*; Tr: *Takifugu rubripes*; Xm: *Xiphophorus maculatus*; Nf: *Nothobranchius furzeri*; Mz: *Maylandia zebra*; Lo: *Lepisosteus oculatus*; Lc: *Latimeria chalumnae*; Rt: *Rhincodon typus*; Np: *Nanorana parkeri*; Xt: *Xenopus tropicalis*; Ac: *Anolis carolinensis*; Pv: *Pogona vitticeps*; Ap: *Anas platyrhynchos*; Mg: *Meleagris gallopavo*. The residues Ile650 and Cys668 are shown in blue and underlined in the EZH2 SET sequence, while Lys661 and Val662 are in red. Non-teleostean fish are indicated in green.



**Figure 2.** Proposed phylogeny of vertebrates based on *Ezh1-Ezh2* evolution. Proposed phylogenetic tree showing changes that have occurred in *Ezh1/2* SET domains during evolution. Ancient vertebrates contain only *Ezh2* in their genome (Green circle). *Ezh1* arose from *Ezh2* gene duplication in the common ancestor to elasmobranchs and bony vertebrates (Blue circle). Additional substitutions such as the SET I650L change occurred in the common ancestor to bony vertebrates (Violet circle). Ray-finned fish including at least teleostans and holosteans harbor specific *Ezh1* SET domain substitution such as K661R and V662I (Red circle). The drawings were done by Shaghayegh Hasanpour.

We next searched for the presence of the *ezh1* gene in the cartilaginous fish genomes. Cartilaginous fishes (Chondrichthyes) comprise two subclasses, the elasmobranchs including the sharks, the rays, the skates and sawfish, and the holocephalans composed of several chimera species, thought to have diverged from a common ancestor more than 400 million years ago (Supplementary Fig. S2). Using human and zebrafish *Ezh1* sequences in TBLASN searches, we identified two genomic sequences in the whale shark (*Rhincodon typus*) genome<sup>40</sup> (Supplementary Table S1). LOC109924829, located at one extremity of the genomic clone NW\_018048782.1, codes for a partial protein containing 2 SANT domains and a part of a CXC domain with high homology to *Ezh1*, whereas LOC109922287, within the genomic clone NW\_018043781.1, potentially contains the last 5 exons of a gene coding for an *Ezh1* SET domain (Supplementary Fig. S3). It is then likely that LOC109924829 and LOC109922287 correspond to two parts of the whale shark *ezh1* gene. Alignment of the whale shark *Ezh1* SET domain with other *Ezh1* and *Ezh2* SET domains reveals that the *Ezh1* SET domain from the whale shark does not contain the K661R and V662I substitutions specific to the *Ezh1* SET domains of the ray-finned fish. However, it possesses the *Ezh1*-specific substitutions I631T, C668S and M705V, but not I650L (Supplementary Fig. S3 and 1D). The little skate (*Leucoraja erinacea*) is the second elasmobranch for which genomic and transcriptomic sequences are available in databases<sup>41,42</sup>. Using human and zebrafish *Ezh1* sequences in a TBLASN search in the Skatebase (<http://skatebase.org>), we identified a partial transcript from the little skate (LS-transcriptB2-ctg80080) that codes for a SET domain highly related to *Ezh1*. As for the whale shark, the *Ezh1* SET domain of the little skate possesses the *Ezh1*-specific substitutions C668S and M705V, but not I650L nor the ray-finned fish-specific K661R and V662I changes (Supplementary Fig. S3).

In contrast, we failed to detect *ezh1* in a jawless vertebrate, the sea lamprey (*Petromyzon marinus*) while the *ezh2* gene could be found (ENSPMAG00000004401), suggesting that *ezh2* is more ancient and that *ezh1* might have arisen from the duplication of *ezh2*.

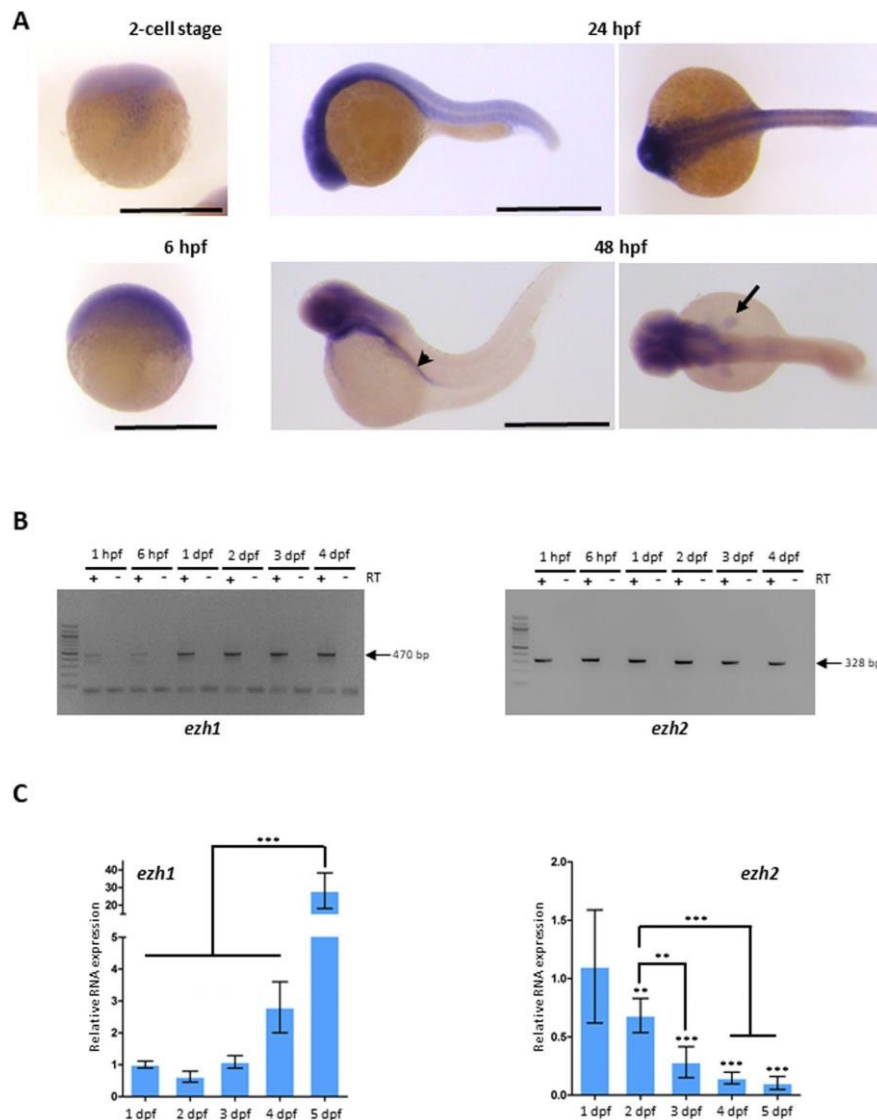
The elephant shark (*Callorhynchus milii*) is the only holocephalan having its genome sequenced<sup>43</sup>. Using TBLASTN searches, we also failed to identify an *ezh1* gene in the elephant shark genome while *ezh2* could be found (NCBI Gene ID: 103187005). This suggests that the *ezh2* duplication might have occurred after the holocephalans and elasmobranchs have diverged. It also implies that there is an ancestor common to elasmobranchs and to bony vertebrates, but not to holocephalans in which this *ezh2* duplication happened (Fig. 2).

Consequently, we propose that like in *Drosophila* and other invertebrates that have a single *ezh* gene [*E(z)*], the first vertebrates possessed only *ezh2* in their genomes. That is still the case for lampreys and chimeras (Fig. 2). *Ezh1*

harboring several amino-acid substitutions such as C668S and M705V in the SET domain, appeared in the common ancestor of elasmobranchs and bony vertebrates following an *ezh2* gene duplication. In the common ancestor of ray-finned fish, lobe-finned fish and tetrapods additional changes like the SET domain I650L substitution, were introduced in *Ezh1*. Finally, ray-finned fish including teleosts and holosteans contain other substitutions including the SET domain K661R and V662I changes (Fig. 2).

### *Ezh1* expression during zebrafish development.

In zebrafish, zygotic transcription starts at the midblastula transition at about cell cycle 10–13 [3–4 hours post fertilization (hpf)]. Before this stage, all developmental processes depend on maternally deposited gene products<sup>44,45</sup>. We studied the expression pattern of *ezh1* before the midblastula transition (2-cell stage), as well as at later developmental stages (6, 24 and 48 hpf). Whole-mount *in situ* hybridization reveals that the *ezh1* transcript is maternally loaded into the embryos since a signal could be detected before midblastula transition at the 2-cell stage (Fig. 3A). To demonstrate that the hybridization conditions are stringent enough and that the *ezh1* antisense probe used in the experiments is specific to *ezh1* transcripts and does not hybridize with mRNAs encoding other SET domain-containing histone methyltransferases including *ezh2*, we performed whole-mount *in situ* hybridization on wild-type and maternal-zygotic *MZezh1<sup>ul3/ul3</sup>* mutant (see below) embryos at the 1 to 2-cell stages (1 hpf) (Supplementary Fig. S4). A clear *ezh1* signal is detected in the wild-type embryos whereas the signal is strongly reduced in *MZezh1<sup>ul3/ul3</sup>* mutant embryos and absent when an *ezh1* sense probe is used in the experiments, thus demonstrating the specificity of the *ezh1* RNA probe (Supplementary Fig. S4). The reduction of the *ezh1* signal in *MZezh1<sup>ul3/ul3</sup>* mutant embryos probably relies on a non-sense mediated decay mechanism which appears to be less intense at later developmental stages, as observed for *ezh2* in zygotic *Zezh2<sup>ul2/ul2</sup>* mutants<sup>26</sup>. The presence of *ezh1* mRNAs in embryos before the midblastula transition is further confirmed by RT-PCR (Supplementary Fig. S5). The PCR primers used in the RT-PCR experiments amplify the *ezh1* transcript (*ezh1-001*, ENSDART00000039170.8) as well as a predicted shorter transcript (*ezh1-002*, ENSDART00000101965.6) potentially coding for an *Ezh1* isoform lacking the SANT and SET domains. However, these two transcripts could be distinguished by RT-PCR according to their size since the intron 4 is retained in the short transcript, but not in *ezh1* (*ezh1-001*) mRNA. Then, RT-PCR amplification is expected to give 470 bp and 871 bp fragments corresponding to the *ezh1* (*ezh1-001*) transcript and the predicted shorter *ezh1-002* transcript, respectively. RT-PCR experiments performed on mRNAs extracted from 1 hpf embryos identify an amplified DNA fragment at about 470 bp, but not at 871 bp suggesting that *ezh1* (*ezh1001*) transcripts, but not the predicted shorter *ezh1-002* transcripts are maternally loaded in zebrafish embryos. The sequencing of the amplicon confirms that it corresponds to *ezh1-001*, but not to *ezh2*, nor to *ezh1-002* as the intron 4 is spliced out (Fig. 3B, Supplementary Fig. S5). Altogether, our results based on both *in situ* hybridization and RT-PCR followed by sequencing, demonstrate that *ezh1* transcripts are maternally deposited into the zebrafish embryos.



**Figure 3.** *Ezh1* expression during zebrafish development. (A) *In situ* hybridization at the 2-cell stage, 6 hpf, 24 hpf and 48 hpf showing that *ezh1* transcripts are maternally provided (2-cell stage) and showing zygotic *ezh1* mRNA distribution during early development. The arrowhead shows the gut and the arrow the pectoral fin bud. Scale bar is 500  $\mu$ m. (B) RT-PCR analysis showing detection of *ezh1* and *ezh2* mRNAs at 1 hpf, 6 hpf, 1 dpf, 2 dpf, 3 dpf and 4 dpf. The reverse transcriptase (RT) was included (+) or not (-) in the reaction, as indicated. (C) RT-qPCR showing that relative *ezh1* expression increases during larval development from 1 dpf to 5 dpf whereas *ezh2* expression decreases. Three independent experiments were performed and error bars represent standard deviation. Statistical analysis was performed using a one-way ANOVA with Tukey's *post hoc* test. \*\*P < 0.01; \*\*\*P < 0.001.

At later developmental stages, zygotic *ezh1* expression is detected at the 50%-epiboly stage (6 hpf) (Fig. 3A). At 24 hpf, *ezh1* mRNAs are ubiquitously present in the embryo and expression becomes more restricted in the developing brain and in the pectoral fin buds at 48 hpf (Fig. 3A). RT-PCR experiments show that *ezh1* RNA levels are higher at 24 hpf compared to 1 or 6 hpf suggesting that maternal levels of *ezh1* transcripts are low (Fig. 3B).

Quantitative RT-qPCR reveals that *ezh1* mRNA levels increase during early larval development between 24 hpf and 5 dpf. This contrasts with *ezh2* mRNA levels which are decreasing during the same larval developmental period (Fig. 3C).

#### TALEN-mediated *ezh1* inactivation in zebrafish.

To gain insights into the function of *ezh1* in zebrafish development, we generated *ezh1* loss-of-function mutants using the transcription activator-like effector nuclease (TALEN)-based technology. TALENs were designed to

target the third exon of *ezh1* in order to introduce a frameshift upstream of all known Ezh1 conserved domains, such as the SANT, CXC and SET domains. In addition, the targeted region was chosen to contain a BamHI restriction site used to screen for mutations and for genotyping analyses (Fig. 4A). TALENs were assembled using the Golden Gate Cloning methodology<sup>46</sup> and *in vitro* transcribed mRNAs encoding each TALEN pair were injected into one-cell stage embryos. At 3 days post injection, genomic DNA was extracted from single embryos and PCR amplification of the targeted region, followed by BamHI digestion revealed the efficacy of the TALENs and of the BamHI restriction strategy for genotyping (Fig. 4B).

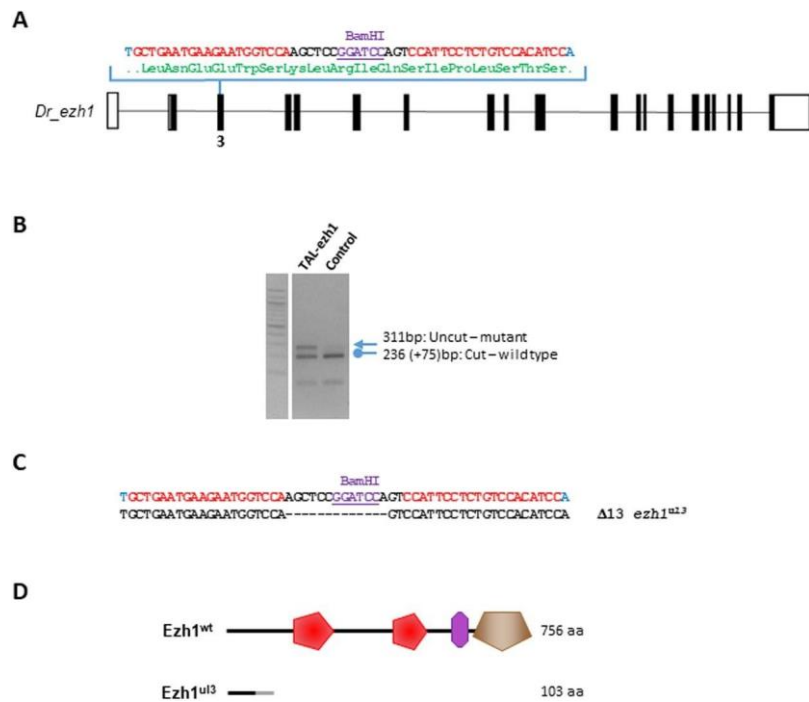
Restriction analysis of genomic DNA showed that the mutation rate at the *ezh1* locus was 100% (29 out of 29 injected embryos tested). Then, we raised TALENs-injected embryos to establish an adult F0 founder population. To evaluate the efficiency of germ line transmission of the mutations, individual F0 fish carrying *ezh1* mutations (loss of the BamHI restriction site) were outcrossed to wild-type TU to obtain F1 offspring. Genomic DNA was isolated from individual F1 embryos from each F0 fish and analyzed by BamHI restriction. Embryos from 3 of 8 individual F0 fish were heterozygous mutants, demonstrating successful germ line transmission of the mutations. Among them, one mutation causes a 13 bp deletion leading to a frame shifting of the coding sequence and appearance of a premature stop codon (Fig. 4C). This  $\Delta 13$  *ezh1* allele, hereafter called *ezh1<sup>ul3</sup>*, codes for a predicted protein of 103 amino acids, lacking all conserved protein domains (Fig. 4D, Supplementary Fig. S6) and was selected to raise a heterozygous *ezh1<sup>+/ul3</sup>* zebrafish line. The heterozygous *ezh1<sup>+/ul3</sup>* fish are viable, fertile and do not show any phenotype. Among siblings from heterozygous *ezh1<sup>+/ul3</sup>* incrosses, we successfully identified males and females carrying the homozygous  $\Delta 13$  mutation (*ezh1<sup>ul3/ul3</sup>*). The homozygous mutants are viable and fertile allowing us to generate a maternal-zygotic *MZezh1<sup>ul3/ul3</sup>* line used for further phenotypic studies. Fish from this *MZezh1<sup>ul3/ul3</sup>* line are still fertile after at least two incross generations and do not present any phenotype (Fig. 5A) demonstrating that both maternal and zygotic *ezh1* products are dispensable to zebrafish development.

Since Ezh1 is a histone methyltransferase able to promote H3K27me2/3 methylation, we next evaluated the H3K27me2/3 status of the *ezh1<sup>ul3/ul3</sup>* line. Total histones were extracted from mutant and wildtype larvae at 9 dpf and H3K27me3 was analyzed by western blot using specific anti-H3K27me3 and anti-H3K27me2 antibodies. Figure 5B reveals that global levels of H3K27me2/3 are similar in *MZezh1<sup>ul3/ul3</sup>* and wild-type total larvae, indicating that loss of *ezh1* function does not affect global levels of H3K27me2/3 methylation in zebrafish.

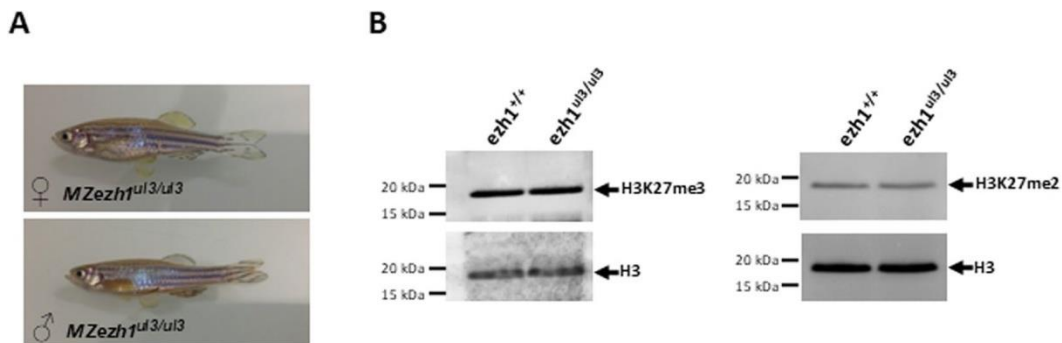
### *Loss of ezh1 function leads to an increase in ezh2 expression.*

Given that H3K27me2/3 levels are not decreased in *ezh1<sup>ul3/ul3</sup>* mutant fish, we next tested if *ezh2* expression is changed in absence of *ezh1* function. At 48 hpf, *ezh2* expression is restricted to specific regions such as the pectoral fin buds, the optic tectum, the mid-hindbrain region, the branchial arches and the retina (Fig. 6A)<sup>26,28</sup>. Whole-mount *in situ* hybridization seems to show a slight increase in the *ezh2* signal in the pectoral fin buds and at both sides of the mid-hindbrain region in *MZezh1<sup>ul3/ul3</sup>* embryos at 48 hpf when compared to wild-type (Fig. 6A). This increase in *ezh2* expression in *MZezh1<sup>ul3/ul3</sup>* mutant was confirmed by RT-qPCR on mRNAs extracted from embryos from 1 to 5 dpf. Quantification of transcripts shows a significant increase in *ezh2* mRNA abundance in *MZezh1<sup>ul3/ul3</sup>* mutants when compared to wild-type at 2-, 3-, 4- and 5 dpf (Fig. 6B).





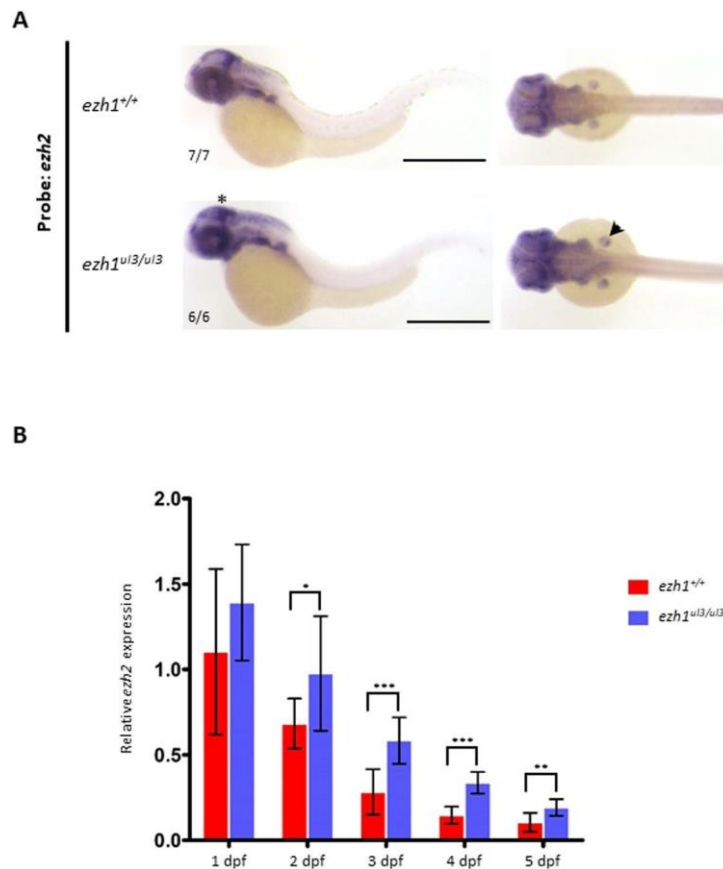
**Figure 4.** Generation of *ezh1* mutant zebrafish using the TALEN technology. (A) Schematic representation of the genomic structure of the *ezh1* gene, with coding and untranslated sequences depicted as solid and open boxes, respectively. The location of the *ezh1* TALEN in exon 3 (3) is indicated. The *ezh1* TALEN target sequence with Left and Right TALEN binding sites in red is shown. The BamHI restriction site is underlined and indicated in violet. (B) Identification of mutant embryos using restriction fragment length polymorphism. Genomic DNA was prepared from an uninjected (Control) and an *ezh1* TALEN injected (TAL-*ezh1*) embryo, amplified by PCR and subjected to BamHI digestion. The TAL-*ezh1* injected embryo contains undigested material (arrow at 311 bp), indicating that the BamHI diagnostic restriction site has been disrupted. (C) Sequence of the mutant allele compared to its wild-type counterpart. Dashes indicate deleted nucleotides. The mutated *ezh1<sup>ul3</sup>* allele has a deletion of 13 nucleotides. (D) Schematic representation of predicted the wild-type (*Ezh1<sup>wt</sup>*) and mutant (*Ezh1<sup>ul3</sup>*) proteins. The gray line in the predicted mutant protein indicates residues read out of frame prior to encountering a premature stop codon, whereas the red, violet and brown motifs in the wildtype protein correspond to SANT (SMART: SM00717), CXC (SMART: SM01114) and SET (SMART: SM00317) domains, respectively. Sizes of the predicted proteins are indicated.



**Figure 5.** *Ezh1 ul3/ul3* zebrafish mutants present a normal phenotype. (A) Pictures of 6 month-old *MZezh1<sup>ul3/ul3</sup>* mutant female (top) and male (bottom) zebrafish from the F2 generation of incross mutant fish. (B) Global trimethylation (H3K27me3) and demethylation (H3K27me2) of lysine 27 of histone H3 is not affected in *ezh1<sup>ul3/ul3</sup>* mutants. Total histones from pools of 10 larvae at 9 dpf from wild-type or *ezh1<sup>ul3/ul3</sup>* crosses as indicated, were extracted and 5 µg of total histones were analyzed by western blotting using specific antiH3K27me3 or anti-H3K27Me2 antibodies and an anti-H3 antibody as a control.

**Ezh1 contributes to zebrafish development in absence of *ezh2*.** We have previously generated an *ezh2<sup>+/ul2</sup>* mutant line using the TALEN technology<sup>26</sup>. *Ezh2*-deficient zebrafish mutants present a normal body plan but die at around 12 dpf with defects in the intestine wall, due to enhanced cell death. In order to uncover the role of *Ezh1* in zebrafish

development, lines harboring mutations in both *ezh1* and *ezh2* genes were generated. Fish from the lines *ezh1*<sup>+/*ul3*</sup> *ezh2*<sup>+/*ul2*</sup> and *ezh1*<sup>*ul3/ul3*</sup> *ezh2*<sup>+/*ul2*</sup> are viable, fertile, healthy and do not show any phenotype.



**Figure 6.** Loss of *ezh1* function leads to an increase in *ezh2* expression. **(A)** *In situ* hybridization showing *ezh2* expression in representative embryos at 48 hpf from wild-type (Top) and *ezh1*<sup>*ul3/ul3*</sup> crosses. A slight increase in the *ezh2* signal at both sides of the midbrain-hindbrain boundary (asterisk) and in the pectoral fin buds (arrowhead) of *MZezh1*<sup>*ul3/ul3*</sup> embryos is observed. Scale bar is 500  $\mu$ m. The numbers indicate the number of embryos with the displayed compared to the total number of embryos analyzed. **(B)** Expression analysis of *ezh2* in wild-type and *MZezh1*<sup>*ul3/ul3*</sup> embryos and larvae at 1, 2, 3, 4 and 5 dpf quantified by RT-qPCR show an increase of *ezh2* expression in *MZezh1*<sup>*ul3/ul3*</sup> mutants. Three independent experiments were performed and statistical analysis was conducted using Student's t-test. \*P < 0.05; \*\*P < 0.01; \*\*\*P < 0.001.

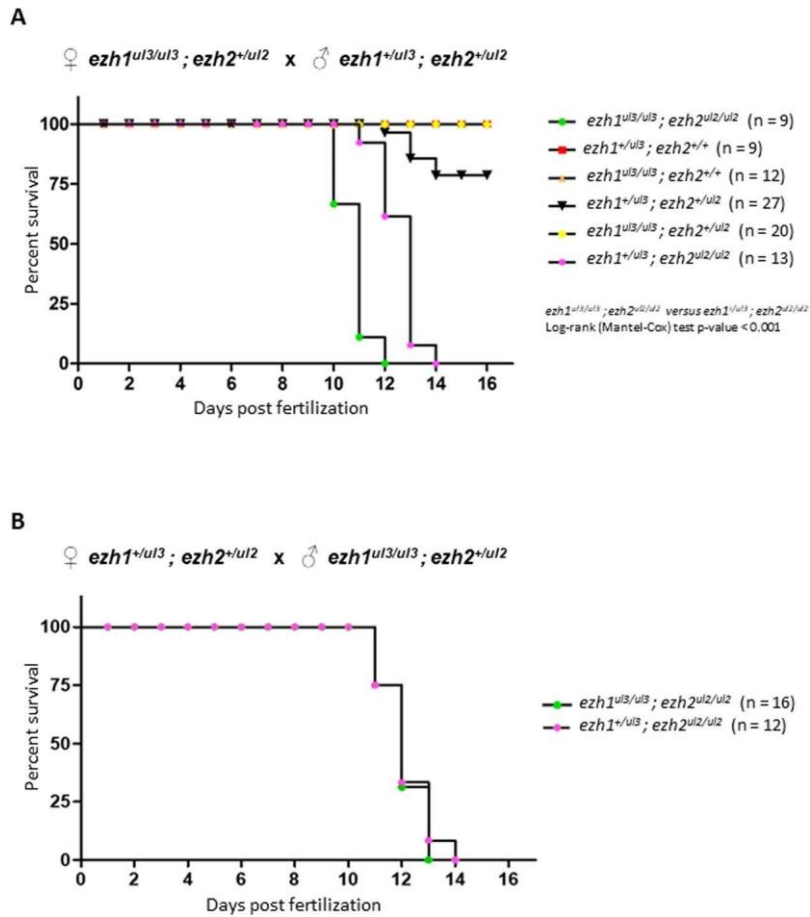
To study the role of Ezh1 in absence of Ezh2, incross of *ezh1*<sup>*ul3/ul3*</sup> *ezh2*<sup>+/*ul2*</sup> fish was performed and larvae were genotyped at 11 dpf. At this stage, we failed to identify larvae with the *ezh1*<sup>*ul3/ul3*</sup>; *ezh2*<sup>*ul2/ul2*</sup> genotype out of 32 individuals checked, indicating that loss of maternal and zygotic *ezh1* expression could be responsible for a premature death of *ezh2*<sup>*ul2/ul2*</sup> larvae.

To further investigate the effect of *ezh1* deficiency on the survival of *ezh2*<sup>*ul2/ul2*</sup> larvae, we performed Kaplan-Meier survival curves. First, *ezh1*<sup>*ul3/ul3*</sup>; *ezh2*<sup>+/*ul2*</sup> females were crossed with *ezh1*<sup>+/*ul3*</sup>; *ezh2*<sup>+/*ul2*</sup> males. At 3dpf, siblings were genotyped and embryos were pooled into distinct tanks according to their genotype and their survival followed to generate Kaplan-Meier plots (Fig. 7A). As previously shown for *ezh2*<sup>*ul2/ul2*</sup> mutants<sup>26</sup>, lethality of *ezh1*<sup>+/*ul3*</sup> *ezh2*<sup>*ul2/ul2*</sup> fish occurs at around 12 dpf (50% survival: 13 dpf). In contrast, mutants lacking *ezh1* function (*MZezh1*<sup>*ul3/ul3*</sup>; *ezh2*<sup>*ul2/ul2*</sup>) are dying about two days earlier (50% survival: 11 dpf) indicating that maternal and/or zygotic *ezh1* expression contributes to zebrafish development in the absence of zygotic *ezh2* function. However, no morphological differences between *ezh1*<sup>+/*ul3*</sup> *ezh2*<sup>*ul2/ul2*</sup> and *MZezh1*<sup>*ul3/ul3*</sup>; *ezh2*<sup>*ul2/ul2*</sup> larvae could be observed.

To determine the role of maternal *ezh1* expression in the survival of *ezh2*-deficient larvae, a cross between *ezh1*<sup>+/*ul3*</sup> *ezh2*<sup>+/*ul2*</sup> females and *ezh1*<sup>*ul3/ul3*</sup>; *ezh2*<sup>+/*ul2*</sup> males was set up to generate paternal-zygotic (PZ) *PZezh1*<sup>*ul3/ul3*</sup>; *ezh2*<sup>*ul2/ul2*</sup> mutants. In this situation, all embryos will benefit from *ezh1* maternal products. Figure 7B reveals that in presence

## PARTIE I : LA REPRESSION POLYCOMB

of maternal *ezh1* products, *ezh1*<sup>+/<sup>ul3</sup></sup>; *ezh2*<sup>ul2/ul2</sup> and *ezh1*<sup>ul3/ul3</sup>; *ezh2*<sup>ul2/ul2</sup> have similar survival rates indicating that *ezh1* products from maternal origin are crucial for larval development in absence of *ezh2*.



**Figure 7.** *Ezh1* contributes to zebrafish development in absence of *ezh2*. Kaplan-Meier survival curves over 16 days for siblings from the indicated genotypes from crosses between ♀*ezh1*<sup>ul3/ul3</sup>; *ezh2*<sup>+/<sup>ul2</sup></sup> and ♂*ezh1*<sup>+/<sup>ul3</sup></sup>; *ezh2*<sup>+/<sup>ul2</sup></sup> (A) and from crosses between ♀*ezh1*<sup>+/<sup>ul3</sup></sup>; *ezh2*<sup>+/<sup>ul2</sup></sup> and ♂*ezh1*<sup>ul3/ul3</sup>; *ezh2*<sup>+/<sup>ul2</sup></sup> (B). The number of fishes of each genotype is indicated. Statistical significance was assessed using a log-rank (Mantel-Cox) test.

### Discussion

Considering that *Drosophila* and other invertebrates have a single gene coding for the H3K27me2/3 histone methyltransferase while mammalian genomes encode two orthologs, we searched for the phylogenetic origin of *Ezh1* and *Ezh2* in vertebrate evolution. We found that *ezh2*, but not *ezh1* is present in the sea lamprey genome suggesting that the vertebrate ancestors contained only one gene, *ezh2*, encoding this histone lysine methyltransferase activity. *Ezh1* might have then arisen through *ezh2* gene duplication during evolution. This is in agreement with the fact that *Ezh2* being the closest to its invertebrate homologs<sup>11</sup>.

In the cartilaginous fish, only *ezh2* was found in the elephant shark genome whereas both *ezh1* and *ezh2* were identified in the genomes of the whale shark and of the little skate. This implies that the current vertebrate phylogenetic tree based on fossils and proposing that holocephalans and elasmobranchs diverged from a common ancestor distinct from the bony vertebrate ancestor should be revised. We propose that holocephalans including the elephant shark diverged before the appearance of a common ancestor of the elasmobranchs including the whale shark and the little skate and of the bony vertebrates including the ray-finned fish, the lobe-finned fish and the tetrapods in which the *ezh2* gene duplication occurred. Obviously, this hypothesis should be tested once genomic data from other lampreys, chimeras and sharks will be available.



A whole-genome duplication (Ts3R) and a subsequent massive loss of duplicated orthologs occurred in the teleost lineage<sup>32,33</sup>. As a consequence of these genomic events, some genes were differentially lost in different teleost species. For instance, Ring1 and Rnf2 are components of the PRC1 protein complex and are the E3 ubiquitin-protein ligases that mediate monoubiquitination of Lysine 119 of histone H2A (H2AK119ub1)<sup>10,47</sup>. Analysis of teleost genomes revealed that both *ring1* and *rnf2* genes are present in the medaka genome, but *ring1* has been lost in zebrafish, in green spotted puffer (*Tetraodon nigroviridis*) and in Japanese puffer (*Takifugu rubripes*) suggesting that *ring1* and *rnf2* are redundant and that Rnf2 function alone is sufficient and could bypass Ring1 requirement<sup>48</sup>. A search for *ezh1* in 25 teleost species having their genome sequenced shows that the gene is present even in fish with particularly small genome size from the tetraodontidae family (e.g., the pufferfish *Takifugu rubripes* and *Tetraodon nigroviridis*). Since gene duplications can provide the raw material on which evolution can act, since they lead to redundant gene copies that are freed up to evolve novel gene functions<sup>49</sup>, one could then argue that Ezh1 and Ezh2 have at least slightly different functions. However, in the green spotted puffer (*Tetraodon nigroviridis*), a gene coding for *ezh1* could be identified (ENSTNIG00000011694) while *ezh2* seems to be absent. This shows that Ezh1 and Ezh2 are functionally redundant and it indicates that *ezh1* could bypass the *ezh2* loss in this pufferfish.

To question the function of *ezh1*, we use the zebrafish model and first investigate *ezh1* expression during development. Previous reports were not able to detect *ezh1* mRNAs in zygotes and embryos prior to maternal to zygotic transition neither by *in situ* hybridization<sup>50,51</sup> nor by RNA sequencing in a large scale approach<sup>52</sup>, but we do. Using *in situ* hybridization, we could detect maternally loaded *ezh1* transcripts at the 1- to 2-cell stages, and the specificity of the *ezh1* probe was assessed by showing that *in situ* hybridization experiments on embryos from *ezh1<sup>ul3/ul3</sup>* incrosses gave a reduced *ezh1* signal. Furthermore, RT-PCR assays followed by sequencing of the amplicon confirmed the presence of *ezh1* mRNAs in the zebrafish before the midblastula transition, even if the maternal *ezh1* mRNA levels seem to be low. Our identification of maternally deposited *ezh1* transcripts in zebrafish embryos then contradicts the results from previous studies<sup>50-52</sup>. A possible explanation to this discrepancy could be linked to the considerable genetic variation between zebrafish strains<sup>53</sup>. Indeed, in the course of our studies, we identified a polymorphism within our TU strain at the *ezh1* locus in the region of exon 2/intron 2-3 (Supplementary Fig. S7). Even if we have no evidence indicating that this polymorphism alters *ezh1* expression or mRNA stability, one might be tempted to speculate that interstrain genetic variability may affect gene expression especially for genes coding for non-essential proteins.

At 24 hpf, the *ezh1* mRNAs are ubiquitously distributed in the embryo but *ezh1* expression becomes more restricted in the developing brain and in the pectoral fin buds at 48 hpf. However, quantitative RT-qPCR reveals that *ezh1* mRNA levels increase during larval development between 24 hpf and 5 dpf, whereas *ezh2* mRNA levels decrease during the same period, as previously described<sup>28</sup>. These results parallel experiments performed in mouse suggesting that PRC2-Ezh2 complexes would exert their functions in proliferative tissues while PRC2-Ezh1 would have a more important role in more differentiated cells<sup>13,17</sup>.

To further investigate the Ezh1 function, we generated an *ezh1*-loss of function zebrafish line using the TALEN technology. Homozygous *ezh1<sup>ul3/ul3</sup>* zebrafish are viable and fertile, indicating that maternal and zygotic *ezh1* products are dispensable to zebrafish development. It is similar to what was observed in mouse since *Ezh1*-deficient mice are viable, fertile and healthy<sup>54-58</sup>. Also, as in mammalian cells in culture where it has been shown that knockdown of *Ezh1* does not result in global reduction of H3K27me3 levels<sup>13</sup>, these global levels are not affected in zebrafish larvae deficient for *ezh1*. Investigation of *ezh2* expression in *ezh1<sup>ul3/ul3</sup>* mutants reveals that *ezh2* levels are increased when *ezh1* expression is lost. This increase in *ezh2* expression in addition to the functional redundancy of Ezh1 and Ezh2, could account at least in part, for the absence of phenotype and for H3K27me2/3 levels unchanged in *MZezh1<sup>ul3/ul3</sup>* mutant zebrafish.

Using Cre-mediated conditional mutagenesis in mouse, it has been shown that in absence of Ezh2 function, Ezh1 plays a role in hair follicle formation and maintenance<sup>55</sup>, in hepatic homeostasis and regeneration<sup>57</sup>, in skeletal growth<sup>56</sup>, as well as in the pathogenesis of hematopoietic malignancies<sup>58</sup>. Here, we show loss of maternal and zygotic *ezh1* expression impairs the survival of *ezh2*-deficient zebrafish larvae. Furthermore and consistent with the fact that *ezh1* transcripts are maternally provided, we show that *ezh1* maternal products are involved in this effect.

Although Ezh1 and Ezh2 have been shown to be functionally redundant *in vitro*<sup>13,17</sup>, it remains unclear whether this is the case in the context of tissue development and homeostasis. Here we show on one hand that *ezh1*<sup>ul3/ul3</sup> are viable and fertile and on the other hand that loss of both maternal-zygotic *ezh1* and zygotic *ezh2* functions results in an earlier mortality of zebrafish larvae when compared to zygotic *ezh2*-deficient fish. These findings underscores the functional redundancy between Ezh2 and Ezh1 in controlling zebrafish larval development.

### Methods

#### *Database searches and analyses.*

Using the sequences of human and zebrafish Ezh1 and Ezh2 SET domains as search queries, TBLASTN analyses were performed using the NCBI (<https://blast.ncbi.nlm.nih.gov/Blast.cgi>), the Ensembl (<http://www.ensembl.org/Multi/Tools/Blast?db=core>), the EFish Genomics ([https://efishgenomics.integrativebiology.msu.edu/blast\\_search/](https://efishgenomics.integrativebiology.msu.edu/blast_search/)) or the Skatebase (<http://skatebase.org/skateblast-skatebase%E2%80%8B/>) Genome Servers. Amino acid sequences of protein domains of various species were obtained from SMART database (<http://smart.embl-heidelberg.de/>). Multiple sequence alignments were performed using Clustal Omega (<https://www.ebi.ac.uk/Tools/msa/clustalo/>) and 3D-structure analysis of EZH2 SET obtained from the Protein Data Bank (<https://www.rcsb.org/>) with Jalview (<http://www.jalview.org/>).

#### *Zebrafish maintenance, embryo preparation and animal ethics statements.*

Zebrafish (TU strain) were maintained at 27.5 °C in a 14/10 h light/dark cycle. The evening before spawning, males and females were separated into individual breeding tanks (Tecniplast). Spontaneous spawning occurred the following morning when the light turned on and the plastic divider removed. Embryos or larvae were then collected and staged according to Kimmel *et al.*<sup>59</sup>. The chorions were removed from embryos by the action of 1% pronase (Sigma) for 1 min. Zebrafish embryos or larvae were fixed overnight in 4% paraformaldehyde in PBS (phosphate-buffered saline, Invitrogen), dehydrated gradually to 100% methanol and kept at -20 °C.

Zebrafish were handled in compliance with local care regulations according to the French and European Union guidelines for the handling of laboratory animals (Directive 2010/63/EU of the European Parliament and of the Council of 22 September 2010 on the protection of animals used for scientific purposes). The experimental procedures carried out on zebrafish were reviewed and approved by the local Ethics Committee, CEEA 75 Nord Pas-de-Calais (APAFiS approval number 2018011722529804).

#### *TALEN design and assembly.*

A TALEN pair was designed to target exon 3 of *ezh1* using the online TAL Effector-Nucleotide Targeter tool (<https://tale-nt.cac.cornell.edu/>)<sup>60</sup>. The target site also contains a BamHI restriction site used to determine TALEN efficacy and for genotyping purposes based on restriction fragment length polymorphism analysis.

*Ezh1*-specific TALEN vectors were constructed using the TALEN Golden Gate assembly system described by Cermak *et al.*<sup>46</sup>. The TALEN expression backbones, pCS2TAL3DD and pCS2TAL3RR<sup>20</sup>, and the plasmids providing repeat variable diresidues (NI, HD, NG and NN recognizing A, C, T and G bases, respectively)<sup>46</sup> for Golden Gate Cloning were obtained from Addgene.

#### *mRNA injection into zebrafish embryos.*

Capped mRNAs were synthesized using the SP6 mMESSAGE mMACHINE kit (Ambion) from linearized plasmid templates. mRNAs (50–100 pg) were injected into 1-cell zebrafish embryos using a FemtoJet microinjector (Eppendorf).

#### *Genotype analyses.*

To achieve genotyping, DNA was extracted using sodium hydroxide and Tris. Pieces of caudal fin, or paraformaldehyde-fixed embryos and larvae were incubated in 20 µL 50 mM NaOH and heated for 20 min at 95

°C. The tubes were then cooled to 4 °C and 2 µL of 1 M Tris-HCl, pH7.4 was added to neutralize the basic solution. Genotype analysis of *ezh1* was performed by PCR on 2.5 µL of samples using the primer set *ezh1\_5\_4892fwd* (5'-GGCTCTGTTCCAGTCAAATCGCCGT-3') and *ezh1\_3\_5202rev* (5'-AGCTTTGGGAAATGGCGAGGCAAA-3') followed by PCR product digestion with the BamHI restriction enzyme. Genotype analysis of *ezh2* was performed using the primer set *TAL\_ezh2\_5'\_S21Ac* (5'-GGTATGGTTGTTGCAGTTCACAGAC-3') and *TAL\_ezh2\_3'\_S21Ac* (5'-AACACCAAACCTCTACACAAGCAGCA-3') and digestion with DdeI<sup>26</sup>.

Sequence determinations (GATC-Biotech, Germany) were performed after cloning of the PCR products into pCR2.1-TOPO (Thermofisher) according to the manufacturer's instructions.

### *Whole-mount in situ hybridization.*

Antisense- and sense-RNA probes were synthesized with the DIG RNA Labeling Kit (SP6/T7) (Roche, 11175025910), according to manufacturer's instructions. The cDNA clone MGC:152758 IMAGE:2639510 purchased at imaGenes GmbH (Berlin) was used as a template for *ezh2*. *ezh1* antisense and sense probes were generated using RT-PCR from total mRNA extracted from zebrafish larvae at 5 dpf using the RNeasy Mini Kit (Qiagen). After Reverse Transcription, cDNAs were amplified by PCR using the probe specific primers, coupled to the T7 sequence for forward primers and the SP6 sequence for reverse primers. The primers used for *ezh1* probe generation were:

ISH\_*ezh1*\_F: TAATACGACTCACTATAGGGGAGGAAGCGACCACGAAACCACC

ISH\_*ezh1*\_R: GATTTAGGTGACACTATAGGGAGACCTGTTTGCTGTCCCAGT

*In situ* hybridization experiments were performed as previously described<sup>61</sup>. The embryos were imaged using a Leica MZ10F stereomicroscope equipped with a Leica DFC295 digital camera.

### *Histone extraction and western blot analysis.*

Histone extraction and Western blot analyses were performed from 10 larvae as described previously<sup>62</sup>.

Primary antibodies used were mouse anti-H3K27me3 (1:1,000; ab6002, Abcam), rabbit anti-H3K27me2 (1:500; ab24684, Abcam) and rabbit anti-H3 (1:5,000; ab1791, Abcam). The secondary antibodies were peroxidase conjugated anti-mouse antibody (1:10,000; 115-035-003, Jackson ImmunoResearch) and peroxidase conjugated anti-rabbit antibody (1:10,000; 711-035-152, Jackson ImmunoResearch).

### *RNA extraction and RT-PCR.*

Total RNAs were purified from 1 hpf, 6, hpf, 1 dpf, 2 dpf, 3 dpf, 4 dpf and 5 dpf wild-type and mutant embryos or larvae using Trizol as previously described<sup>62</sup>. cDNA was synthesized using Superscript III (18080-044, Invitrogen) according to manufacturer's instructions. About 40 embryos were used in total RNA extractions and 1 µg total RNAs were used to perform the reverse transcription experiments. Primers used were:

Dr\_*ezh1*\_cDNA\_5a:5'-CGTCTAGTGAGGTCTGAGGATG-3'

Dr\_*ezh1*\_cDNA\_3a: 5'-CCTCGTCCTGTTCCAACACTTC-3'

ZA027\_*ezh2*\_cDNA\_5:5'-GAGGTGAAAGGACCCTCTACC-3'

ZA027\_*ezh2*\_cDNA\_3:5'-CTCAGTTTCCATTCTGATTTAAG-3'

Dr\_cDNA\_ube2a\_F: 5'-TGACTGTTGACCCACCTTACAG-3'

Dr\_cDNA\_ube2a\_R: 5'-CAAATAAAAGCAAGTAACCCCG-3'

Sequence determination (GATC-Biotech, Germany) was performed after cloning of the PCR products into pCR2.1-TOPO (Thermofisher) following the manufacturer's instructions.

PCR reactions were performed as follow: 95 °C 4 min, [95 °C 45 sec, 55 °C 45 sec, 72 °C 1 min] 40 cycles, 72 °C 10 min.

The quantitative qPCR reaction was performed in triplicate using a Bio-Rad CFX96 Real-Time System using SYBR Green Supermix (Bio-Rad). Relative mRNA expression of each gene was normalized to *ube2a* levels.

### *Kaplan-Meier analysis.*

At 3 dpf, the tip of the caudal fin was transected within the pigment gap distal to the circulating blood for genotyping purposes<sup>62</sup>. The embryos were pooled into separate 1-liter tanks according to their genotype and placed into an incubator set up at 28 °C. Larvae were fed from 5 dpf on with early stage zebrafish nutrition (Gemma Micro ZF 75, Planktovie) three times per day and regularly checked directly in the tanks using a Leica EZ4 stereomicroscope with no or minimal manipulation, during a period of 16 days post fertilization. Larvae were declared dead when heart beats could not be detected under stereomicroscopic examination. Kaplan-Meier plots were generated and log-rank (Mantel-Cox) test p-values calculated with Prism (GraphPad Software).

### Data Availability

All relevant data are within the paper and its Supplementary Information Files.

### References

1. Di Croce, L. & Helin, K. Transcriptional regulation by Polycomb group proteins. *Nat. Struct. Mol. Biol.* **20**, 1147–1155 (2013).
2. Margueron, R. & Reinberg, D. The Polycomb complex PRC2 and its mark in life. *Nature.* **469**, 343–349 (2011).
3. Sauvageau, M. & Sauvageau, G. Polycomb group proteins: multi-faceted regulators of somatic stem cells and cancer. *Cell Stem Cell.* **7**, 299–313 (2010).
4. Surface, L. E., Thornton, S. R. & Boyer, L. A. Polycomb group proteins set the stage for early lineage commitment. *Cell Stem Cell.* **7**, 288–298 (2010).
5. Cao, R. *et al.* Role of histone H3 lysine 27 methylation in Polycomb-group silencing. *Science.* **298**, 1039–1043 (2002).
6. Czermin, B. *et al.* Drosophila enhancer of Zeste/ESC complexes have a histone H3 methyltransferase activity that marks chromosomal Polycomb sites. *Cell.* **111**, 185–196 (2002).
7. Kuzmichev, A., Nishioka, K., Erdjument-Bromage, H., Tempst, P. & Reinberg, D. Histone methyltransferase activity associated with a human multiprotein complex containing the Enhancer of Zeste protein. *Genes Dev.* **16**, 2893–2905 (2002).
8. Müller, J. *et al.* Histone methyltransferase activity of a Drosophila Polycomb group repressor complex. *Cell.* **111**, 197–208 (2002).
9. Francis, N. J., Kingston, R. E. & Woodcock, C. L. Chromatin compaction by a polycomb group protein complex. *Science.* **306**, 1574–1577 (2004).
10. Wang, H. *et al.* Role of histone H2A ubiquitination in Polycomb silencing. *Nature.* **431**, 873–878 (2004).
11. Laible, G. *et al.* Mammalian homologues of the Polycomb-group gene Enhancer of zeste mediate gene silencing in Drosophila heterochromatin and at *S. cerevisiae* telomeres. *EMBO J.* **16**, 3219–3232 (1997).
12. Cao, R. & Zhang, Y. SUZ12 is required for both the histone methyltransferase activity and the silencing function of the EED-EZH2 complex. *Mol. Cell.* **15**, 57–67 (2004).
13. Margueron, R. *et al.* Ezh1 and Ezh2 maintain repressive chromatin through different mechanisms. *Mol. Cell.* **32**, 503–518 (2008).
14. Nekrasov, M., Wild, B. & Müller, J. Nucleosome binding and histone methyltransferase activity of Drosophila PRC2. *EMBO Rep.* **6**, 348–353 (2005).
15. Pasini, D., Bracken, A. P., Jensen, M. R., Lazzerini Denchi, E. & Helin, K. Suz12 is essential for mouse development and for EZH2 histone methyltransferase activity. *EMBO J.* **23**, 4061–4071 (2004).
16. Tie, F., Stratton, C. A., Kurzhals, R. L. & Harte, P. J. The N terminus of Drosophila ESC binds directly to histone H3 and is required for E(Z)-dependent trimethylation of H3 lysine 27. *Mol. Cell. Biol.* **27**, 2014–2026 (2007).
17. Shen, X. *et al.* EZH1 mediates methylation on histone H3 lysine 27 and complements EZH2 in maintaining stem cell identity and executing pluripotency. *Mol. Cell.* **32**, 491–502 (2008).
18. Völkel, P., Dupret, B., Le Bourhis, X. & Angrand, P. O. Diverse involvement of EZH2 in cancer epigenetics. *Am. J. Transl. Res.* **7**, 175–193 (2015).
19. Calebiro, D. *et al.* Recurrent EZH1 mutations are a second hit in autonomous thyroid adenomas. *J. Clin. Invest.* **126**, 3383–3388 (2016).
20. Dahlem, T. J. *et al.* Simple methods for generating and detecting locus-specific mutations induced with TALENs in the zebrafish genome. *PLoS Genet.* **8**, e1002861 (2012).
21. Bedell, V. M. *et al.* *In vivo* genome editing using a high-efficiency TALEN system. *Nature.* **491**, 114–118 (2012).
22. Moore, F. E. *et al.* Improved somatic mutagenesis in zebrafish using transcription activator-like effector nucleases (TALENs). *PLoS One.* **7**, e37877 (2012).
23. Chang, N. *et al.* Genome editing with RNA-guided Cas9 nuclease in zebrafish embryos. *Cell Res.* **23**, 465–472 (2013).
24. Hwang, W. Y. *et al.* Efficient genome editing in zebrafish using a CRISPR-Cas system. *Nat. Biotechnol.* **31**, 227–229 (2013).

25. Xiao, A. *et al.* Chromosomal deletions and inversions mediated by TALENs and CRISPR/Cas in zebrafish. *Nucleic Acids Res.* **41**, e141 (2013).
26. Dupret, B. *et al.* The histone lysine methyltransferase Ezh2 is required for maintenance of the intestine integrity and for caudal fin regeneration in zebrafish. *Biochim. Biophys. Acta.* **1860**, 1079–1093 (2017).
27. O'Carroll, D. *et al.* The polycomb-group gene Ezh2 is required for early mouse development. *Mol. Cell. Biol.* **21**, 4330–4336 (2001).
28. San, B. *et al.* Normal formation of a vertebrate body plan and loss of tissue maintenance in the absence of ezh2. *Sci. Rep.* **6**, 24658 (2016).
29. Völkel, P. & Angrand, P. O. The control of histone lysine methylation in epigenetic regulation. *Biochimie.* **89**, 1–20 (2007).
30. Antonysamy, S. *et al.* Structural context of disease-associated mutations and putative mechanism of autoinhibition revealed by X-ray crystallographic analysis of the EZH2-SET domain. *PLoS One.* **8**, e84147 (2013).
31. Wu, H. *et al.* Structure of the catalytic domain of EZH2 reveals conformational plasticity in cofactor and substrate binding sites and explains oncogenic mutations. *PLoS One.* **8**, e83737 (2013).
32. Christoffels, A. *et al.* Fugu genome analysis provides evidence for a whole-genome duplication early during the evolution of rayfined fishes. *Mol. Biol. Evol.* **21**, 1146–1151 (2004).
33. Jaillon, O. *et al.* Genome duplication in the teleost fish *Tetraodon nigroviridis* reveals the early vertebrate proto-karyotype. *Nature.* **431**, 946–957 (2004).
34. Bian, C. *et al.* The Asian arowana (*Scleropages formosus*) genome provides new insights into the evolution of an early lineage of teleosts. *Sci. Rep.* **6**, 24501 (2016).
35. Larhammar, D. & Risinger, C. Molecular genetic aspects of tetraploidy in the common carp *Cyprinus carpio*. *Mol. Phylogenet. Evol.* **3**, 59–68 (1994).
36. Xu, P. *et al.* Genome sequence and genetic diversity of the common carp, *Cyprinus carpio*. *Nat. Genet.* **46**, 1212–1219 (2014).
37. Berthelot, C. *et al.* The rainbow trout genome provides novel insights into evolution after whole-genome duplication in vertebrates. *Nat. Commun.* **5**, 3657 (2014).
38. Braasch, I. *et al.* The spotted gar genome illuminates vertebrate evolution and facilitates human-teleost comparisons. *Nat. Genet.* **48**, 427–437 (2016).
39. Amemiya, C. T. *et al.* The African coelacanth genome provides insights into tetrapod evolution. *Nature.* **496**, 311–316 (2013).
40. Read, T. D. *et al.* Draft sequencing and assembly of the genome of the world's largest fish, the whale shark: *Rhincodon typus* Smith 1828. *BMC Genomics.* **18**, 532 (2017).
41. Wang, Q. *et al.* Community annotation and bioinformatics workforce development in concert—Little Skate Genome Annotation Workshops and Jamborees. *Database (Oxford).* **2012**, bar064 (2012).
42. Wyffels, J. *et al.* SkateBase, an elasmobranch genome project and collection of molecular resources for chondrichthyan fishes. *F1000Res.* **3**, 191 (2014).
43. Venkatesh, B. *et al.* Elephant shark genome provides unique insights into gnathostome evolution. *Nature.* **505**, 174–179 (2014).
44. Pelegri, F. Maternal factors in zebrafish development. *Dev. Dyn.* **228**, 535–554 (2003).
45. Dosch, R. *et al.* Maternal control of vertebrate development before the midblastula transition: mutants from the zebrafish I. *Dev. Cell.* **6**, 771–680 (2004).
46. Cermak, T. *et al.* Efficient design and assembly of custom TALEN and other TAL effector-based constructs for DNA targeting. *Nucleic Acids Res.* **39**, e82 (2011).
47. Cao, R., Tsukada, Y. & Zhang, Y. Role of Bmi-1 and Ring1A in H2A ubiquitylation and Hox gene silencing. *Mol. Cell.* **20**, 845–854 (2005).
48. Le Faou, P., Völkel, P. & Angrand, P. O. The zebrafish genes encoding the Polycomb repressive complex (PRC) 1. *Gene.* **475**, 10–21 (2011).
49. Stephens, S. G. Possible significances of duplication in evolution. *Adv. Genet.* **4**, 247–265 (1951).
50. Chrispijn, N. D., Andralojc, K. M., Castenmiller, C. & Kamminga, L. M. Gene expression profile of a selection of Polycomb Group genes during zebrafish embryonic and germ line development. *PLoS One.* **13**, e0200316 (2018).
51. Sun, X. J. *et al.* Genome-wide survey and developmental expression mapping of zebrafish SET domain-containing genes. *PLoS One.* **3**, e1499 (2008).
52. White, R. J. *et al.* A high-resolution mRNA expression time course of embryonic development in zebrafish. *Elife.* **6**, e30860 (2017).
53. Guryev, V. *et al.* Genetic variation in the zebrafish. *Genome Res.* **16**, 491–497 (2006).
54. Ai, S. *et al.* Divergent Requirements for EZH1 in Heart Development Versus Regeneration. *Circ. Res.* **121**, 106–112 (2017).
55. Ezhkova, E. *et al.* EZH1 and EZH2 cogovern histone H3K27 trimethylation and are essential for hair follicle homeostasis and wound repair. *Genes Dev.* **25**, 485–498 (2011).
56. Lui, J. C. *et al.* EZH1 and EZH2 promote skeletal growth by repressing inhibitors of chondrocyte proliferation and hypertrophy. *Nat. Commun.* **7**, 13685 (2016).
57. Bae, W. K. *et al.* The methyltransferases enhancer of zeste homolog (EZH) 1 and EZH2 control hepatocyte homeostasis and regeneration. *FASEB J.* **29**, 1653–1662 (2015).
58. Mochizuki-Kashio, M. *et al.* Ezh2 loss in hematopoietic stem cells predisposes mice to develop heterogeneous malignancies in an Ezh1-dependent manner. *Blood.* **126**, 1172–1183 (2015).
59. Kimmel, C. B., Ballard, W. W., Kimmel, S. R., Ullmann, B. & Schilling, T. F. Stages of embryonic development of the zebrafish. *Dev. Dyn.* **203**, 253–310 (1995).

60. Doyle, E. L. *et al.* TAL Effector-Nucleotide Targeter (TALE-NT) 2.0: tools for TAL effector design and target prediction. *Nucleic Acids Res.* **40**, W117–122 (2012).
61. Dupret, B., Völkel, P., Le Bourhis, X. & Angrand, P. O. The Polycomb Group Protein Pcgf1 is dispensable in zebrafish but involved in early growth and aging. *PLoS One.* **11**, e0158700 (2016).
62. Dupret, B., Völkel, P., Follet, P., Le Bourhis, X. & Angrand, P. O. Combining genotypic and phenotypic analyses on single mutant zebrafish larvae. *MethodsX.* **5**, 244–256 (2018).

### Acknowledgements

We are grateful to Shaghayegh Hasanpour for helpful discussions and her drawings of fish in Fig. 2. This work was supported by the Inserm, the University of Lille and the GIP Cancéropôle Nord-Ouest. BD was supported by a fellowship from the Région Hauts-de-France, Nord-Pas de Calais, Picardie and the University of Lille.

### Author Contributions

P.O.A. conceived and supervised the work. L.R. and P.O.A. performed the phylogenetic analyses. P.V. and P.O.A. generated the *ezh1<sup>ul3/ul3</sup>* zebrafish line using the TALEN methodology. P.V., A.B., L.R., A.C. and B.D. performed the *in situ* hybridizations and the phenotypic characterizations. P.V., L.R., B.D., X.L.B. and P.O.A. analyzed the data. X.L.B. and P.O.A. provided critical resources. P.O.A. wrote the original draft with the help of P.V. and X.L.B. P.V., L.R. and P.O.A. revised and edited the final manuscript. All authors read and approved the final manuscript.

### 2.3 Conclusion

L'activité méthyltransférase du complexe PRC2 est portée par le seul gène *Enhancer of zeste* ou *E(z)* chez la drosophile. Deux gènes orthologues sont retrouvés dans le génome des mammifères. La comparaison des séquences protéiques de EZH1 et EZH2 avec *E(z)* montrent que l'homologie de séquence est plus importante entre EZH2 et *E(z)* (Laible *et al.*, 1997). L'absence de *Ezh1* chez les agnathes confirme que EZH2 est l'orthologue le plus proche de *E(z)*, tandis que *Ezh1* est apparu plus tardivement dans l'évolution par duplication de *Ezh2*. La comparaison de séquences protéiques de *Ezh2* et *Ezh1* chez les téléostéens permet de retracer l'histoire évolutive de ces deux protéines. La conservation des deux protéines montrent qu'elles ont toutes deux une fonction propre. Néanmoins, l'absence de *Ezh2* chez le poisson globe suggère une redondance fonctionnelle de ces deux protéines. L'expression de *ezh1* a été analysée par hybridation *in situ* et RT-PCR. Le transcrit *ezh1* est déposé dans les œufs et est observable dans les stades les plus précoces du développement. Ces résultats s'opposent à plusieurs études publiées (Sun *et al.*, 2008 ; Chrispijn *et al.*, 2018). Au cours du développement, le profil d'expression de *ezh1* ressemble fortement à ce qu'il a été décrit pour *ezh2* (Dupret *et al.*, 2017), ce qui pourrait suggérer une redondance fonctionnelle. Néanmoins, la quantification des transcrits *ezh1* et *ezh2* évolue inversement au cours du développement. Le transcrit *ezh2* est fortement exprimé durant les phases de développement précoce. L'expression de *ezh1* est très réduite dans les stades précoces et augmente au cours du développement. Cette observation semble confirmer le rôle de *Ezh1* dans les cellules différenciées. Pour avoir une meilleure compréhension des fonctions de *Ezh1*, les TALENs ont été utilisés afin de générer un mutant perte de fonction de *Ezh1*. La protéine *Ezh1* n'est pas indispensable au développement embryonnaire. Néanmoins, la perte de *Ezh1* augmente l'expression de *Ezh2* et diminue la survie des mutants *ezh2<sup>ul2/ul2</sup>*. L'ensemble de ces résultats confirment les conclusions de l'équipe de Margueron (2008), la redondance fonctionnelle de *Ezh1* et *Ezh2* ne semblent concerner qu'une faible proportion de cibles génomiques.

### 3. Eed : Loss of Polycomb Repressive Complex 2 Function Alters Digestive Organ Homeostasis and Neuronal Differentiation in Zebrafish

#### 3.1 Contexte scientifique

La répression Polycomb exerce un contrôle transcriptionnel par deux complexes protéiques, PRC1 et PRC2. Le complexe PRC2 dépose la marque H3K27me3 qui est alors reconnue par les protéines CBX du complexe PRC1. Le complexe PRC1 catalyse l'ajout d'un groupement ubiquitine sur la lysine 119 de l'histone H2A. L'activité des deux complexes est liée autant dans leur fonction de répression transcriptionnelle que par leur recrutement. Néanmoins, leur composition peut varier et ainsi moduler leur fonction (Gil & O'Loughlen, 2014 ; Schuettengruber *et al.*, 2017). Le complexe PRC2 est composé de plusieurs sous-unités protéiques. Les protéines EZH1 ou EZH2 portent l'activité enzymatique du complexe et ajoutent des groupements méthyles sur la lysine 27 de l'histone H3. D'autres acteurs protéiques sont présents dans le complexe PRC2 et sont indispensables à son activité. Les protéines SUZ12, EED et RBBP4/7 représentent avec EZH1/2 la formation minimale nécessaire à l'activité du PRC2. La perte de fonction de SUZ12 chez la souris conduit à un arrêt du développement embryonnaire précoce en raison de la perte de la marque H3K27me3 (Pasini *et al.*, 2004). De la même façon, les embryons murins portant une mutation de *Eed* voient leur développement s'arrêter pendant la gastrulation pour les mêmes raisons. Ces protéines sont donc essentielles à l'activité méthyltransférase de PRC2 (Faust *et al.*, 1995 ; 1998). Le complexe PRC2 est conservée chez le poisson zèbre. Le laboratoire a préalablement étudié le complexe PRC2 via la perte de fonction de *ezh2*. Les embryons mutants meurent après la mise en place correcte du plan du corps autour de 12 jpf (Dupret *et al.*, 2017). Les embryons *ezh1*<sup>-/-</sup> et *ezh2*<sup>-/-</sup> ont une survie réduite, ce qui suggère un effet de compensation de Ezh1 en cas de perte de Ezh2. Néanmoins, il a été relevé dans la littérature que la protéine Ezh2 pouvait avoir une activité indépendamment du complexe PRC2 (Völkel *et al.*, 2015), et même activer la transcription de certains gènes (Xu *et al.*, 2012). Ainsi, la perte de fonction de Ezh2 chez le poisson zèbre ne permet pas d'étudier uniquement sa fonction dans le complexe PRC2. Afin d'étudier la fonction du PRC2 et de la marque H3K27me3 dans le développement précoce du poisson zèbre, nous avons adopté la stratégie de cibler spécifiquement une autre protéine. Le gène *suz12* chez le poisson zèbre est conservé en deux orthologues, *suz12a* et *suz12b*, et nécessite la génération d'une double mutation. De ce fait, les TALENs ont été utilisés pour cibler spécifiquement le gène *eed*. De la même façon que les mutants *ezh2*, les embryons mutants *eed* présentent une mortalité tardive en



comparaison du modèle murin et laisse l'opportunité d'étudier le rôle du complexe PRC2 dans le développement tardif, la spécification et la maintenance des tissus.

## 3.2 Article

### **Loss of Polycomb Repressive Complex 2 Function Alters Digestive Organ Homeostasis and Neuronal Differentiation in Zebrafish**

**Ludivine Raby** <sup>2,t</sup>, **Pamela Völkel** <sup>1,t</sup>, **Shaghayegh Hasanpour** <sup>2</sup>, **Julien Cicero** <sup>1,3</sup>, **Robert-Alain Toillon** <sup>1</sup>, **Eric Adriaenssens** <sup>1</sup>, **Isabelle Van Seuning** <sup>1</sup>, **Xuefen Le Bourhis** <sup>1</sup> and **Pierre-Olivier Angrand** <sup>1,\*</sup>

<sup>1</sup> Univ. Lille, CNRS, Inserm, CHU Lille, UMR 9020-U 1277 – CANTHER – Cancer Heterogeneity Plasticity and Resistance to Therapies, F-59000 Lille, France; ludivine.raby.etu@univ-lille.fr (L.R.); pamela.voelkel@univ-lille.fr (P.V.); julien.cicero.etu@univ-lille.fr (J.C.); robert-alain.toillon@univ-lille.fr (R.-A.T.); eric.adriaenssens@univ-lille.fr (E.A.); isabelle.vanseuning@inserm.fr (I.V.S.); xuefen.le-bourhis@univ-lille.fr (X.L.B.)

<sup>2</sup> Department of Fisheries and Animal Sciences, Faculty of Natural Resources, University of Tehran, Karaj 31587-77871, Iran; shaghayegh.hasanpour88@yahoo.com

<sup>3</sup> Univ. Artois, UR 2465, Laboratoire de la Barrière Hémato-Encéphalique (LBHE), F-62300 Lens, France

Received: 12 October 2021, Accepted: 9 November 2021, Published: 12 November 2021

#### **Abstract**

Polycomb repressive complex 2 (PRC2) mediates histone H3K27me3 methylation and the stable transcriptional repression of a number of gene expression programs involved in the control of cellular identity during development and differentiation. Here, we report on the generation and on the characterization of a zebrafish line harboring a null allele of *eed*, a gene coding for an essential component of the PRC2. Homozygous *eed*-deficient mutants present a normal body plan development but display strong defects at the level of the digestive organs, such as reduced size of the pancreas, hepatic steatosis, and a loss of the intestinal structures, to die finally at around 10–12 days post fertilization. In addition, we found that PRC2 loss of function impairs neuronal differentiation in very specific and discrete areas of the brain and increases larval activity in locomotor assays. Our work highlights that zebrafish is a suited model to study human pathologies associated with PRC2 loss of function and H3K27me3 decrease.

#### **Introduction**

Polycomb group (PcG) proteins are epigenetic repressors involved in a variety of biological processes, including stem cell renewal and control of cell identity during differentiation and development [1–3]. PcG proteins associate into two main histone-modifying protein complexes named polycomb repressive complexes 1 and 2 (PRC1 and PRC2) [4–6]. According to the current canonical model, PRC2 mono-, di-, and trimethylates lysine 27 of histone H3 (H3K27me1/2/3) [7–9] leading to the recruitment of PRC1 that in turn monoubiquitinylates lysine 119 of histone H2A (H2AK119ub) [10,11]. These histone modifications induce chromatin compaction and transcriptional silencing at numerous PRC2 target gene expression programs involved in cell fate, cell differentiation, lineage specification, or development [12–19]. Thus, PRC2 that initiates polycomb repression appears central to epigenetic regulation. The PRC2 core complex consists of four stoichiometric factors EZH2 or EZH1, SUZ12, EED, and RBBP4/7. The first three proteins, EZH1/2, SUZ12, and EED, form the minimal complex possessing the histone methyltransferase activity. EZH2 and EZH1 are two mutually exclusive paralogs harboring the catalytic activity, but EZH2 appears to be the main catalytic subunit of PRC2 during development and remains more efficient than EZH1 at methylating

---

H3K27 [20]. However, the catalytic activity of EZH1/2 requires the physical association with EED and SUZ12 as their absence inhibits PRC2 activity. RBBP4/7 is required for PRC2 binding to nucleosomes and enhances its methyltransferase activity. In addition, a number of facultative subunits have been reported to bind to the PRC2 core complex at a substoichiometric level defining at least two distinct multimeric complexes, PRC2.1 and PRC2.2. PRC2.1 contains one of the PCL proteins (PHF1, MTF2, or PHF19), together with EPOP or PALI1, whereas PRC2.2 is associated with AEBP2 and JARID2. These additional proteins are involved in PRC2 recruitment to chromatin, stabilization of the protein complex, and/or modulation of its activity [5,21–23]. In this landscape, SUZ12 and EED are the only two components associated with all PRC2 alternate protein complexes and are absolutely required for deposition and spreading of the H3K27me2/3 epigenetic marks into the epigenome. Indeed, these components are essential for early mouse development, and mice lacking *Suz12* or *Eed* functions are not viable and die during early postimplantation stages [24,25]. In humans, EED mutations and variants have been identified in different pathologies, including cancers [26–29] as well as abnormal developmental syndromes [30–34].

Here, we describe the use of the zebrafish model to investigate the function of *Eed*. Zebrafish (*Danio rerio*) is a widely used vertebrate model for studying gene functions in development, morphogenesis as well as in physiopathological conditions. Furthermore, this organism is particularly suited to study the role of the PRC2 complex and H3K27me3 methylation. Indeed, it has been previously shown that in contrast to mice, loss of function of several core subunits of the PRC2 complex, including *Ezh1* [35,36], *Ezh2* [37,38], and *Suz12a/b* [39], does not lead to zebrafish lethality at gastrulation. Instead, *ezh2* and *suz12a; suz12b* homozygous mutants die at around 12 days post fertilization (dpf), after body plan formation and tissue specification [37–39], whereas *ezh1* homozygous mutants are viable and fertile [35,36]. This absence of embryonic lethality in zebrafish allows the study of the role of H3K27me3 methylation in tissue development and homeostasis without the requirement of conditional knockout strategies. In this study, we generated and characterized null alleles of *eed* to investigate the role of PRC2 function and H3K27me2/3 methylation in zebrafish development and organogenesis. Homozygous *eed*<sup>-/-</sup> mutants die at around 10–12 dpf with strong defects at the level of the digestive organs, including reduced size of the pancreas, hepatic steatosis, and loss of the intestinal structures. However, this relatively late lethality allowed us to focus on the expression of markers of the developmental brain. We found that PRC2 loss of function alters neuronal differentiation in very specific and discrete areas of the brain. Furthermore, locomotor activity assays highlighted the role of PRC2 in the zebrafish larval behavior. Altogether, our observations showed that this *eed*-deficient zebrafish line could constitute a valuable model to study the impact of H3K27me3 decrease and PRC2 loss of function in development as well as in several human pathologies.

### Materials and Methods

#### *Zebrafish Maintenance, Embryo Preparation, and Animal Ethics Statements*

Zebrafish from the TU strain and Tg(actb2:GFP-Hsa.UTRN)<sup>e116</sup> transgenics [40] are maintained at 27.5 °C in a 14/10 h light/dark cycle. The evening before spawning, males and females were separated into individual breeding tanks (Tecniplast, Decines-Charpieu, France). Spontaneous spawning occurred the following morning when the plastic divider was removed. Embryos were collected and staged according to Kimmel *et al.* [41]. The chorions were removed from embryos using 1% pronase (Sigma, St. Louis, MO, USA) for 1 min. Zebrafish embryos were fixed overnight in 4% paraformaldehyde in PBS (phosphatebuffered saline), dehydrated gradually to 100% methanol, and stored at -20 °C.

Zebrafish were maintained in compliance with the French and European Union guidelines for the handling of laboratory animals (Directive 2010/63/EU of the European Parliament and of the Council of 22 September 2010 on the protection of animals used for scientific purposes). The experimental procedures carried out on zebrafish were reviewed and approved by the local Ethics Committee, CEEA 75 Nord Pas-de-Calais, and the French Ministry of Higher Education and Research (APAFiS approval number 135272018011722529804\_v3).

#### *TALEN Design and Assembly*

TALENs targeting *eed* were designed and assembled as previously described [42]. The *eed* TALEN target site was selected using the online TAL Effector-Nucleotide Targeter tool (<https://tale-nt.cac.cornell.edu/>), last

accessed on 11 October 2021 [43]) in exon 3 (ENSDART00000162928.1) with the presence of a restriction site (BseNI) within the spacer sequence for screening and genotyping purposes.

Eed-specific TALEN constructs were engineered using the TALEN Golden Gate assembly system described by Cermak *et al.* [44] using the expression backbones, pCS2TAL3DD and pCS2TAL3RR [45]. Plasmids for Golden Gate Cloning were obtained from Addgene.

### *mRNA Injection into Zebrafish Embryos*

Capped mRNAs were synthesized using SP6 mMACHINE kit (Ambion) from linearized plasmid templates. mRNAs (50–100 pg) were injected into 1-cell zebrafish embryos using a FemoJet microinjector (Eppendorf, Matesson, France).

### *Genotype Analyses*

For genotyping, embryos, larvae or caudal biopsies of larvae were incubated in 10  $\mu$ L PCR extraction buffer (10 mM Tris-HCl pH 8.0, 2 mM EDTA, 0.2% Triton X-100, 100  $\mu$ g/mL proteinase K) and placed at 50 °C for 4 h prior to proteinase K inactivation at 95 °C for 5 min. Genotype analysis was performed by PCR on 2.5  $\mu$ L of samples using the primer set TAL\_eed\_5<sup>0</sup> (CAGCCTAAAGGTGAGGCCGC) and TAL\_eed\_3<sup>a</sup> (CGATCTCGACCCGGAACAC) followed by PCR product digestion with the BseNI restriction enzyme. An alternate genotyping strategy consists of the PCR amplification of genomic DNAs with the primer set TAL\_eed5\_new GTGTGCAGGAGGATCATGGTCAG) and TAL\_eed3\_new (GTCCGTCACACTCTGTTGCTG) followed by a 2% agarose gel electrophoresis without enzymatic digestion.

To genotype paraformaldehyde-fixed embryos and larvae, DNA was extracted using sodium hydroxide and Tris [46]. Single embryos or larvae were placed into microcentrifuge tubes containing 20  $\mu$ L 50 mM NaOH and heated 20 min at 95 °C. The tubes were then cooled to 4 °C, and 2  $\mu$ L of 1 M Tris-HCl, pH 7.4, was added to neutralize the basic solution. Genotype analysis was performed on 2.5  $\mu$ L of samples by PCR-BseNI digestion or PCR without digestion, as described above.

Sequence determination (GATC-Eurofin, Ebersberg, Germany) was performed after cloning of the PCR products into pCR-XL-TOPO (Invitrogen-ThermoFisher, Waltham, MA, USA) according to the manufacturer's instructions.

### *Kaplan–Meier Analysis*

For the establishment of survival curves, the tip of the caudal fin of embryos from an *eed*<sup>+/-</sup> incross was transected within the pigment gap distal to the circulating blood for genotyping purposes [46]. The embryos were then pooled according to their genotypes, placed into separate 1-L tanks, and incubated at 28 °C. Larvae were fed from 5 dpf on, 3 times per day, with early-stage zebrafish nutrition (Gemma Micro ZF75, Planktovie, France). The larvae were regularly examined directly in the tanks using a Leica EZ4 stereomicroscope with no or minimal manipulation during a period of 18 days after the spawn. Larvae were declared dead when heartbeats could not be detected. Dead larvae were immediately removed from the tanks.

### *Oil Red-O Staining*

Zebrafish larvae were fixed overnight in 4% paraformaldehyde and pigments were bleached 2 h in 1% KOH, 3% H<sub>2</sub>O<sub>2</sub>. Larvae were stained with oil red-O (Hamiya Biomedical Company, Tukwila, WA, USA) for 15 min. Stained larvae were washed with PBS, stored in 70% glycerol, and imaged on a bright-field Leica MZ125 stereomicroscope equipped with a Leica DFC295 digital camera. After imaging, DNA was extracted using sodium hydroxide and Tris and genotyped.

### *Smurf Assay*

Smurf assays were conducted according to Dambroise *et al.* [47] by placing zebrafish larvae in 50 mL Falcon tubes containing 2.5% blue #1 (erionoglucine disodium salt, Sigma) for 30 min at room temperature. They were then rinsed under aquarium water until no more blue coloration could be found in the eluate, anesthetized in 0.016% tricaine (MS-222;

3-amino benzoic acid ethyl ester, Sigma) and imaged on a bright-field stereomicroscope (Leica MZ125) equipped with a Leica DFC295 digital camera. After imaging, DNA was extracted for genotyping analyses.

### *Histological Analysis*

For histological analyses, tail biopsies were used for genotyping before paraformaldehyde fixation of the larvae. Larvae were embedded in paraffin and cut into 5  $\mu\text{m}$ -thick sections. These were mounted on sylanated glass slides, deparaffinated, rehydrated, and stained with hematoxylin and eosin for histological analysis as previously described [46].

### *Biometric Measurements*

Standard length and eye diameter were measured according to Parichy *et al.* [48] on anesthetized larvae using ImageJ. After imaging and measurement, the larvae were genotyped.

Measurements of the surface of the pancreas labeled by in situ hybridization with a *prss1* probe were performed by equal intensity pixel counting using ImageJ.

Quantification of macrovesicular in the liver was performed on at least 3 consecutive hematoxylin and eosin-stained slides from 4 wild-type and 4 *eed*<sup>-/-</sup> mutant larvae using ImageJ. Each image was anonymized and transformed into an 8-bit grayscale picture before counting the hepatic nuclei and the number of vesicles.

For all biometric measurements, the software GraphPad Prism was used to build the graphs and to perform the statistical analyses.

### *Confocal Microscopy*

For confocal imaging, tail biopsies were transected for genotyping before paraformaldehyde fixation of the larvae. After fixation in 4% paraformaldehyde, zebrafish larvae were maintained on glass-bottom dishes in 1.5% agarose. Fluorescent images were captured with an LSM 880 microscope and an AiryScan detector (ZEISS, Oberkoden, Germany). A Z-stack slice was made every 5  $\mu\text{m}$ .

### *Whole-Mount In Situ Hybridization*

Antisense-RNA probes were generated by RT-PCR from total mRNA extracted from zebrafish larvae at 5 dpf using the RNeasy Mini Kit (Qiagen, Courtaboeuf, France), following the manufacturer's protocol. After reverse transcription (Superscript III, Invitrogen), cDNAs were amplified by PCR using the probe-specific primers, coupled to the T7 sequence for forward primers and the SP6 sequence for reverse primers. DIG-labeled antisense-RNA probes were synthesized using the DIG RNA Labeling Kit (SP6) (Roche-Sigma Aldrich Chimie S.a.r.l., Saint-Quentin-Follavie, France), following the manufacturer's instructions. The primers used for probe generation were:

ISHa\_eeF: TAATACGACTCACTATAGGG GCTGTGGCGGCTGGACTCTG

ISHa\_eeR: GATTTAGGTGACACTATAG CGCCAGATGGAGGCGTCGT

ISH\_fabp2\_F: TAATACGACTCACTATAGGG ATGACCTTCAACGGGACCTGG

ISH\_fabp2\_R: GATTTAGGTGACACTATAG AGCCCTCTTGAATCCTCTTGGC

ISH\_ace2\_F: TAATACGACTCACTATAGGGCTGTTGGAGAGATCATGTCGCTTTCT

ISH\_ace2\_R: GATTTAGGTGACACTATAGTGTCTTCTCAAAGGCTTTGTTCACT

ISHa\_fabp10a\_F: TAATACGACTCACTATAGGG AAAGCATGGCCTCAGCGGGA ISHa\_fabp10a\_R:  
GATTTAGGTGACACTATAG GTCAGCGTCTCCACCATTTCTCCA

ISH\_prss1\_F: TAATACGACTCACTATAGGG TGCTCACTGCTACAAGTCCCGT

ISH\_prss1\_R: GATTTAGGTGACACTATAG CCGAGCTTAGTTGGAGTTCATGGT

ISH\_phox2bb\_F: TAATACGACTCACTATAGGG GGGCCTAACCCGAACCCTACCTC

ISH\_phox2bb\_R: GATTTAGGTGACACTATAG GAGCGCACATCGCAGTCTATCGG

ISH\_pcna\_F: TAATACGACTCACTATAGGG GGCAACATCAAGCTCTCACA

ISH\_pcna\_R: GATTTAGGTGACACTATAG AAATCCCACAGATGACAGGC

ISH\_ccna2\_F: TAATACGACTCACTATAGGGGGAAGGATGTCAACACAAGGAAG

ISH\_ccna2\_R: GATTTAGGTGACACTATAGGAGAGAAGTGTCAAGCACCAGATG

ISH\_ccnd1\_F: TAATACGACTCACTATAGGGTTTGTCTGCGAAGTGGATACCA  
 ISH\_ccnd1\_R: GATTTAGGTGACACTATAGAACAGTTTGGGCGTGCTGAGT  
 ISH\_sox2\_F: TAATACGACTCACTATAGGG GTCCGAGAGCGAGAAGCGACC  
 ISH\_sox2\_R: GATTTAGGTGACACTATAGGCTGTAGGTGGGCGAGCCGT  
 ISH\_mycn\_F: TAATACGACTCACTATAGGGTGCAGACTCCAAACGGGAGGCA  
 ISH\_mycn\_R: GATTTAGGTGACACTATAGGCTCACTTCGGGCGCTTTGACT  
 ISH\_neurod1\_F: TAATACGACTCACTATAGGGTCGAGACGCTCCGACTAGCCAA  
 ISH\_neurod1\_R: GATTTAGGTGACACTATAGGCGTCGAGCCCCGCGTAAAGA  
 ISH\_mag\_F: TAATACGACTCACTATAGGGCCGTGAGGGTGTTCAGTGTGTGT  
 ISH\_mag\_R: GATTTAGGTGACACTATAGCGTCTCCCGTGCCTTCCTCT

RNA probes for *bmi1a* were synthesized with the DIG RNA Labeling Kit (SP6/T7) (Roche) on 1 µg of linearized cDNA clone MGC:56403 IMAGE:5605189 (imaGenes GmbH, Berlin, Germany).

Whole-mount in situ hybridizations were performed according to Thisse and Thisse [49]. Briefly, the fixed embryos were rehydrated and permeabilized with 10 µg/mL proteinase K for 30 sec (1-cell stage embryos), 10 min (24 hpf embryos), or 30 min (48 hpf embryos and 5 dpf larvae) at room temperature. Ten to 15 embryos from the TU strain of each time point, or about 50 larvae from *eed<sup>+/-</sup>* incrosses at 5 dpf, were hybridized with digoxigeninlabeled antisense-RNA probes at 70 °C. After extensive washing, the probes were detected with anti-digoxigenin-AP Fab fragment (Roche Diagnostics, 1093274, diluted at 1:10,000), followed by staining with BCIP/NBT (5-bromo-4-chloro-3-indolyl-phosphate/nitro blue tetrazolium) alkaline phosphate substrate. The embryos were imaged using a Leica MZ10F stereomicroscope equipped with a Leica DFC295 digital camera.

#### ***Histone Extraction and Western Blot Analysis***

Histone extraction and Western blot analyses were performed from 10 larvae as described previously [46]. Primary antibodies used were mouse anti-H3K27me3 (1:1000; ab6002, Abcam, Paris, France), rabbit anti-H3K27me2 (1:500; ab24684, Abcam), rabbit anti-H3K27me1 (1:500; ab84932, Abcam), rabbit anti-H2AK119ub (1:2000; DC27C4; Cell Signaling), rabbit antiH3K9me3 (1:1000; AB8898; Millipore-Sigma Aldrich S.a.r.l., Saint-Quentin-Follavier, France), rabbit anti-H4K20me3 (1:1000; ab9053; Abcam), rabbit anti-H3K27ac (1:1000; ab4729; Abcam) and rabbit anti-H3 (1:5000; ab1791, Abcam). The secondary antibodies were peroxidase conjugated anti-mouse antibody (1:10,000; 115-035-003, Jackson ImmunoResearch, Ely, UK) and peroxidase conjugated anti-rabbit antibody (1:10,000; 711-035-152, Jackson ImmunoResearch).

#### ***RNA Extraction, RT-PCR and RT-qPCR***

RNA extraction and RT-PCR experiments were conducted as previously described [50]. cDNA was synthesized using Superscript III (18080-044, Invitrogen, Carlsbad, CA, USA) following the manufacturer's instructions. A quantity of 1 µg of total RNAs were used to perform the reverse transcription experiments. Primers used were:

Dr_eed_cDNA_5 <sup>0</sup> : GGAAACGAGATGCCGAACAA	
Dr-eed_cDNA_3 <sup>0</sup> a: CAGCCGGATCTCTCCCTGAG	
Dr_cDNA_ube2a_F: TGACTGTTGACCCACCTTACAG	
Dr_cDNA_ube2a_R: CAAATAAAAGCAAGTAACCCCG	
Dr_beta-actin_cDNA_5 <sup>0</sup> : CGTGACATCAAGGAGAAGCT	Dr_beta-actin_cDNA_3 <sup>0</sup> :
ATCCACATCTGCTGGAAGGT	qPCR_fasn_5 <sup>0</sup> :
CCAGCCATAAGAACGTCAGCCGAGA	qPCR_fasn_3 <sup>0</sup> :
CACCTTCCCGTCACACACCTCGT	qPCR_srebf1_5 <sup>0</sup> :
ACAGCGCGGCTAATGGCAGG	qPCR_srebf1_3 <sup>0</sup> : TGCCCAGGAGCCGACAGGAA
qPCR_pparg_5 <sup>0</sup> : GCACAGGCGCTTCAGTGTTCAG	qPCR_pparg_3 <sup>0</sup> :
CCAGCGAGTGCGTGTCTCGTCC	qPCR_nr1h3_5 <sup>0</sup> :
ACCCAGCGATTGACAGCATCACCT	qPCR_nr1h3_3 <sup>0</sup> :
CCTGCACGTTTGGTCCGTCTGCT	

PCR reactions were performed as follow: 95 °C 4 min, (95 °C 45 s, 65 °C 45 s, 72 °C 1 min) 35 cycles, 72 °C 10 min. The quantitative qPCR reaction was performed in triplicate using a QuantStudio 3 Real-Time System (Thermo Fisher Scientific, Waltham, MA, USA) using SYBR Green Supermix (Bio-Rad, Marne-la-Coquette, France). Relative mRNA expression of each gene was normalized to *ube2a* levels.

### ***Behavioral Studies***

Behavioral studies were performed on 5 dpf larvae from *eed*<sup>+/-</sup> incrosses in plates that were handled minimally before placement in a ZebraBox chamber (ViewPoint Life Sciences, Lyon, France) equipped with an infrared light-emitting floor and a top-mounted infrared camera allowing video recording of the whole plate under both light and dark conditions. Larval behavior measurements were achieved using the ZebraLab software with a detection threshold set at 35 and an xmin set at 3 (ViewPoint Life Sciences, Lyon, France). After recording, the larvae were euthanized, and their genotype was determined.

The locomotor activity assays were conducted in 48-well plates. The protocol consists of a 10 min initial acclimating period in the dark, followed by six alternating 10 min light and dark phases.

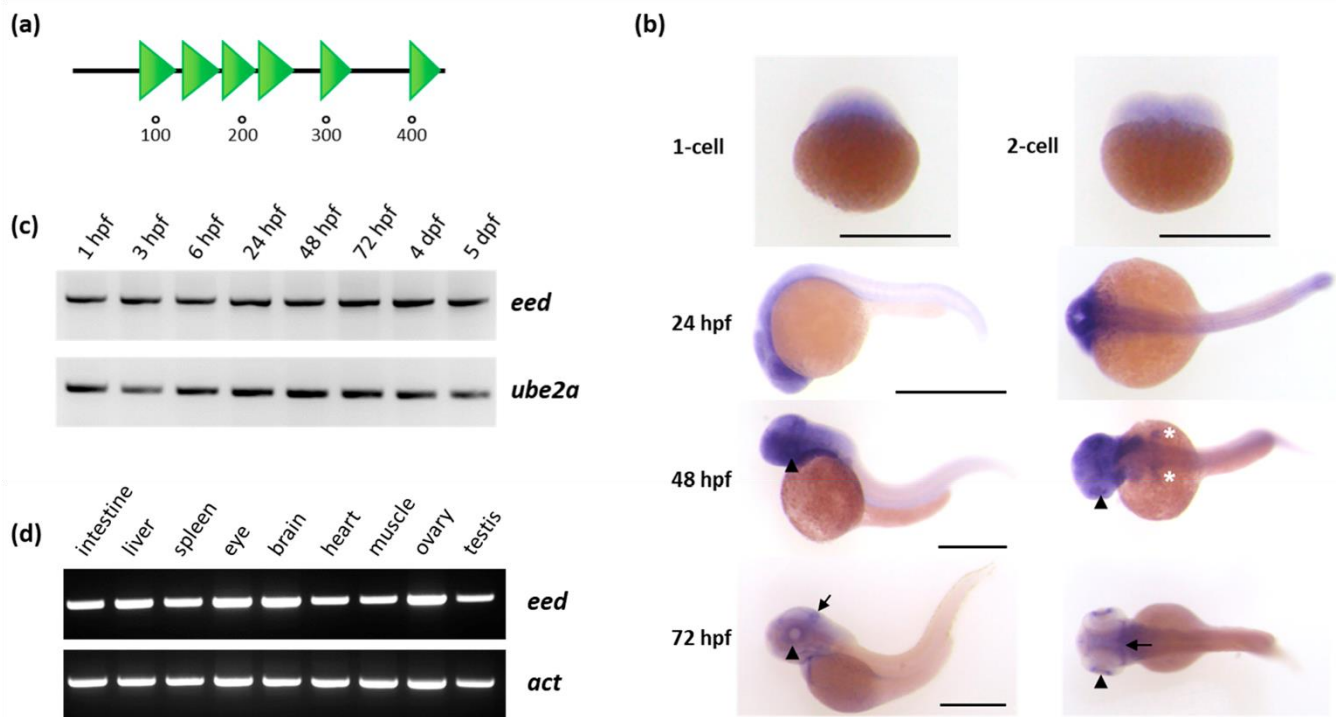
Thigmotaxis assays were conducted in 24-well plates. The timeline of the thigmotaxis assay protocol consists of a 6 min acclimatization step in the light followed by a 4 min dark challenge period. The arena was divided into two areas, inner and outer, of the equivalent surface. The diameter of the inner area corresponds to the total diameter of the well divided<sup>√</sup> by 2. Thigmotaxis was presented as the percentage of the total distance moved or the total time spent by each larva into the outer zone during the challenge period.

## **Results**

### ***eed Expression during Zebrafish Development***

Human EED protein is encoded by 12 exons containing gene located on chromosome 11, upstream of the *HIKESHI* (*heat shock protein nuclear import factor hikiishi*) gene. The protein is composed of 441 amino acids and contains six WD40 domains adopting a  $\beta$ -propeller architecture forming multiple van der Waals interactions and hydrogen bonds with a domain located at the N-terminus of EZH2 [51]. EED is conserved across vertebrate species, both at the protein and genomic levels. In zebrafish, the *eed* gene is located on chromosome 1. It is organized in 12 exons positioned at the 5' of the *hikiishi* gene, thus defining an evolutionarily conserved syntenic block. The gene also codes for a 441-amino-acids-long protein sharing 91.4% identity with its human ortholog and containing six WD40 domains as well (Figure 1a).





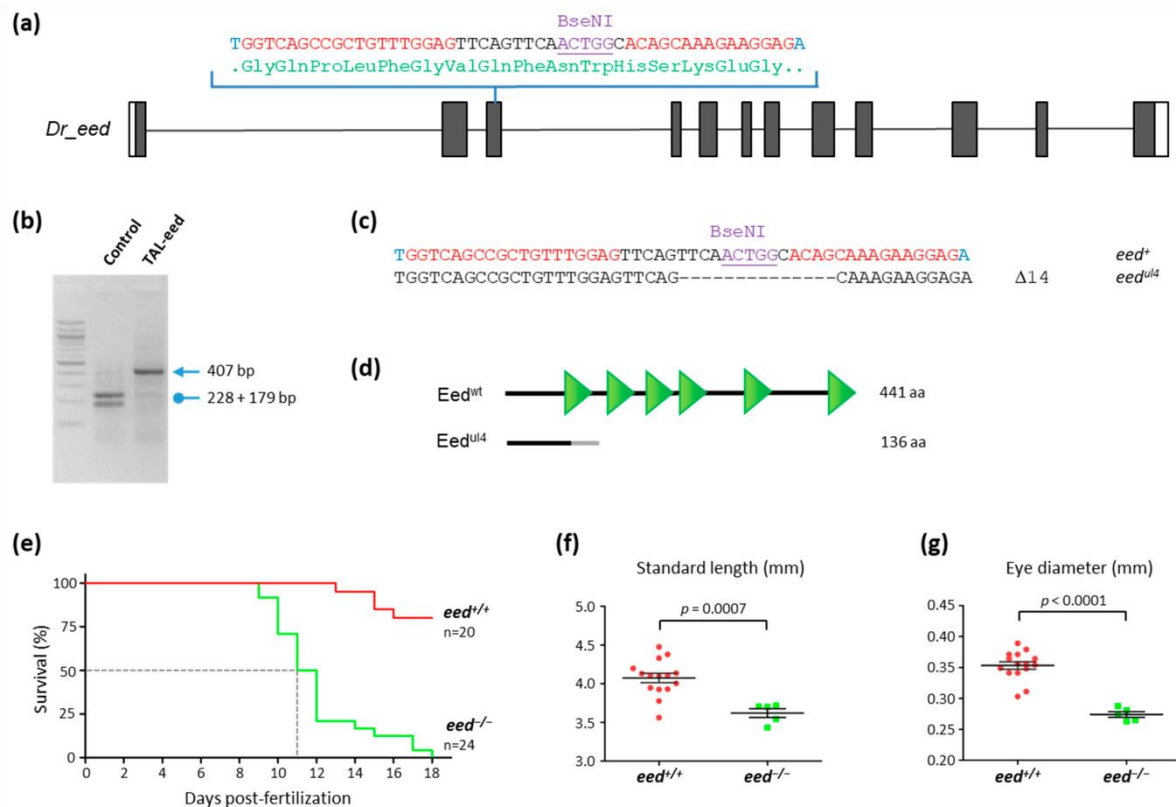
**Figure 1.** Organization of the Eed protein and *eed* mRNA expression in zebrafish: (a) schematic representation of the Eed protein. The zebrafish protein, as its human ortholog, is composed of 441 amino acids and contains six WD40 domains (SMART: SM000320) shown as green triangles; (b) whole-mount RNA in situ hybridization showing maternally provided *eed* transcripts at the 1-cell and 2-cell stage, and zygotic *eed* mRNA distribution at 24, 48 and 72 hpf. At these later stages, a lateral view is at the left and a dorsal view at the right. The arrowheads show the retina, the white asterisks the pectoral fin buds, and the arrow identify the midbrain-hindbrain boundary. Scale bar is 500  $\mu$ m; (c) RT-PCR experiment showing the detection of *eed* transcripts at 1 hpf, 3 hpf, 6 hpf, 24 hpf, 48 hpf, 72 hpf, 4 dpf, and 5 dpf. *Ube2a* is used as a control; (d) RT-PCR experiment showing *eed* mRNAs in adult zebrafish tissues. Beta-actin (*act*) is used as a control. (See Supplementary Materials).

In zebrafish, zygotic transcription is activated at about cell cycle 10–13, around 3.5 h post fertilization (hpf). Before this mid-blastula transition (MBT) stage, all developmental events depend on the maternally provided gene products [52,53]. Using in situ hybridization at 1-cell and 2-cell stages (Figure 1b) and RT-PCR at 1 and 3 hpf (Figure 1c), we found in agreement with previous studies [54,55] that *eed* mRNA is maternally loaded into the zebrafish embryo. After MBT, during tissue specification, *eed* zygotic expression is detected by RT-PCR (Figure 1c) and appears enriched in the anterior region of the embryo and in highly proliferative tissues such as the retina, the midbrain-hindbrain boundary, the pharyngeal arches, the pectoral fin buds or the intestine as shown by in situ hybridization at 24, 48 and 72 hpf (Figure 1b). Finally, in adult zebrafish, RT-PCR experiments showed that *eed* is ubiquitously expressed in all the tissues we tested (Figure 1d).

### TALEN-Mediated Inactivation of *eed* in Zebrafish

To generate *eed* mutations in zebrafish, we designed TALENs targeting the exon 3 within the sequence coding for the first WD40 repeat (Figure 2a). The targeted site was chosen to cover a BseNI restriction site used to screen for TALEN-induced mutations and for genotyping purposes (Figure 2b). Following TALEN mRNA injection in zebrafish embryos from the TU strain at the 1-cell stage, we isolated a line carrying an allele, *eed<sup>ul4</sup>*, with a deletion of 14 base pairs (Figure 2c). This mutation produces a frameshift leading to the synthesis of a predicted protein of 136 amino acids lacking all the WD40 domains (Figure 2d). Therefore, *eed<sup>ul4</sup>* is very likely a null allele and henceforth referred to as *eed<sup>-</sup>*.

## PARTIE I : LA REPRESSION POLYCOMB



**Figure 2.** Loss of *eed* gene function leads to zebrafish death at around 10–12 dpf: (a) schematic representation of the genomic structure of the *eed* gene, with coding and untranslated sequences depicted as solid and open boxes, respectively. The sequence targeted by the TALEN in exon 3 is indicated with left and right TALEN binding sites shown in red, while the BseNI restriction sequence is underlined; (b) identification of mutant embryos using restriction fragment length polymorphism. Genomic DNA was prepared from an uninjected (Control) and an *eed* TALEN injected (TAL-*eed*) embryo, amplified by PCR and subjected to BseNI digestion. The TAL-*eed* injected embryo contains undigested material (arrow at 407 bp), indicating that the BseNI diagnostic restriction site has been disrupted; (c) sequence of the mutant allele compared to its wild-type counterpart. Dashes show deleted nucleotides. The mutated *eed*<sup>ul4</sup> allele possesses a deletion of 14 nucleotides; (d) schematic representation of wild-type (*Eed*<sup>wt</sup>) and predicted mutant (*Eed*<sup>ul4</sup>) proteins. The gray line in the predicted mutant protein corresponds to residues read out of frame prior to encountering a premature STOP codon. The green triangles in the wild-type protein show the WD40 domains; (e) Kaplan–Meier survival curves over 18 days for *eed*<sup>+/+</sup> (red curve) and *eed*<sup>-/-</sup> (green curve) siblings from a cross between heterozygous *eed*<sup>+/+</sup> fish. The number of fish considered is indicated; (f) measurement of the total length of *eed*<sup>+/+</sup> (red) and *eed*<sup>-/-</sup> (green) siblings from a cross between heterozygous *eed*<sup>+/+</sup> fish at 11 dpf. Statistical significance was assessed using a t-test; (g) measurement of the eye diameter of *eed*<sup>+/+</sup> (red) and *eed*<sup>-/-</sup> (green) siblings from a cross between heterozygous *eed*<sup>+/+</sup> fish at 11 dpf. Statistical significance was assessed using a Student *t*-test.

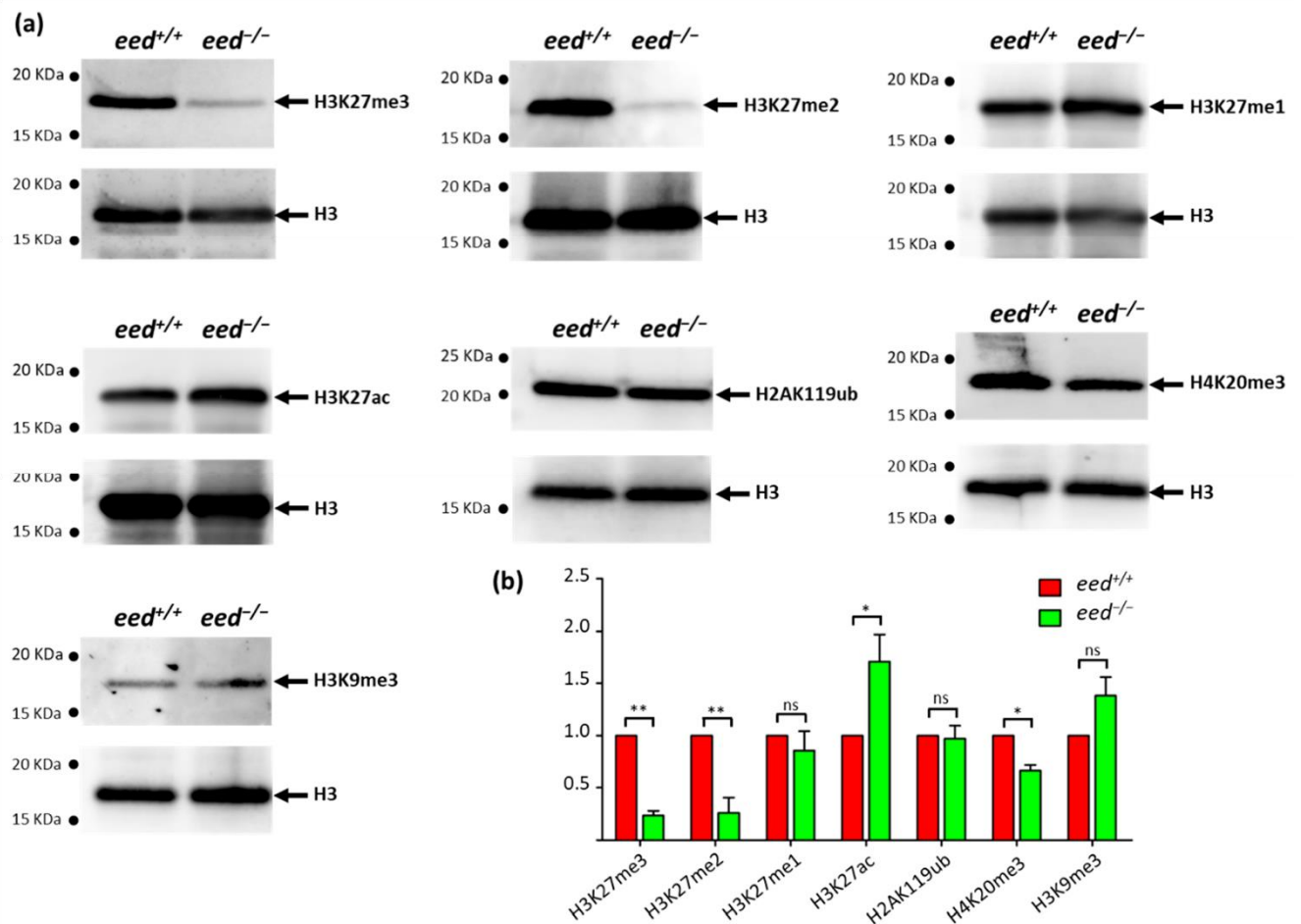
Adult heterozygous *eed*<sup>+/+</sup> fish are viable, fertile, and do not show any obvious phenotype. In contrast, and similarly to *ezh2*<sup>-/-</sup> [37,38] and *suz12a*<sup>-/-</sup>; *suz12b*<sup>-/-</sup> fish [39], homozygous *eed*<sup>-/-</sup> mutants die at around 10–12 dpf (Figure 2e). Analysis of the living *eed*<sup>-/-</sup> mutants at 11 dpf showed that they do not present gross morphological alterations while they appear significantly smaller than wild-type siblings (Figure 2f,g). Then, *eed* function is not required for early development and normal formation of the body plan but is necessary for zebrafish survival after 10–12 dpf.

### Loss of *eed* Function Alters H3K27me2/3 Levels, but Not H3K27me1 nor H2AK119ub

To investigate the effect of *eed* loss of function on histone modifications, we performed a series of Western blot analyses using antibodies recognizing specific histone marks (Figure 3). Histones were extracted from *eed*<sup>+/+</sup> and *eed*<sup>-/-</sup> siblings at 9 dpf and subjected to Western blot analyses to determine global levels of several histone modifications.



## PARTIE I : LA REPRESSION POLYCOMB



**Figure 3.** Effects of *eed* loss of function on global levels of histone modifications: (a) after genotyping of the caudal extremity of larvae, 10 to 15  $\mu$ g of bulk histones from a pool of 10 *eed*<sup>+/+</sup> or *eed*<sup>-/-</sup> siblings at 9 dpf were analyzed by Western blot using specific antibodies recognizing H3K27me3, H3K27me2, H3K27me1, H2AK119ub, H3K27ac, H3K9me3 or H4K20me3. After stripping, the membranes were reprobed with an anti-histone H3 as a control; (b) comparison between *eed*<sup>+/+</sup> (red) and *eed*<sup>-/-</sup> (green) siblings, of histone modification signals normalized by total histone H3 and expressed relative to the levels found in the wild-type fish. Quantified levels are the mean  $\pm$  SD of at least three independent histone extractions, except for H3K9me3, where two histone extraction were performed, followed by Western blot analysis. Statistical significance was assessed using a Student *t*-test; ns, non-significant; \*,  $p < 0.05$ ; \*\*,  $p < 0.01$ . (See Supplementary Materials).

Not surprisingly, a great reduction in bulk H3K27me3 and H3K27me2 levels is observed in *eed*<sup>-/-</sup> homozygous larvae at 9 dpf. However, faint signals for H3K27me2/3 methylations are still detected in homozygous mutant fish, presumably due to the activity of the maternal products initially deposited in the embryos. In contrast, we did not find significant differences in H3K27me1 levels between wild-type and mutant larvae. This puzzling observation might suggest that in the absence of PRC2 activity, other histone lysine methyltransferases could monomethylate H3K27 and compensate for PRC2 loss of activity. Because mammalian G9a (EHMT2) and Glp (EHMT1) histone lysine methyltransferases have been shown to target H3K27 in vivo [56,57], their zebrafish orthologs (Ehmt1a, Ehmt1b, and Ehmt2) are potential candidates that could maintain H3K27me1 levels in the absence of PRC2 function. In agreement with other studies using knockout embryonic stem cells (ESCs) and in pathological situations where the loss of H3K27me2/3 correlates with increased acetylation at H3K27 [58,59], we also found that *eed*<sup>-/-</sup> homozygous larvae display a slight increase in global H3K27ac levels at 9 dpf (Figure 3).

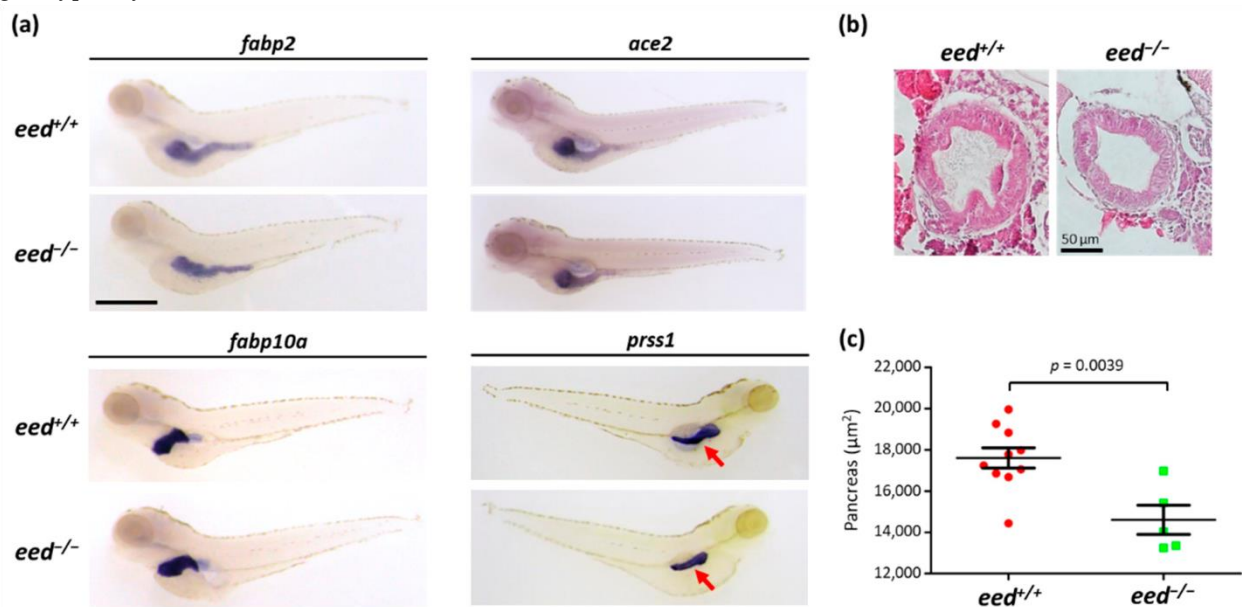
Monoubiquitylation of H2AK119 contributes to polycomb repression and is achieved by the action of the catalytic subunits RING1/RNF2 of PRC1 [10,11]. PRC1 complexes consist of a variety of protein assemblies with diverse compositions and are classified as canonical (cPRC1) and non-canonical (ncPRC1) complexes [60,61]. This classification is mainly based on the presence of a chromobox (CBX) protein in cPRC1 complexes and YY1-binding

protein (RYBP), or its homolog YAF2, in ncPRC1 complexes. Whereas cPRC1 is recruited to H3K27me3 via its CBX subunit, ncPRC1 complexes mediate H2A ubiquitylation at polycomb target sites independently of PRC2 action and of the presence of H3K27me3 epigenetic marks [60–62]. The examination of bulk H2AK119ub did not show differences between *eed*<sup>-/-</sup> and *eed*<sup>+/+</sup> homozygous larvae at 9 dpf, indicating that in the absence of PRC2 function, ncPRC1 complexes are able to maintain normal levels of H2AK119ub (Figure 3).

Finally, to check whether an increase in the heterochromatin marks H3K9me3 and H4K20me3 could compensate for the decrease in H3K27me2/3, we investigated their relative abundance in *eed*<sup>-/-</sup> homozygous mutant larvae at 9 dpf, compared to wild-type. If a decrease in H4K20me3 levels could be found in PRC2-deficient fish, we were not able to detect significant differences in bulk H3K9me3 (Figure 3). Then, the normal formation of the body plan in *eed*<sup>-/-</sup> mutants cannot be explained by the compensation of H3K27me2/3 loss through the increase in other repressive marks.

### Loss of *eed* Function Alters the Homeostasis of Digestive Organs in Zebrafish

Because it has been shown that *ezh2*-deficient zebrafish develop intestinal defects [38,63], we investigated the potential effect of *eed* loss of function on the development of digestive organs. First, we examined the expression of markers of the intestine (*fabp2*, fatty acidbinding protein 2, intestinal and *ace2*, angiotensin I converting enzyme 2), the liver (*fabp10a*, fatty acid-binding protein 10a, liver basic), and the exocrine pancreas (*prss1*, serine protease 1 also known as trypsin, *try*) in 5-dpf old larvae by in situ hybridization (Figure 4a). Pools of larvae from heterozygous *eed*<sup>+/-</sup> incrosses were subjected to in situ hybridization and individually pictured before being genotyped by RFLP.

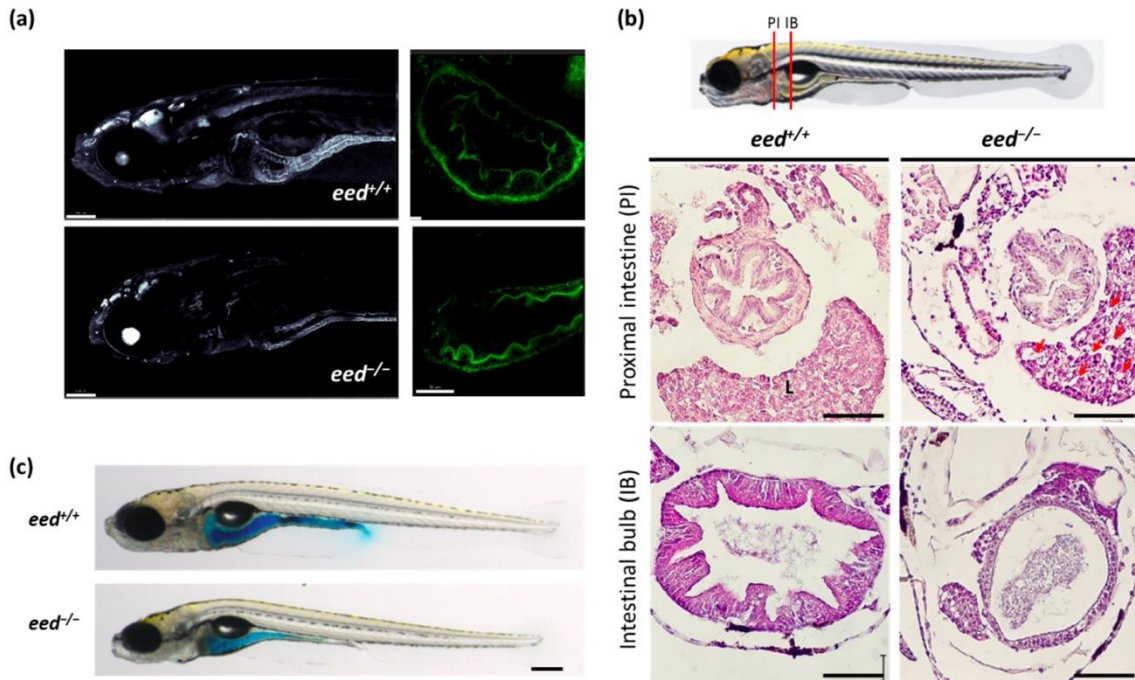


**Figure 4.** Organization of the digestive organs at 5 dpf: (a) whole-mount RNA in situ hybridization to detect the expression of markers of the intestine (*fabp2*, *ace2*), the liver (*fabp10a*), and the exocrine pancreas (*prss1*) in *eed*<sup>+/+</sup> and *eed*<sup>-/-</sup> siblings at 5 dpf. The red arrow shows the pancreas. Scale bar is 500  $\mu$ m; (b) intestinal bulb sections from *eed*<sup>+/+</sup> (left) and *eed*<sup>-/-</sup> (right) larvae at 5 dpf stained with hematoxylin and eosin. Scale bar is 50  $\mu$ m; (c) measurement of the surface of the pancreas labeled by in situ hybridization using a *prss1* probe at 5 dpf for *eed*<sup>+/+</sup> (red) and *eed*<sup>-/-</sup> (green) larvae. Statistical significance was assessed by a Student t-test, and the corresponding *p*-value is indicated.

The expression of *fabp2*, *ace2*, and *fabp10a* revealed that the intestine and the liver are normally developed in *eed*<sup>-/-</sup> larvae at 5 dpf. Furthermore, histological studies showed that the structure of the intestine is identical in *eed*<sup>+/+</sup> and *eed*<sup>-/-</sup> zebrafish at this stage (Figure 4b). In contrast, and similar to what was observed in *ezh2*<sup>-/-</sup> mutants [38], the size of the pancreas appears smaller in *eed*-deficient fish at 5 dpf (Figure 4a,c).

To visualize the intestine structure in the whole larvae at a later developmental stage by confocal microscopy, we used the transgenic line Tg(*actb2*:GFP-Hsa.UTR)<sup>e116</sup> in which actin is labeled by the green fluorescent protein

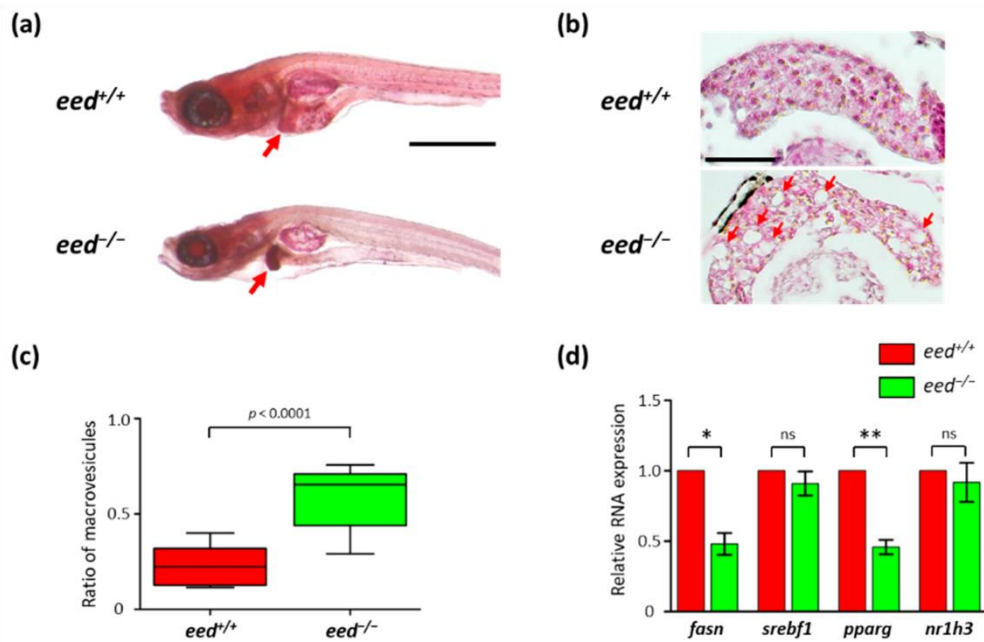
(GFP) [40,64,65]. Fluorescence imaging revealed an alteration of the intestine structure of the *eed*<sup>-/-</sup> transgenics at 9 dpf. Indeed, the number of cells forming the thickness of the intestinal wall and visible through GFP expression is decreased in *eed*-deficient zebrafish as of 9 dpf (Figure 5a). Then, while the intestine develops normally until 5 dpf, its integrity fails to be maintained at later developmental stages.



**Figure 5.** Structure of the intestine at 9–11 dpf: (a). confocal imaging of the anterior part (left, scale bar is 150  $\mu$ m) and of the intestine bulb (right, scale bar is 50  $\mu$ m) of transgenic Tg (*actb2*:GFP-Hsa.UTRN)<sup>e116</sup> zebrafish larvae, wild-type (up) or lacking *eed* function (down) at 9 dpf; (b) histological sections stained with hematoxylin and eosin at the levels of the proximal intestine (PI) and the intestinal bulb (IB) as indicated from *eed*<sup>+/+</sup> (left) and *eed*<sup>-/-</sup> (right) siblings at 11 dpf. Red arrows show macrovesicles. L, liver. Scale bar is 50  $\mu$ m; (c) Smurf assay performed on *eed*<sup>+/+</sup> and *eed*<sup>-/-</sup> siblings at 11 dpf. Scale bar is 200  $\mu$ m.

The analysis of histological sections of larvae at 11 dpf confirmed that the intestine wall of *eed*<sup>-/-</sup> mutants is strongly reduced and lacks folds at the level of the intestinal bulb (Figure 5b). By contrast, the structure of the proximal intestine does not show these strong defects, as previously observed in *ezh2*<sup>-/-</sup> mutants [38]. Because the intestine looks normal at 5 dpf but altered at 9–11 dpf, we conclude that in the absence of *eed* and PRC2 functions, the development of the intestine occurs, but the integrity of its structure is not maintained. To investigate the permeability of the intestine in mutants, we performed a Smurf assay [47]. In this assay, the permeability of the intestine was assessed by evaluating the diffusion throughout the body of blue food dye present in the fish water. Figure 5c shows that in spite of a strong alteration of the structure of the intestine bulb, the intestinal permeability is not increased.

Neutral lipid dye oil red-O staining of larval zebrafish is commonly used to examine steatosis caused by consumption of a high-fat diet or induced by various drugs [66–70]. Oil red-O staining clearly shows a strong accumulation of neutral lipids in the liver of *eed*<sup>-/-</sup> larvae à 11 dpf when compared to wild-type larvae (Figure 6a). Such a strong oil red-O staining of the liver was also found in *ezh2*<sup>-/-</sup> homozygous mutants [38], whereas exposure to Ezh1/2 inhibitor PF-06726304 acetate increases lipid accumulation in larval zebrafish [71]. Moreover, hematoxylin and eosin stain on histological sections of zebrafish larvae at 11 dpf showed signs of macrovesicular steatosis in *eed*<sup>-/-</sup> larvae (Figures 5b and 6b,c).



**Figure 6.** Liver alterations in *eed*-deficient larvae: (a) representative images of lateral views of *eed*<sup>+/+</sup> and *eed*<sup>-/-</sup> larvae stained with oil red-O at 11 dpf. The red arrow shows the liver. Scale bar is 500 µm; (b) histological sections stained with hematoxylin and eosin showing the liver of *eed*<sup>+/+</sup> and *eed*<sup>-/-</sup> larvae at 11 dpf. Red arrows show macrovesicles. Scale bar is 50 µm; (c) quantification of macrovesicular steatosis *eed*<sup>+/+</sup> (red) and *eed*<sup>-/-</sup> (green) larvae at 11 dpf. The ratio of macrovesicular structures relative to total hepatic surface was calculated on 3 histological slices per larvae from 4 different larvae per genotype. Statistical significance was assessed using a Mann-Whitney *t*-test;

(d) relative levels of *fasn*, *srebf1*, *pparg*, and *nr1h3* RNA abundances in 9 dpf old *eed*<sup>+/+</sup> (red) and *eed*<sup>-/-</sup> (green) larvae quantified by RT-qPCR. Three independent experiments were performed, and error bars represent standard deviation. Statistical analysis was performed using a one-way ANOVA test followed by a Dunn's multiple test comparison; ns, non-significant; \*, *p* < 0.05; \*\*, *p* < 0.01.

To investigate the impact of *eed* loss of function on the expression of lipogenic factors, we performed RT-qPCRs to determine RNA abundance for *fasn* (fatty acid synthetase), *srebf1* (sterol regulatory element-binding transcription factor 1), *pparg* (peroxisome proliferator-activated receptor gamma), and *nr1h3* (nuclear receptor subfamily 1, group H, member 3, also known as liver X receptor *lxr*) at 9 dpf. Then, in spite of showing an increase in lipid accumulation, RNA levels for lipogenic factors remain either unchanged (*srebf1* and *nr1h3*) or reduced (*fasn* and *pparg*) in *eed*<sup>-/-</sup> larvae (Figure 6d). This suggests that *eed* loss of function-induced steatosis is likely the result of an increase in lipid storage in the liver rather than an increase in lipid synthesis.

Taken together, our study of *eed*-deficient zebrafish mutants demonstrates that PRC2 loss of function has various effects on the homeostasis of digestive organs, ranging from a delay in pancreatic development, a loss of maintenance of the intestine structures to hepatic steatosis.

### Loss of *eed* Function Alters Neuronal Differentiation

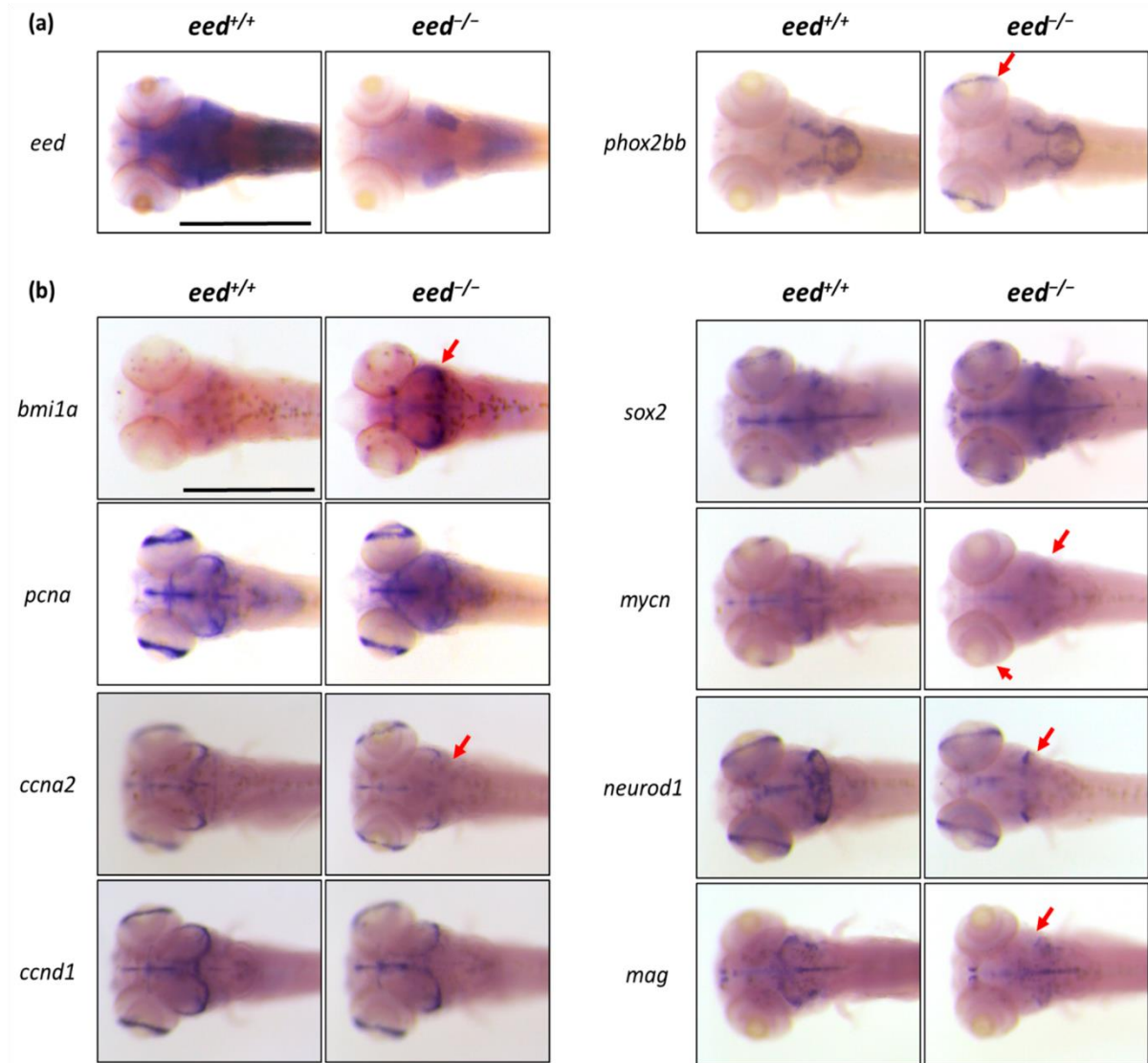
To characterize in more detail our *eed*-deficient zebrafish mutants, we performed additional whole-mount RNA in situ hybridization experiments at 5 dpf with a focus on the brain. Mutant and wild-type siblings were treated in one batch for each probe and documented photographically. The genotype of each embryo was then determined by RFLP on genomic DNA extracted from the fixed and stained material. First, we showed that *eed* expression is strongly reduced in *eed*<sup>-/-</sup> larvae, presumably due to nonsensemediated decay (NMD) of the mutant transcript (Figure 7a). Second, we investigated the expression profile of *phox2bb* (*paired like homeobox 2Bb*), a transcription factor expressed by a subset of cells in the posterior regions of the brain. In *eed*<sup>-/-</sup> mutants, *phox2bb* appears additionally expressed ectopically in the retina (Figure 7a). This *phox2bb* misexpression in the retina was also observed in *ezh2*-deficient zebrafish larvae [38]. Then, *eed* loss of function recapitulates the defects found in



*ezh2*<sup>-/-</sup> mutants at the level of the digestive organs, but also in terms of *phox2bb* expression in the brain and the retina. Moreover, while *eed* is expressed in large territories of the brain, its loss of function impairs *phox2bb* specifically in the retina, indicating that PRC2 controls gene expression programs in a very subtle and discreet manner, depending on the cell type.

To gain insights into the possible role of PRC2 in brain development, we applied in situ hybridizations at 5 dpf on several genes known to be expressed in the midbrain-hindbrain boundary region that contains neural stem cells and dividing progenitors [72], and where *eed* transcripts are abundant (Figures 1a and 7a) [54]. The genes chosen for the study were the stemness markers *bmi1a* (*bmi1 polycomb ring finger oncogene 1a*) and *sox2* (SRY-box transcription factor 2); the proliferation marker *pcna* (proliferating cell nuclear antigen); *mycn* (MYCN proto-oncogene, bHLH transcription factor), which is involved in several signaling pathways promoting cell growth, proliferation and metabolism of progenitor cells in different developing organs and tissues; two genes coding for cyclins, *ccna2* (cyclin A2) and *ccnd1* (cyclin D1); *neurod1* (*neuronal differentiation 1*), a marker of neuronal precursors; and *mag* (myelin-associated glycoprotein), a marker of glial precursors (Figure 7b).

Analysis of expression profiles of the stemness markers *bmi1a* and *sox2* revealed a strong increase in *bmi1a* labeling at the midbrain-hindbrain boundary of *eed*<sup>-/-</sup> mutants suggesting that the number of stem cells is increased, possibly due to differentiation defects in this region. In contrast, *sox2* expression seems less affected by the loss of *eed* function, with eventually a slight increase in *sox2* RNA abundance in the hindbrain. The RNA expression profile of *pcna* appeared globally similar in *eed*<sup>+/+</sup> and *eed*<sup>-/-</sup> larvae indicating that PRC2 loss of function has little or no effect on cell proliferation in the developing brain at 5 dpf. However, while expression of the cyclin gene *ccnd1* is unchanged in *eed*<sup>-/-</sup> larvae, transcription of *ccna2* appears in part slightly deregulated. We observed that the expression of *mycn* is downregulated in the hindbrain and the retina of the *eed* mutants. Finally, using probes for neuronal precursors (*neurod1*) and glial precursors (*mag*), whole-mount RNA in situ hybridizations showed a reduction in the labeled areas indicating that the numbers of these precursors are diminished in the hindbrain or midbrain of *eed*<sup>-/-</sup> larvae, respectively. Altogether, our results suggest that PRC2 loss of function disturbs the differentiation and the development of a subset of neurogenic cells in the midbrain-hindbrain region and in the retina without dramatically affecting cell proliferation, as judged by *pcna* expression. Thus, remarkably, while *eed* loss of function drastically affects *bmi1a* expression, the effect of the mutation on the expression of other markers such as *ccna2*, *mycn*, *neurod1*, or *mag* remains relatively thin and affects discrete areas of the brain.



**Figure 7.** Whole-mount RNA in situ hybridization of the brain region of *eed*<sup>+/+</sup> and *eed*<sup>-/-</sup> siblings at 5 dpf: (a) whole-mount in situ hybridization to detect *eed* (left) and *phox2bb* RNA expression. The red arrow shows the *phox2bb* ectopic expression in the retina of *eed* mutants. Scale bar is 500  $\mu$ m; (b) whole-mount in situ hybridization to detect the expression profiles of stemness (*bmi1a*, *sox2*), proliferation (*pcna*, *mycn*), cell cycle (*ccna2*, *ccnd1*), and neuronal/glial precursor (*neurod1*, *mag*) markers. The red arrows emphasize expression profile differences between *eed*<sup>+/+</sup> and *eed*<sup>-/-</sup> larvae. Scale bar is 500  $\mu$ m.

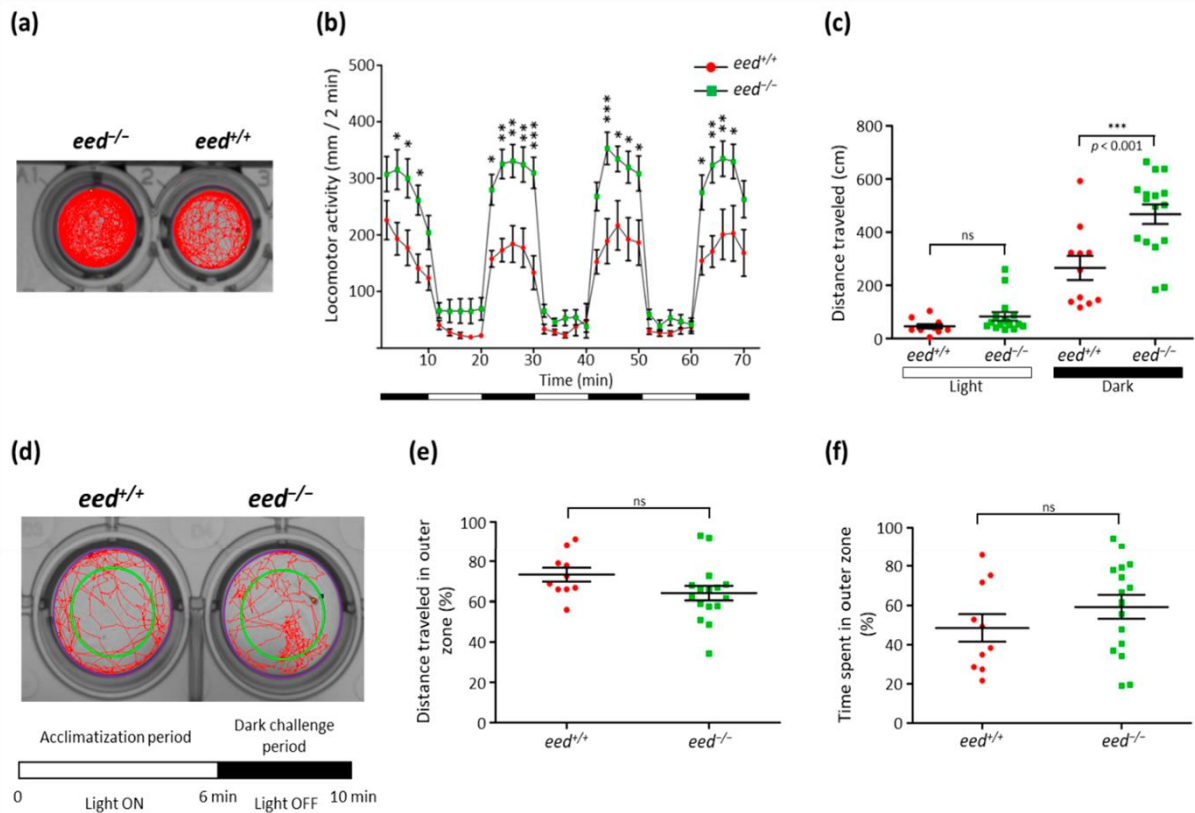
### Loss of *eed* Function Alters Locomotor Activity

Loss of *eed* function alters the differentiation of neurogenic cells at 5 dpf. Then, to investigate whether *eed* deficiency could also affect larval behavior, we performed locomotor and thigmotaxis assays. Zebrafish larvae from heterozygous *eed*<sup>+/-</sup> incrosses were individually disposed of into dish wells. The plates were handled minimally before placement in a Zebrafish chamber (ViewPoint Life Sciences, Lyon, France) equipped with an infrared light-emitting floor and a top-mounted infrared camera allowing video recording of the whole plate under both light and dark conditions. Larval behavior measurements

were achieved using the ZebraLab software (Version 5.18.0.0, ViewPoint Life Sciences, Lyon, France). After recording, the larvae were euthanized and their genotype determined by RFLP.

The locomotor activity assays were conducted in 48-well plates following a protocol that consists of a 10 min initial acclimating period in the dark, followed by six alternating 10 min light and dark phases (Figure 8a). As previously described [73], switching to light dramatically decreases larval activity, whereas return to darkness is associated with an increase in locomotor activity. Additional cycles of alternating light and dark produce alternating levels of

low and high activity of the zebrafish larvae, respectively (Figure 8b). Behaviorally, *eed*<sup>-/-</sup> larvae are hyperactive compared with wild-type larvae at 5 dpf. This hyperactive phenotype of *eed*<sup>-/-</sup> mutant is particularly prominent and significant during the dark cycles (Figure 8b,c).



**Figure 8.** Behavioral comparison between wild-type and *eed*<sup>-/-</sup> mutants at 5 dpf: (a) locomotor tracking for a 70 min recording of an *eed*<sup>-/-</sup> mutant (left) and a wild-type (right) larvae; (b) distance traveled throughout a 70 min session for wild-type (red, *n* = 11) and *eed*<sup>-/-</sup> mutant (green, *n* = 16). Data are presented as mean ± SD of the distance moved (in mm) in 2 min intervals. Black and white bars at the bottom indicate dark and light conditions, respectively. Statistical analysis was performed using two-way ANOVA followed by Bonferroni posttest comparisons; ns, non-significant; \*, *p* < 0.05; \*\*, *p* < 0.01; \*\*\*, *p* < 0.001; (c) cumulative distance traveled for each wild-type (red) and mutant (green) larvae during the light (left) and dark (right) periods. Statistical analysis was performed using one-way ANOVA followed by Bonferroni posttest comparisons; ns, non-significant; \*\*\*, *p* < 0.001; (d) the experimental procedure for the thigmotaxis assay is composed of a 6 min acclimatization period with the light ON followed by a 4 min visual motor challenge period with the lights OFF. Experiments are performed in a 24-well plate format. Inner and outer zones are delineated by the green circle in a way that the two zones cover equivalent spatial areas. The distance traveled during the dark challenge period is shown for two larvae; (e) cumulative distance traveled in the outer zone of wild-type (red, *n* = 10) and *eed*<sup>-/-</sup> mutant (green, *n* = 16) larvae.

Each point represents individual larvae. No statistical (ns) difference was found between wild-type and *eed*<sup>-/-</sup> mutants; (f) cumulative time spent in the outer zone of wild-type (red, *n* = 10) and *eed*<sup>-/-</sup> mutant (green, *n* = 16) larvae. Each point represents unique larvae. No statistical (ns) difference was found between wild-type and *eed*<sup>-/-</sup> mutants.

Thigmotaxis (or “wall hugging”) is a common behavioral endpoint used in preclinical studies employing rodents but also applied to the zebrafish model [74,75]. Animals engaged in thigmotaxic behavior strongly avoid the center of an arena and reduce their exploration behavior to stay or move in close proximity to the boundaries of the environment. Thigmotaxis assays were conducted in 24-well plates using a protocol composed of a 6 min acclimatization period in the light followed by a 4 min dark challenge period. The arena was divided into two areas, inner and outer, of equivalent surface, and thigmotaxis was presented as the percentage of the total distance moved or the total time spent by each larva into the outer zone (Figure 8d). Our measurements showed that there is no significant difference in thigmotaxis behavior between *eed*<sup>-/-</sup> mutants and wild-type larvae at 5 dpf (Figure 8e,f).

Thus, in spite of having a hyperactive phenotype, the *eed*<sup>-/-</sup> mutants do not present an increase in their exploration behavior when compared to wild-type siblings.

### Discussion

PRC2 is a histone-modifying protein complex conserved from fruit flies to humans and able to deposit H3K27me1/2/3 epigenetic marks responsible in turn to local chromatin compaction and gene repression. Genome-wide studies using human or mouse embryonic stem cells showed that PRC2 and H3K27me3 marks are enriched at promoters of numerous developmental genes, leading to the proposal that PRC2 could be involved in the maintenance of the pluripotency of stem cells by keeping developmental genes repressed [12–19]. Differentiation is then associated with the relocation of the PRC2 complexes turning stem cell genes expression off and switching on gene programs specific to given developmental lineages. PRC2 is composed of the core subunit EZH1/2, EED, SUZ12, and RBBP4/7 together with a wide range of substoichiometric subunits defining distinct alternate protein complexes [22]. However, PRC2 catalytic activity relies only on the presence of the two mutually exclusive histone methyltransferases EZH1 or EZH2, associated with EED and SUZ12. In zebrafish, as the result of a genome duplication that occurred in the teleost lineage, the Suz12 subunit is encoded by two ohnologues, *suz12a* and *suz12b*. Then, *eed* is the sole zebrafish gene coding for a PRC2 subunit present in all alternate complexes. Consequently, since EED is required for EZH1/2 methyltransferase activity, it is expected that inactivation of the zebrafish *eed* gene will result in the loss of activity of all PRC2 complexes.

Using the TALEN technology, we generated an *eed* allele harboring a 14 bp deletion within the third exon. This deletion conducts to a frameshift within the coding sequence leading to the production of a putative protein lacking all the WD40 domains, predicted to be non-functional, and therefore likely producing an *eed* null allele. Heterozygous *eed*<sup>+/-</sup> fish are viable, fertile, and do not present visible gross phenotypes. By opposition, *eed* homozygosity leads to larval death at around 12 dpf. This time point is similar to what was observed for zebrafish lacking the function of other PRC2 components, such as for *ezh2* mutants [37,38] or *suz12a; suz12b* double-mutants [39], but contrasts to *Eed* knockout mouse phenotype [25]. While mouse embryos homozygous for the *Eed* mutation fail to gastrulate normally, *eed* embryonic expression is not required for the establishment of the zebrafish body plan. This relatively late lethality makes the zebrafish a compelling model to study PRC2 function during the development of various tissues without the requirement of conditional knockout strategies. It also reinforces the idea that PRC2 function plays little or no role in the implementation of developmental decisions but is mainly required to maintain these developmental instructions.

Zebrafish *eed*<sup>-/-</sup> mutants display various alterations of the digestive organs. These defects include a smaller pancreas at 5 dpf, hepatic steatosis characterized by an increase in lipids and macrovesicles in the liver, and the loss of maintenance of the intestine structures particularly marked at the level of the intestinal bulb. Studies in mice also pointed out the role of *Eed* in intestinal development [76]. Indeed, the conditional knockout of *Eed* in postnatal intestinal crypts decreases the stem cell pools, reduces the proliferation potential, increases the secretory cell differentiation, and results in smaller mouse intestines. Zebrafish larvae hatch with a yolk sac, which is depleted at 6–7 dpf. Then, after that time, their growth relies on external food uptake. If left unfed, the larvae will usually die of starvation at around 12 dpf [77]. Thus, it is very likely that the defects of the digestive organs found in the *eed*<sup>-/-</sup> homozygous mutants account for their larval death at 12 dpf. Furthermore, the marked *eed*<sup>-/-</sup> phenotype at the level of the digestive organs closely resembles the phenotype of the *ezh2*<sup>-/-</sup> homozygous mutants we have previously described [38]. These indistinguishable phenotypes exclude critical roles of *Ezh1* and/or PRC2-independent *Ezh2* functions in the development of the digestive organs, at least until 12 dpf.

Investigations on bulk histone modifications in *eed*-deficient larvae at 9 dpf revealed a dramatic decrease in H3K27me2/3 marks, as expected. However, *eed*<sup>-/-</sup> larvae retained some H3K27me2 and H3K27me3 residual methylation marks. We assume that the remaining H3K27me2/3 signals were deposited by the maternal PRC2 activity and retained in slowly dividing or quiescent cells. Maternal PRC2 contribution was previously demonstrated by showing that maternal-zygotic *MZezh2*<sup>-/-</sup> mutants display a much stronger phenotype than zygotic *ezh2*<sup>-/-</sup> mutants, with small eyes, accumulation of blood near the yolk extension, stringy heart, heart edema, and absence of pectoral fins at 2 dpf [37]. EZH1 and EZH2 are both able to deposit H3K27me1, but the role of PRC2 in the control of H3K27me1 monomethylation is still a matter of debate [20,58,78,79]. Studies using *Eed* and *Suz12* knockout mouse ESC lines, as well as knockdown experiments in mouse ESCs, showed that the inhibition of PRC2



activity is responsible for a global loss of H3K27me2 and H3K27me3 levels, but at best for a partial reduction in H3K27me1 marks [24,80,81]. Our analysis in zebrafish revealed that H3K27me1 global levels are not significantly changed in *eed*<sup>-/-</sup> larvae indicating that in the absence of PRC2 function, H3K27 monomethylation is achieved by other histone methylation routes. It has been shown that the histone lysine methyltransferases G9a (EHMT2) and Glp (EHMT1), in addition to being potent H3K9 methyltransferases, are able to monomethylate H3K27 both in vitro and in mouse ESCs [56,57,82]. Thus, their zebrafish orthologs, *Ehmt1a*, *Ehmt1b*, and *Ehmt2*, appear as suitable candidates to account for monomethylation at H3K27 targets in the absence of PRC2 activity.

Polycomb repression is achieved by the combined action of both PRC2 and PRC1. While PRC2 methylates H3K27, PRC1 monoubiquitylates H2AK119. Multiple PRC1 subcomplexes have been shown to coexist in cells. These complexes are commonly referred to as canonical (cPRC1) or non-canonical (ncPRC1), based on their dependence on H3K27me3 recruitment to target loci [6,83]. We have shown that global H2AK119ub levels were unchanged in *eed*<sup>-/-</sup> larvae indicating H2K119 monoubiquitylation is maintained in the absence of PRC2 function, presumably thanks to ncPRC1 action. This situation is similar to what was observed in ESCs, where *Ezh1/Ezh2* double knockout was reported to do not affect H2K119ub levels [58]. H2K119ub maintenance in *eed*<sup>-/-</sup> mutants also explains the milder phenotype observed in *eed*-deficient fish when compared to *rnf2*<sup>-/-</sup> mutants. In zebrafish, *Rnf2* is the only catalytic subunit common to all PRC1 subcomplexes [84]. Zebrafish deficient in *rnf2* gene function lack H2K119ub marks, display several defects including jaw malformation, pericardial edema, diminished blood circulation, absence of pectoral fins, severe craniofacial phenotype, and die at around 4–5 dpf [85,86]. However, our results contrast with another study using maternal-zygotic *MZezh2* zebrafish mutants [87]. In this report, it was shown that loss of *Ezh2* function from maternal and zygotic origins resulted in dramatic impairment of PRC1 recruitment at the chromatin and complete loss of H3K119ub marks. This observation highlighted the role of the PRC2 in the functional recruitment of cPRC1 complexes and led us to propose that in the absence of zygotic *Eed* function, maternal PRC2 activity is sufficient to deposit enough H3K27me3 marks allowing cPRC1 recruitment at early developmental stages. Then, at later stages, H2AK119ub marks are propagated and transmitted through cell generations in a PRC2-independent fashion via the action of ncPRC1 complexes. This hypothesis would explain why *rnf2*<sup>-/-</sup> and *MZezh2*<sup>-/-</sup> mutants display more severe phenotypes than *ezh2* and *eed* homozygous zygotic mutants do. Furthermore, it has been demonstrated that ncPRC1 complexes are able to propagate H2AK119ub during cell division via their recruitment to the chromatin through the binding of the RYBP/YAF2 subunits to H2AK119ub marks [88].

Sequencing studies on pediatric high-grade gliomas, including glioblastoma multiformes (GBMs) and diffuse intrinsic pontine gliomas (DIPGs), identify recurrent somatic lysine to methionine substitutions at position 27 (K27M) in histone H3. The H3<sup>K27M</sup> mutation primarily occurs in *H3F3A* (>70%), one of the two genes coding for the histone variant H3.3, with a lower frequency in *HIST1H3B* and *HIST1H3C* encoding H3.1, and more rarely in *HIST2H3C* coding for H3.2 [89–92]. The H3.3<sup>K27M</sup> mutated histone binds to EZH2 and inhibits the methyltransferase activity of the PRC2, leading to a dramatic decrease in global H3K27me3 marks and alterations in gene expression programs [93–97]. These changes in gene expression would sustain the undifferentiated phenotype of the tumor cells that contributes to the aggressiveness of the gliomas and to the dismal associated prognosis [94]. In addition to *H3F3A*<sup>K27M</sup> mutations, DIPGs also contain genetic alterations targeting canonical signaling pathways, most frequently *TP53* loss of function (about 60%) and gain-of-function mutations or amplifications in platelet-derived growth factor receptor alpha (*PDGFRA*; about 40%) [89,98], as well as numerous other mutations that might contribute to significant inter- and intratumoral DIPG heterogeneity and found at lower-frequencies [99]. However, the individual contribution of each of the different mutations, including *H3F3A*<sup>K27M</sup>, in glioma genesis remains poorly understood. In this context, we took advantage of our *eed*-deficient zebrafish line to investigate the effects of PRC2 loss of activity on neuronal differentiation and proliferation. We found that in *eed*<sup>-/-</sup> larvae, expression of some stemness (*bmi1a*), as well as neuronal (*neurod1*) or glial (*mag*) precursor markers, is altered, indicating that neural differentiation is impaired in the absence of functional PRC2 activity. Interestingly, a number of human DIPGs have been shown positive for BMI1 expression, which then appears as a potential therapeutic target [100–102]. By contrast, our in situ hybridization experiments revealed that the *pcna* expression profile remained unchanged in the brain of *eed*<sup>-/-</sup> larvae suggesting that PRC2 loss of function and reduction in H3K27me3 marks do not affect the proliferation potential of neural cells. Thus, we propose that in DIPGs, the *H3F3A*<sup>K27M</sup> mutation is

responsible for the maintenance of an undifferentiated state, whereas other mutations might contribute to proliferation.

Surprisingly, we also showed that while expressed ubiquitously in the brain of wildtype larvae, *eed* loss of function results in gene expression alterations in relatively thin and discrete areas of the brain. For instance, ectopic expression of *phox2bb* is specifically found in the retina, whereas *neurod1* expression is lost in a subset part of the hindbrain. Then, the effects of PRC2 loss of function appear strictly cell-specific and context-dependent, differentially affecting cell fates. These subtle changes in gene expression would indubitably escape to global transcriptomic analyses. Thus, our observations also outline the importance of performing whole-mount in situ hybridizations, but not only global transcriptomic analyses, for the characterization of zebrafish lines harboring mutations in polycomb group genes. In this context, it is worth emphasizing that the loss of *eed* function gives different phenotypes depending on the considered organ. It affects neuronal differentiation in specific brain regions, the development of the pancreas, the maintenance of the intestine wall structures, or the metabolism of lipids in the liver. This variation in effects and phenotypes might reflect the fact that *Eed* controls different gene expression programs in the different organs. It also reinforces the idea that a better deciphering of PRC2 function would rely on transcriptomic analyses at the single-cell level.

Cohen-Gibson syndrome (COGIS, OMIM #617561) is an overgrowth disorder characterized by dysmorphic facial features, advanced bone age, skeletal abnormalities, associated with variable intellectual disability, and caused by de novo mutations in *EED* [30,103]. To date, COGIS (*EED*-related overgrowth) has been reported in eight individuals, with at least one patient showing hyperactivity [104]. Interestingly, our behavioral studies revealed that *eed*<sup>-/-</sup> larvae also present a hyperactive phenotype, raising the possibility that the *eed*-deficient zebrafish line could constitute a pertinent model to study the neurological defects leading to behavioral disabilities in COGIS.

### Supplementary Materials

The following are available online at <https://www.mdpi.com/article/10.3390/cells10113142/s1>, Figure S1: Uncropped RT-PCR gels used in Figure 1; Figure S2: Uncropped Western blot used in Figure 3.

### Author Contributions

Conceptualization and supervision: P.-O.A.; Methodology and data collection: L.R., P.V., S.H., J.C. and P.-O.A.; Formal Analysis: L.R., P.V., S.H., R.-A.T., E.A., I.V.S., X.L.B., and P.-O.A.; Resources and funding acquisition: R.-A.T., E.A., I.V.S., X.L.B. and P.-O.A.; Original draft preparation: P.-O.A.; Revision of the manuscript: L.R., P.V., R.-A.T., I.V.S. and X.L.B. All authors have read and agreed to the published version of the manuscript.

### Funding

This work was supported by the CNRS, Inserm, the University of Lille, the Ligue Contre le Cancer—Comité du Pas-de-Calais and the GIP Cancéropôle Nord-Ouest. This study was supported by a grant from Contrat de Plan Etat-Région CPER Cancer 2015–2020.

### Institutional Review Board Statement

Zebrafish were maintained in compliance with the French and European Union guidelines (Directive 2010/63/EU of the European Parliament and of the Council of 22 September 2010 on the protection of animals used for scientific purposes) and the experimental procedures on zebrafish were reviewed and approved by the local Ethics Committee, CEEA 75 Nord Pas-de-Calais, and the French Ministry of Higher Education and Research (APAFIS approval number 13527-2018011722529804\_v3).

### Acknowledgments

We are grateful to Barbara Dupret, Anaïs Chapart, Aurélie Bary, Amélie Boudelle, and Adeline Duquesnoy for technical assistance. S.H. was recipient of a visiting research scholarship from the Ministry of Science, Research and Technology of the Islamic Republic of Iran.

### References

1. Di Croce, L.; Helin, K. Transcriptional regulation by Polycomb group proteins. *Nat. Struct. Mol. Biol.* **2013**, *20*, 1147–1155. [[CrossRef](#)] [[PubMed](#)]
2. Sauvageau, M.; Sauvageau, G. Polycomb group proteins: Multi-faceted regulators of somatic stem cells and cancer. *Cell Stem Cell* **2010**, *7*, 299–313. [[CrossRef](#)]
3. Surface, L.E.; Thornton, S.R.; Boyer, L.A. Polycomb group proteins set the stage for early lineage commitment. *Cell Stem Cell* **2010**, *7*, 288–298. [[CrossRef](#)]
4. Geng, Z.; Gao, Z. Mammalian PRC1 Complexes: Compositional Complexity and Diverse Molecular Mechanisms. *Int. J. Mol. Sci.* **2020**, *21*, 8594. [[CrossRef](#)]
5. Glancy, E.; Ciferri, C.; Bracken, A.P. Structural basis for PRC2 engagement with chromatin. *Curr. Opin. Struct. Biol.* **2021**, *67*, 135–144. [[CrossRef](#)] [[PubMed](#)]
6. Piunti, A.; Shilatifard, A. The roles of Polycomb repressive complexes in mammalian development and cancer. *Nat. Rev. Mol. Cell Biol.* **2021**, *22*, 326–345. [[CrossRef](#)]
7. Cao, R.; Wang, L.; Wang, H.; Xia, L.; Erdjument-Bromage, H.; Tempst, P.; Jones, R.S.; Zhang, Y. Role of histone H3 lysine 27 methylation in Polycomb-group silencing. *Science* **2002**, *298*, 1039–1043. [[CrossRef](#)]
8. Kuzmichev, A.; Nishioka, K.; Erdjument-Bromage, H.; Tempst, P.; Reinberg, D. Histone methyltransferase activity associated with a human multiprotein complex containing the Enhancer of Zeste protein. *Genes Dev.* **2002**, *16*, 2893–2905. [[CrossRef](#)]
9. Müller, J.; Hart, C.M.; Francis, N.J.; Vargas, M.L.; Sengupta, A.; Wild, B.; Miller, E.L.; O'Connor, M.B.; Kingston, R.E.; Simon, J.A. Histone methyltransferase activity of a Drosophila Polycomb group repressor complex. *Cell* **2002**, *111*, 197–208. [[CrossRef](#)]
10. Wang, H.; Wang, L.; Erdjument-Bromage, H.; Vidal, M.; Tempst, P.; Jones, R.S.; Zhang, Y. Role of histone H2A ubiquitination in Polycomb silencing. *Nature* **2004**, *431*, 873–878. [[CrossRef](#)]
11. de Napoles, M.; Mermoud, J.E.; Wakao, R.; Tang, Y.A.; Endoh, M.; Appanah, R.; Nesterova, T.B.; Silva, J.; Otte, A.P.; Vidal, M.; *et al.* Polycomb group proteins Ring1A/B link ubiquitylation of histone H2A to heritable gene silencing and X inactivation. *Dev. Cell* **2004**, *7*, 663–676. [[CrossRef](#)]
12. Boyer, L.A.; Plath, K.; Zeitlinger, J.; Brambrink, T.; Medeiros, L.A.; Lee, T.I.; Levine, S.S.; Wernig, M.; Tajonar, A.; Ray, M.K.; *et al.* Polycomb complexes repress developmental regulators in murine embryonic stem cells. *Nature* **2006**, *441*, 349–353. [[CrossRef](#)] [[PubMed](#)]
13. Azuara, V.; Perry, P.; Sauer, S.; Spivakov, M.; Jørgensen, H.F.; John, R.M.; Gouti, M.; Casanova, M.; Warnes, G.; Merkenschlager, M.; *et al.* Chromatin signatures of pluripotent cell lines. *Nat. Cell Biol.* **2006**, *8*, 532–538. [[CrossRef](#)]
14. Bernstein, B.E.; Mikkelsen, T.S.; Xie, X.; Kamal, M.; Huebert, D.J.; Cuff, J.; Fry, B.; Meissner, A.; Wernig, M.; Plath, K.; *et al.* A bivalent chromatin structure marks key developmental genes in embryonic stem cells. *Cell* **2006**, *125*, 315–326. [[CrossRef](#)]
15. Bracken, A.P.; Dietrich, N.; Pasini, D.; Hansen, K.H.; Helin, K. Genome-wide mapping of Polycomb target genes unravels their roles in cell fate transitions. *Genes Dev.* **2006**, *20*, 1123–1136. [[CrossRef](#)] [[PubMed](#)]
16. Lee, T.I.; Jenner, R.G.; Boyer, L.A.; Guenther, M.G.; Levine, S.S.; Kumar, R.M.; Chevalier, B.; Johnstone, S.E.; Cole, M.F.; Isono, K.; *et al.* Control of developmental regulators by Polycomb in human embryonic stem cells. *Cell* **2006**, *125*, 301–313. [[CrossRef](#)] [[PubMed](#)]
17. Mikkelsen, T.S.; Ku, M.; Jaffe, D.B.; Issac, B.; Lieberman, E.; Giannoukos, G.; Alvarez, P.; Brockman, W.; Kim, T.K.; Koche, R.P.; *et al.* Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. *Nature* **2007**, *448*, 553–560. [[CrossRef](#)]
18. Pan, G.; Tian, S.; Nie, J.; Yang, C.; Ruotti, V.; Wei, H.; Jonsdottir, G.A.; Stewart, R.; Thomson, J.A. Whole-genome analysis of histone H3 lysine 4 and lysine 27 methylation in human embryonic stem cells. *Cell Stem Cell* **2007**, *1*, 299–312. [[CrossRef](#)]
19. Zhao, X.D.; Han, X.; Chew, J.L.; Liu, J.; Chiu, K.P.; Choo, A.; Orlov, Y.L.; Sung, W.K.; Shahab, A.; Kuznetsov, V.A.; *et al.* Wholegenome mapping of histone H3 Lys4 and 27 trimethylations reveals distinct genomic compartments in human embryonic stem cells. *Cell Stem Cell* **2007**, *1*, 286–298. [[CrossRef](#)]
20. Margueron, R.; Li, G.; Sarma, K.; Blais, A.; Zavadil, J.; Woodcock, C.L.; Dynlacht, B.D.; Reinberg, D. Ezh1 and Ezh2 maintain repressive chromatin through different mechanisms. *Mol. Cell* **2008**, *32*, 503–518. [[CrossRef](#)]
21. Kouznetsova, V.L.; Tchekanov, A.; Li, X.; Yan, X.; Tsigelny, I.F. Polycomb repressive 2 complex-Molecular mechanisms of function. *Protein Sci.* **2019**, *28*, 1387–1399. [[CrossRef](#)]
22. Van Mierlo, G.; Veenstra, G.J.C.; Vermeulen, M.; Marks, H. The Complexity of PRC2 Subcomplexes. *Trends Cell Biol.* **2019**, *29*, 660–671. [[CrossRef](#)]
23. Chammas, P.; Mocavini, I.; Di Croce, L. Engaging chromatin: PRC2 structure meets function. *Br. J. Cancer* **2020**, *122*, 315–328. [[CrossRef](#)]
24. Pasini, D.; Bracken, A.P.; Jensen, M.R.; Lazzarini Denchi, E.; Helin, K. Suz12 is essential for mouse development and for EZH2 histone methyltransferase activity. *EMBO J.* **2004**, *23*, 4061–4071. [[CrossRef](#)]

25. Faust, C.; Schumacher, A.; Holdener, B.; Magnuson, T. The eed mutation disrupts anterior mesoderm production in mice. *Development* **1995**, *121*, 273–285. [[CrossRef](#)]
26. Ueda, T.; Sanada, M.; Matsui, H.; Yamasaki, N.; Honda, Z.I.; Shih, L.Y.; Mori, H.; Inaba, T.; Ogawa, S.; Honda, H. EED mutants impair polycomb repressive complex 2 in myelodysplastic syndrome and related neoplasms. *Leukemia* **2012**, *26*, 2557–2560. [[CrossRef](#)] [[PubMed](#)]
27. De Raedt, T.; Beert, E.; Pasmant, E.; Luscan, A.; Brems, H.; Ortonne, N.; Helin, K.; Hornick, J.L.; Mautner, V.; Kehrer-Sawatzki, H.; *et al.* PRC2 loss amplifies Ras-driven transcription and confers sensitivity to BRD4-based therapies. *Nature* **2014**, *514*, 247–251. [[CrossRef](#)] [[PubMed](#)]
28. Lee, W.; Teckie, S.; Wiesner, T.; Ran, L.; Prieto Granada, C.N.; Lin, M.; Zhu, S.; Cao, Z.; Liang, Y.; Sboner, A. PRC2 is recurrently inactivated through EED or SUZ12 loss in malignant peripheral nerve sheath tumors. *Nat. Genet.* **2014**, *46*, 1227–1232. [[CrossRef](#)] [[PubMed](#)]
29. Erokhin, M.; Chetverina, O.; Gyorffy, B.; Tatarskiy, V.V.; Mogila, V.; Shtil, A.A.; Roninson, I.B.; Moreaux, J.; Georgiev, P.; Cavalli, G.; *et al.* Clinical Correlations of Polycomb Repressive Complex 2 in Different Tumor Types. *Cancers* **2021**, *13*, 3155. [[CrossRef](#)]
30. Cohen, A.S.; Tuysuz, B.; Shen, Y.; Bhalla, S.K.; Jones, S.J.; Gibson, W.T. A novel mutation in EED associated with overgrowth. *J. Hum. Genet.* **2015**, *60*, 339–342. [[CrossRef](#)] [[PubMed](#)]
31. Imagawa, E.; Higashimoto, K.; Sakai, Y.; Numakura, C.; Okamoto, N.; Matsunaga, S.; Ryo, A.; Sato, Y.; Sanefuji, M.; Ihara, K.; *et al.* Mutations in genes encoding polycomb repressive complex 2 subunits cause Weaver syndrome. *Hum. Mutat.* **2017**, *38*, 637–648. [[CrossRef](#)]
32. Cooney, E.; Bi, W.; Schlesinger, A.E.; Vinson, S.; Potocki, L. Novel EED mutation in patient with Weaver syndrome. *Am. J. Med. Genet. A* **2017**, *173*, 541–545. [[CrossRef](#)]
33. Griffiths, S.; Loveday, C.; Zachariou, A.; Behan, L.A.; Chandler, K.; Cole, T.; D'Arrigo, S.; Dieckmann, A.; Foster, A.; Gibney, J.; *et al.* EED and EZH2 constitutive variants: A study to expand the Cohen-Gibson syndrome phenotype and contrast it with Weaver syndrome. *Am. J. Med. Genet. A* **2019**, *179*, 588–594. [[CrossRef](#)]
34. Spellicy, C.J.; Peng, Y.; Olewiler, L.; Cathey, S.S.; Rogers, R.C.; Bartholomew, D.; Johnson, J.; Alexov, E.; Lee, J.A.; Friez, M.J.; *et al.* Three additional patients with EED-associated overgrowth: Potential mutation hotspots identified? *J. Hum. Genet.* **2019**, *64*, 561–572. [[CrossRef](#)] [[PubMed](#)]
35. Völkel, P.; Bary, A.; Raby, L.; Chapart, A.; Dupret, B.; Le Bourhis, X.; Angrand, P.O. Ezh1 arises from Ezh2 gene duplication but its function is not required for zebrafish development. *Sci. Rep.* **2019**, *9*, 4319. [[CrossRef](#)]
36. Soto, R.A.; Najia, M.A.T.; Hachimi, M.; Frame, J.M.; Yette, G.A.; Lummertz da Rocha, E.; Stankunas, K.; Daley, G.Q.; North, T.E. Sequential regulation of hemogenic fate and hematopoietic stem and progenitor cell formation from arterial endothelium by Ezh1/2. *Stem Cell Rep.* **2021**, *16*, 1718–1734. [[CrossRef](#)]
37. San, B.; Chrispijn, N.D.; Wittkopp, N.; van Heeringen, S.J.; Lagendijk, A.K.; Aben, M.; Bakkers, J.; Ketting, R.F.; Kammaing, L.M. Normal formation of a vertebrate body plan and loss of tissue maintenance in the absence of ezh2. *Sci. Rep.* **2016**, *6*, 24658. [[CrossRef](#)] [[PubMed](#)]
38. Dupret, B.; Völkel, P.; Vennin, C.; Toillon, R.A.; Le Bourhis, X.; Angrand, P.O. The histone lysine methyltransferase Ezh2 is required for maintenance of the intestine integrity and for caudal fin regeneration in zebrafish. *Biochim. Biophys. Acta Gene Regul. Mech.* **2017**, *1860*, 1079–1093. [[CrossRef](#)]
39. Oppel, F.; Ki, D.H.; Zimmerman, M.W.; Ross, K.N.; Tao, T.; Shi, H.; He, S.; Aster, J.C.; Look, A.T. *suz12* inactivation in *p53*- and *nfl*-deficient zebrafish accelerates the onset of malignant peripheral nerve sheath tumors and expands the spectrum of tumor types. *Dis. Model Mech.* **2020**, *13*, dmm042341. [[CrossRef](#)] [[PubMed](#)]
40. Maître, J.L.; Berthoumieux, H.; Krens, S.F.; Salbreux, G.; Jülicher, F.; Paluch, E.; Heisenberg, C.P. Adhesion functions in cell sorting by mechanically coupling the cortices of adhering cells. *Science* **2012**, *338*, 253–256. [[CrossRef](#)]
41. Kimmel, C.B.; Ballard, W.W.; Kimmel, S.R.; Ullmann, B.; Schilling, T.F. Stages of embryonic development of the zebrafish. *Dev. Dyn.* **1995**, *203*, 253–310. [[CrossRef](#)] [[PubMed](#)]
42. Dupret, B.; Völkel, P.; Le Bourhis, X.; Angrand, P.O. The Polycomb Group Protein Pcgf1 Is Dispensable in Zebrafish but Involved in Early Growth and Aging. *PLoS ONE* **2016**, *11*, e0158700. [[CrossRef](#)]
43. Doyle, E.L.; Booher, N.J.; Standage, D.S.; Voytas, D.F.; Brendel, V.P.; Vandyk, J.K.; Bogdanove, A.J. TAL Effector-Nucleotide Targeter (TALEN-NT) 2.0: Tools for TAL effector design and target prediction. *Nucleic Acids Res.* **2012**, *40*, W117–W122. [[CrossRef](#)] [[PubMed](#)]
44. Cermak, T.; Doyle, E.L.; Christian, M.; Wang, L.; Zhang, Y.; Schmidt, C.; Baller, J.A.; Somia, N.V.; Bogdanove, A.J.; Voytas, D.F. Efficient design and assembly of custom TALEN and other TAL effector-based constructs for DNA targeting. *Nucleic Acids Res.* **2011**, *39*, e82. [[CrossRef](#)] [[PubMed](#)]
45. Dahlem, T.J.; Hoshijima, K.; Jurynech, M.J.; Gunther, D.; Starker, C.G.; Locke, A.S.; Weis, A.M.; Voytas, D.F.; Grunwald, D.J. Simple methods for generating and detecting locus-specific mutations induced with TALENs in the zebrafish genome. *PLoS Genet.* **2012**, *8*, e1002861. [[CrossRef](#)]

46. Dupret, B.; Völkel, P.; Follet, P.; Le Bourhis, X.; Angrand, P.O. Combining genotypic and phenotypic analyses on single mutant zebrafish larvae. *MethodsX* **2018**, *5*, 244–256. [[CrossRef](#)]
47. Dambroise, E.; Monnier, L.; Ruisheng, L.; Aguilaniu, H.; Joly, J.S.; Tricoire, H.; Rera, M. Two phases of aging separated by the Smurf transition as a public path to death. *Sci. Rep.* **2016**, *6*, 23523. [[CrossRef](#)] [[PubMed](#)]
48. Parichy, D.M.; Elizondo, M.R.; Mills, M.G.; Gordon, T.N.; Engeszer, R.E. Normal table of postembryonic zebrafish development: Staging by externally visible anatomy of the living fish. *Dev. Dyn.* **2009**, *238*, 2975–3015. [[CrossRef](#)]
49. Thisse, C.; Thisse, B. High-resolution in situ hybridization to whole-mount zebrafish embryos. *Nat. Protoc.* **2008**, *3*, 59–69. [[CrossRef](#)]
50. Raby, L.; Völkel, P.; Le Bourhis, X.; Angrand, P.O. The Polycomb Orthologues in Teleost Fishes and Their Expression in the Zebrafish Model. *Genes* **2020**, *11*, 362. [[CrossRef](#)]
51. Han, Z.; Xing, X.; Hu, M.; Zhang, Y.; Liu, P.; Chai, J. Structural basis of EZH2 recognition by EED. *Structure* **2007**, *15*, 1306–1315. [[CrossRef](#)]
52. Pelegri, F. Maternal factors in zebrafish development. *Dev. Dyn.* **2003**, *228*, 535–554. [[CrossRef](#)]
53. Dosch, R.; Wagner, D.S.; Mintzer, K.A.; Runke, G.; Wiemelt, A.P.; Mullins, M.C. Maternal control of vertebrate development before the midblastula transition: Mutants from the zebrafish I. *Dev. Cell* **2004**, *6*, 771–780. [[CrossRef](#)] [[PubMed](#)]
54. Chrispijn, N.D.; Andralojc, K.M.; Castenmiller, C.; Kamminga, L.M. Gene expression profile of a selection of Polycomb Group genes during zebrafish embryonic and germ line development. *PLoS ONE* **2018**, *13*, e0200316. [[CrossRef](#)] [[PubMed](#)]
55. White, R.J.; Collins, J.E.; Sealy, I.M.; Wali, N.; Dooley, C.M.; Digby, Z.; Stemple, D.L.; Murphy, D.N.; Billis, K.; Hourlier, T.; et al. A high-resolution mRNA expression time course of embryonic development in zebrafish. *eLife* **2017**, *6*, e30860. [[CrossRef](#)] [[PubMed](#)]
56. Tachibana, M.; Ueda, J.; Fukuda, M.; Takeda, N.; Ohta, T.; Iwanari, H.; Sakihama, T.; Kodama, T.; Hamakubo, T.; Shinkai, Y. Histone methyltransferases G9a and GLP form heteromeric complexes and are both crucial for methylation of euchromatin at H3-K9. *Genes Dev.* **2005**, *19*, 815–826. [[CrossRef](#)] [[PubMed](#)]
57. Wu, H.; Chen, X.; Xiong, J.; Li, Y.; Li, H.; Ding, X.; Liu, S.; Chen, S.; Gao, S.; Zhu, B. Histone methyltransferase G9a contributes to H3K27 methylation in vivo. *Cell Res.* **2011**, *21*, 365–367. [[CrossRef](#)]
58. Lavarone, E.; Barbieri, C.M.; Pasini, D. Dissecting the role of H3K27 acetylation and methylation in PRC2 mediated control of cellular identity. *Nat. Commun.* **2019**, *10*, 1679. [[CrossRef](#)]
59. Piunti, A.; Hashizume, R.; Morgan, M.A.; Bartom, E.T.; Horbinski, C.M.; Marshall, S.A.; Rendleman, E.J.; Ma, Q.; Takahashi, Y.H.; Woodfin, A.R.; et al. Therapeutic targeting of polycomb and BET bromodomain proteins in diffuse intrinsic pontine gliomas. *Nat. Med.* **2017**, *23*, 493–500. [[CrossRef](#)]
60. Gao, Z.; Zhang, J.; Bonasio, R.; Strino, F.; Sawai, A.; Parisi, F.; Kluger, Y.; Reinberg, D. PCGF homologs, CBX proteins, and RYBP define functionally distinct PRC1 family complexes. *Mol. Cell* **2012**, *45*, 344–356. [[CrossRef](#)]
61. Tavares, L.; Dimitrova, E.; Oxley, D.; Webster, J.; Poot, R.; Demmers, J.; Bezstarosti, K.; Taylor, S.; Ura, H.; Koide, H.; et al. RYBP-PRC1 complexes mediate H2A ubiquitylation at polycomb target sites independently of PRC2 and H3K27me3. *Cell* **2012**, *148*, 664–678. [[CrossRef](#)]
62. Morey, L.; Aloia, L.; Cozzuto, L.; Benitah, S.A.; Di Croce, L. RYBP and Cbx7 define specific biological functions of polycomb complexes in mouse embryonic stem cells. *Cell Rep.* **2013**, *3*, 60–69. [[CrossRef](#)]
63. San, B.; Aben, M.; Elurbe, D.M.; Voeltzke, K.; Den Broeder, M.J.; Rougeot, J.; Legler, J.; Kamminga, L.M. Genetic and Epigenetic Regulation of Zebrafish Intestinal Development. *Epigenomes* **2018**, *2*, 19. [[CrossRef](#)]
64. Behrndt, M.; Salbreux, G.; Campinho, P.; Hauschild, R.; Oswald, F.; Roensch, J.; Grill, S.W.; Heisenberg, C.P. Forces driving epithelial spreading in zebrafish gastrulation. *Science* **2012**, *338*, 257–260. [[CrossRef](#)]
65. Araya, C.; Häkkinen, H.M.; Carcamo, L.; Cerda, M.; Savy, T.; Rookyard, C.; Peyriéras, N.; Clarke, J.D.W. Cdh2 coordinates Myosin-II dependent internalisation of the zebrafish neural plate. *Sci. Rep.* **2019**, *9*, 1835. [[CrossRef](#)]
66. Dai, W.; Wang, K.; Zheng, X.; Chen, X.; Zhang, W.; Zhang, Y.; Hou, J.; Liu, L. High fat plus high cholesterol diet lead to hepatic steatosis in zebrafish larvae: A novel model for screening anti-hepatic steatosis drugs. *Nutr. Metab.* **2015**, *12*, 42. [[CrossRef](#)] [[PubMed](#)]
67. Chen, B.; Zheng, Y.M.; Zhang, J.P. Comparative Study of Different Diets-Induced NAFLD Models of Zebrafish. *Front. Endocrinol.* **2018**, *9*, 366. [[CrossRef](#)] [[PubMed](#)]
68. Howarth, D.L.; Lindtner, C.; Vacaru, A.M.; Sachidanandam, R.; Tsedensodnom, O.; Vasilkova, T.; Buettner, C.; Sadler, K.C. Activating transcription factor 6 is necessary and sufficient for alcoholic fatty liver disease in zebrafish. *PLoS Genet.* **2014**, *10*, e1004335. [[CrossRef](#)] [[PubMed](#)]
69. Yin, G.; Cao, L.; Du, J.; Jia, R.; Kitazawa, T.; Kubota, A.; Teraoka, H. Dexamethasone-induced hepatomegaly and steatosis in larval zebrafish. *J. Toxicol. Sci.* **2017**, *42*, 455–459. [[CrossRef](#)]
70. Katoch, S.; Patial, V. Zebrafish: An emerging model system to study liver diseases and related drug discovery. *J. Appl. Toxicol.* **2021**, *41*, 33–51. [[CrossRef](#)]



71. den Broeder, M.J.; Ballangby, J.; Kamminga, L.M.; Aleström, P.; Legler, J.; Lindeman, L.C.; Kamstra, J.H. Inhibition of methyltransferase activity of enhancer of zeste 2 leads to enhanced lipid accumulation and altered chromatin status in zebrafish. *Epigenet. Chromatin* **2020**, *13*, 5. [[CrossRef](#)]
72. Alunni, A.; Hermel, J.M.; Heuzé, A.; Bourrat, F.; Jamen, F.; Joly, J.S. Evidence for neural stem cells in the medaka optic tectum proliferation zones. *Dev. Neurobiol.* **2010**, *70*, 693–713. [[CrossRef](#)] [[PubMed](#)]
73. MacPhail, R.C.; Brooks, J.; Hunter, D.L.; Padnos, B.; Irons, T.D.; Padilla, S. Locomotion in larval zebrafish: Influence of time of day, lighting and ethanol. *Neurotoxicology* **2009**, *30*, 52–58. [[CrossRef](#)]
74. Schnörr, S.J.; Steenbergen, P.J.; Richardson, M.K.; Champagne, D.L. Measuring thigmotaxis in larval zebrafish. *Behav. Brain Res.* **2012**, *228*, 367–374. [[CrossRef](#)] [[PubMed](#)]
75. Douek, A.M.; Amiri Khabooshan, M.; Henry, J.; Stamatis, S.A.; Kreuder, F.; Ramm, G.; Änkö, M.L.; Wlodkowic, D.; Kaslin, J. An Engineered *sgsh* Mutant Zebrafish Recapitulates Molecular and Behavioural Pathobiology of Sanfilippo Syndrome A/MPS IIIA. *Int. J. Mol. Sci.* **2021**, *22*, 5948. [[CrossRef](#)] [[PubMed](#)]
76. Koppens, M.A.; Bounova, G.; Gargiulo, G.; Tanger, E.; Janssen, H.; Cornelissen-Steijger, P.; Blom, M.; Song, J.Y.; Wessels, L.F.; van Lohuizen, M. Deletion of Polycomb Repressive Complex 2 from Mouse Intestine Causes Loss of Stem Cells. *Gastroenterology* **2016**, *151*, 684–697. [[CrossRef](#)]
77. Lucore, E.C.; Connaughton, V.P. Observational learning and irreversible starvation in first-feeding zebrafish larvae: Is it okay to copy from your friends? *Zoology* **2021**, *145*, 125896. [[CrossRef](#)]
78. Shen, X.; Liu, Y.; Hsu, Y.J.; Fujiwara, Y.; Kim, J.; Mao, X.; Yuan, G.C.; Orkin, S.H. EZH1 mediates methylation on histone H3 lysine 27 and complements EZH2 in maintaining stem cell identity and executing pluripotency. *Mol. Cell* **2008**, *32*, 491–502. [[CrossRef](#)]
79. Lee, C.H.; Holder, M.; Grau, D.; Saldaña-Meyer, R.; Yu, J.R.; Ganai, R.A.; Zhang, J.; Wang, M.; LeRoy, G.; Dobenecker, M.W.; et al. Distinct Stimulatory Mechanisms Regulate the Catalytic Activity of Polycomb Repressive Complex 2. *Mol. Cell* **2018**, *70*, 435–448. [[CrossRef](#)]
80. Cao, R.; Zhang, Y. SUZ12 is required for both the histone methyltransferase activity and the silencing function of the EED-EZH2 complex. *Mol. Cell* **2004**, *15*, 57–67. [[CrossRef](#)]
81. Ferrari, K.J.; Scelfo, A.; Jammula, S.; Cuomo, A.; Barozzi, I.; Stützer, A.; Fischle, W.; Bonaldi, T.; Pasini, D. Polycomb-dependent H3K27me1 and H3K27me2 regulate active transcription and enhancer fidelity. *Mol. Cell* **2014**, *53*, 49–62. [[CrossRef](#)]
82. Tachibana, M.; Sugimoto, K.; Fukushima, T.; Shinkai, Y. Set domain-containing protein, G9a, is a novel lysine-preferring mammalian histone methyltransferase with hyperactivity and specific selectivity to lysines 9 and 27 of histone H3. *J. Biol. Chem.* **2001**, *276*, 25309–25317. [[CrossRef](#)] [[PubMed](#)]
83. Aranda, S.; Mas, G.; Di Croce, L. Regulation of gene transcription by Polycomb proteins. *Sci. Adv.* **2015**, *1*, e1500737. [[CrossRef](#)]
84. Le Faou, P.; Völkel, P.; Angrand, P.O. The zebrafish genes encoding the Polycomb repressive complex (PRC) 1. *Gene* **2011**, *475*, 10–21. [[CrossRef](#)]
85. Van der Velden, Y.U.; Wang, L.; van Lohuizen, M.; Haramis, A.P. The Polycomb group protein Ring1b is essential for pectoral fin development. *Development* **2012**, *139*, 2210–2220. [[CrossRef](#)] [[PubMed](#)]
86. Van der Velden, Y.U.; Wang, L.; Querol Cano, L.; Haramis, A.P. The polycomb group protein ring1b/rnf2 is specifically required for craniofacial development. *PLoS ONE* **2013**, *8*, e73997. [[CrossRef](#)]
87. Rougeot, J.; Chrispijn, N.D.; Aben, M.; Elurbe, D.M.; Andralojc, K.M.; Murphy, P.J.; Jansen, P.W.T.C.; Vermeulen, M.; Cairns, B.R.; Kamminga, L.M. Maintenance of spatial gene expression by Polycomb-mediated repression after formation of a vertebrate body plan. *Development* **2019**, *146*, dev178590. [[CrossRef](#)]
88. Zhao, J.; Wang, M.; Chang, L.; Yu, J.; Song, A.; Liu, C.; Huang, W.; Zhang, T.; Wu, X.; Shen, X.; et al. RYBP/YAF2-PRC1 complexes and histone H1-dependent chromatin compaction mediate propagation of H2AK119ub1 during cell division. *Nat. Cell Biol.* **2020**, *22*, 439–452. [[CrossRef](#)] [[PubMed](#)]
89. Schwartzenuber, J.; Korshunov, A.; Liu, X.Y.; Jones, D.T.; Pfaff, E.; Jacob, K.; Sturm, D.; Fontebasso, A.M.; Quang, D.A.; Tönjes, M.; et al. Driver mutations in histone H3.3 and chromatin remodelling genes in paediatric glioblastoma. *Nature* **2012**, *482*, 226–231. [[CrossRef](#)]
90. Wu, G.; Broniscer, A.; McEachron, T.A.; Lu, C.; Paugh, B.S.; Becksfors, J.; Qu, C.; Ding, L.; Huether, R.; Parker, M.; et al. Somatic histone H3 alterations in pediatric diffuse intrinsic pontine gliomas and non-brainstem glioblastomas. *Nat. Genet.* **2012**, *44*, 251–253. [[CrossRef](#)]
91. Fontebasso, A.M.; Papillon-Cavanagh, S.; Schwartzenuber, J.; Nikbakht, H.; Gerges, N.; Fiset, P.O.; Bechet, D.; Faury, D.; De Jay, N.; Ramkissoon, L.A.; et al. Recurrent somatic mutations in ACVR1 in pediatric midline high-grade astrocytoma. *Nat. Genet.* **2014**, *46*, 462–466. [[CrossRef](#)]

92. Castel, D.; Philippe, C.; Calmon, R.; Le Dret, L.; Truffaux, N.; Boddaert, N.; Pagès, M.; Taylor, K.R.; Saulnier, P.; Lacroix, L.; *et al.* Histone H3F3A and HIST1H3B K27M mutations define two subgroups of diffuse intrinsic pontine gliomas with different prognosis and phenotypes. *Acta Neuropathol.* **2015**, *130*, 815–827. [[CrossRef](#)]
93. Bender, S.; Tang, Y.; Lindroth, A.M.; Hovestadt, V.; Jones, D.T.; Kool, M.; Zapatka, M.; Northcott, P.A.; Sturm, D.; Wang, W.; *et al.* Reduced H3K27me3 and DNA hypomethylation are major drivers of gene expression in K27M mutant pediatric high-grade gliomas. *Cancer Cell* **2013**, *24*, 660–672. [[CrossRef](#)]
94. Chan, K.M.; Fang, D.; Gan, H.; Hashizume, R.; Yu, C.; Schroeder, M.; Gupta, N.; Mueller, S.; James, C.D.; Jenkins, R.; *et al.* The histone H3.3K27M mutation in pediatric glioma reprograms H3K27 methylation and gene expression. *Genes Dev.* **2013**, *27*, 985–990. [[CrossRef](#)] [[PubMed](#)]
95. Herz, H.M.; Morgan, M.; Gao, X.; Jackson, J.; Rickels, R.; Swanson, S.K.; Florens, L.; Washburn, M.P.; Eissenberg, J.C.; Shilatifard, A. Histone H3 lysine-to-methionine mutants as a paradigm to study chromatin signaling. *Science* **2014**, *345*, 1065–1070. [[CrossRef](#)]
96. Lewis, P.W.; Müller, M.M.; Koletsky, M.S.; Cordero, F.; Lin, S.; Banaszynski, L.A.; Garcia, B.A.; Muir, T.W.; Becher, O.J.; Allis, C.D. Inhibition of PRC2 activity by a gain-of-function H3 mutation found in pediatric glioblastoma. *Science* **2013**, *340*, 857–861. [[CrossRef](#)] [[PubMed](#)]
97. Veneti, S.; Garimella, M.T.; Sullivan, L.M.; Martinez, D.; Huse, J.T.; Heguy, A.; Santi, M.; Thompson, C.B.; Judkins, A.R. Evaluation of histone 3 lysine 27 trimethylation (H3K27me3) and enhancer of Zest 2 (EZH2) in pediatric glial and glioneuronal tumors shows decreased H3K27me3 in H3F3A K27M mutant glioblastomas. *Brain Pathol.* **2013**, *23*, 558–564. [[CrossRef](#)]
98. Khuong-Quang, D.A.; Buczkowicz, P.; Rakopoulos, P.; Liu, X.Y.; Fontebasso, A.M.; Bouffet, E.; Bartels, U.; Albrecht, S.; Schwartzentruber, J.; Letourneau, L.; *et al.* K27M mutation in histone H3.3 defines clinically and biologically distinct subgroups of pediatric diffuse intrinsic pontine gliomas. *Acta Neuropathol.* **2012**, *124*, 439–447. [[CrossRef](#)] [[PubMed](#)]
99. Mackay, A.; Burford, A.; Carvalho, D.; Izquierdo, E.; Fazal-Salom, J.; Taylor, K.R.; Bjerke, L.; Clarke, M.; Vinci, M.; Nandhabalan, M.; *et al.* Integrated Molecular Meta-Analysis of 1,000 Pediatric High-Grade and Diffuse Intrinsic Pontine Glioma. *Cancer Cell* **2017**, *32*, 520–537. [[CrossRef](#)]
100. Ballester, L.Y.; Wang, Z.; Shandilya, S.; Miettinen, M.; Burger, P.C.; Eberhart, C.G.; Rodriguez, F.J.; Raabe, E.; Nazarian, J.; Warren, K.; *et al.* Morphologic characteristics and immunohistochemical profile of diffuse intrinsic pontine gliomas. *Am. J. Surg. Pathol.* **2013**, *37*, 1357–1364. [[CrossRef](#)]
101. Kumar, S.S.; Sengupta, S.; Lee, K.; Hura, N.; Fuller, C.; DeWire, M.; Stevenson, C.B.; Fouladi, M.; Drissi, R. BMI-1 is a potential therapeutic target in diffuse intrinsic pontine glioma. *Oncotarget* **2017**, *8*, 62962–62975. [[CrossRef](#)] [[PubMed](#)]
102. Balakrishnan, I.; Danis, E.; Pierce, A.; Madhavan, K.; Wang, D.; Dahl, N.; Sanford, B.; Birks, D.K.; Davidson, N.; Metselaar, D.S.; *et al.* Senescence Induced by BMI1 Inhibition Is a Therapeutic Vulnerability in H3K27M-Mutant DIPG. *Cell Rep.* **2020**, *33*, 108286. [[CrossRef](#)] [[PubMed](#)]
103. Cohen, A.S.; Gibson, W.T. EED-associated overgrowth in a second male patient. *J. Hum. Genet.* **2016**, *61*, 831–834. [[CrossRef](#)] [[PubMed](#)]
104. Sequerra Amram Cohen, A.; Gibson, W.T. EED-Related Overgrowth. In *GeneReviews® [Internet]*; Adam, M.P., Ardinger, H.H., Pagon, R.A., Wallace, S.E., Bean, L.J.H., Mirzaa, G., Amemiya, A., Eds.; University of Washington: Seattle, WA, USA, 2019; pp. 1993–2021.

### 3.3 Conclusion

L'étude de l'expression du gène *eed* au cours du développement suggère un rôle prépondérant dans ce processus. Le transcrit est déposé maternellement dans les œufs. Dans les stades embryonnaires les plus précoces, l'expression de *eed* est ubiquitaire, tandis qu'au fur et à mesure du développement, le profil d'expression est restreint au niveau de zones de proliférations cellulaires, de la même façon que pour les gènes *ezh2* ou *ezh1* (Dupret *et al.*, 2017 ; Völkel *et al.*, 2019). Le complexe PRC2 est donc impliqué dans le développement précoce des poissons zèbres mais pas indispensable, puisque les larves mutantes survivent jusqu'au tour de 11 jpf. Tous les organes sont mis en place, et aucun phénotype homéotique n'est visible précocement mais des altérations des organes digestifs des larves mutantes apparaissent au cours du développement et conduisent à leur létalité précoce. La protéine Eed est indispensable à la constitution du complexe PRC2. Ainsi, l'état de méthylation de la lysine 27 de l'histone H3 a été quantifié. Les di- et triméthylations ont largement diminué, ce qui indique que le complexe PRC2 est non fonctionnel. Néanmoins la mono-méthylation de la lysine 27 ne semble pas être réduite chez les larves *eed*<sup>-/-</sup>. Chez les mammifères, plusieurs études ont montré la capacité des histones méthyl-transférases G9a et Glp à méthyler la lysine 9 de l'histone H3 ainsi que la lysine 27 de l'histone H3 (Tachibana *et al.*, 2005 ; Wu *et al.*, 2011). Le poisson zèbre possède deux orthologues pour de G9a (Ehmt1a/b) et un orthologue de Glp (Ehmt2), qui pourrait être à l'origine de la réalisation de la monométhylation de la lysine 27 de l'histone H3. D'autres marques épigénétiques ont également été étudié. De façon surprenante, la déposition de la marque H2AK119ub1 ne semble pas varier. La marque répressive H3K27ac augmente tandis que la marque répressive H4K20me3 diminue, ce qui indique une dérégulation transcriptionnelle. La modification de l'environnement chromatinien chez les larves reflète les résultats obtenus sur plusieurs lignées cellulaires cancéreuses humaines (Ueda *et al.*, 2016 ; Krug *et al.*, 2019). Notre modèle d'inactivation de PRC2 chez le poisson zèbre résume les modifications épigénétiques observées dans les cancers, et en particulier dans le cas des gliomes pédiatriques diffus de haut grade, dans lesquels une mutation de la lysine 27 de l'histone H3 inhibe l'activité du PRC2. Cette mutation dans les cellules cancéreuses est associée avec des cellules très peu différenciées et hautement résistantes aux thérapies existantes. Le profil d'expression de marqueurs de différenciation a été étudié au niveau neural chez les larves mutantes, et montrent que la perte de fonction de PRC2 altère la différenciation des cellules neurales et maintient certaines cellules neurales aux stades progénitrices. Ces altérations du développement neural sont corrélées avec des altérations comportementales.



# 1. La transplantation de cellules tumorales chez le poisson zèbre : de la recherche translationnelle à la médecine personnalisée

## 1.1 Contexte scientifique

En France, le cancer reste la principale cause de mortalité, en particulier chez les hommes. En 2018, 157400 personnes sont décédées d'un cancer en France, et 57 % d'entre elles étaient des hommes (*Panorama des cancers*, INCA, 2022). Le taux d'incidence des cancers ne semble pas varier, mais le taux de mortalité sur la période 2010-2018 a légèrement diminué. Ces chiffres encourageants sont le résultat de campagnes de prévention et de dépistage qui permettent la détection du cancer dans les stades les plus précoces. De nouvelles stratégies de lutte contre le cancer ont été développées et donnent des résultats prometteurs. Néanmoins, les cancers à mauvais pronostics comme le cancer du pancréas, les cancers du système nerveux central, ou les cancers métastasés demeurent sans solutions thérapeutiques efficaces. La culture cellulaire reste une méthode de choix pour l'étude de certains types de cellules cancéreuses. Le processus métastatique est en revanche un processus complexe qui comprend la migration des cellules cancéreuses, leur intravasation et leur extravasation du système vasculaire, ainsi que la colonisation d'un nouvel environnement. Ces étapes sont associées à des changements morphologiques et moléculaires des cellules cancéreuses, notamment avec la transition épithélio-mésenchymateuse (Fares *et al.*, 2020). L'étude de ces phénomènes nécessite donc un système complexe et malgré les progrès récents des systèmes micro-fluidiques, les modèles animaux restent indispensables en recherche. Le poisson zèbre est historiquement un modèle de recherche pour la biologie du développement. Il a été révélé très rapidement que les poissons zèbres pourraient être un modèle relevant pour étudier la carcinogénèse (Pliss *et al.*, 1982). Grâce aux nombreuses lignées mutantes disponibles, le faible coût d'élevage ou le temps de génération et d'expérimentation court, le poisson zèbre est également devenu un modèle de choix dans les expériences de transplantations de cellules cancéreuses (White *et al.*, 2013).

## 1.2 Article

# La transplantation de cellules tumorales chez le poisson zèbre : de la recherche translationnelle à la médecine personnalisée

Ludivine Raby <sup>1,2</sup>, Pamela Völkel <sup>1,2,3</sup>, Xuefen Le Bourhis <sup>1,2</sup>, Pierre-Olivier Angrand <sup>1,2</sup>

1. Inserm U908, Cell Plasticity & Cancer, 59000 Lille, France, University of Lille, 59000 Lille, France

2. CNRS, 59000 Lille, France

Reçu le 2 avril 2019

Accepté le 5 juin 2019

### Résumé

D'abord utilisé pour l'étude du développement, le modèle poisson zèbre prend maintenant une importance capitale dans la lutte contre le cancer. La transplantation de cellules cancéreuses chez le poisson zèbre est un outil clé dans la recherche préclinique car elle permet d'étudier les événements cellulaires et moléculaires responsables du développement tumoral et permet d'évaluer l'action de molécules thérapeutiques *in vivo*. Appliquée aux cellules cancéreuses dérivées de patients, la xénotransplantation permet d'identifier les protocoles thérapeutiques les plus efficaces dans le cadre de la médecine personnalisée. Cette revue fait le point sur les apports de la transplantation de cellules tumorales chez le poisson zèbre dans la lutte contre le cancer.

### Summary

Cancer cell transplantation in zebrafish: From translational research to personalized medicine

Primarily used in genetic studies of development, the zebrafish (*Danio rerio*) has rapidly emerged as a promising animal model of human cancer. Cancer cell transplantation in zebrafish constitutes a key platform for clinical research since it allows to study cellular and molecular events involved in various aspects of tumorigenesis and to evaluate the efficacy of therapeutic molecules *in vivo*. Applied to patient-derived cells, the xenotransplantation approach in zebrafish allows to define the most appropriate therapeutic strategies for specific alterations found in patients in the context of personalized medicine. This review discusses the zebrafish transplantation model for the study of cancer development and drug discovery.

### Introduction

Le cancer est actuellement une des causes majeures de morbidité et de mortalité chez l'homme. Selon l'Organisation mondiale de la santé, le nombre de nouveaux cas de cancers est estimé à 18,1 millions en 2018 alors que le nombre de décès par cancers était de 9,6 millions [1]. C'est en général la capacité des cellules tumorales à se propager dans l'organisme et coloniser des tissus distants aboutissant à la formation de métastases, qui va déterminer l'évolution et l'issue de la maladie. La métastase est un processus complexe qui implique le détachement de la tumeur primaire, l'invasion du tissu environnant, l'intravasation dans les vaisseaux sanguins et lymphatiques, l'extravasation dans des tissus secondaires déterminés, l'enracinement dans un nouvel environnement cellulaire puis la prolifération tumorale. Ainsi, la compréhension des mécanismes moléculaires et cellulaires responsables de la formation des métastases est essentielle à la lutte contre la progression tumorale [2]. Dans ce contexte, la transplantation de cellules cancéreuses est un outil de choix pour la recherche préclinique en cancérologie car elle permet d'étudier de nombreux événements cellulaires responsables du développement tumoral et de sa dissémination *in vivo*. Historiquement, la souris est l'organisme modèle le plus utilisé dans les expériences de xénotransplantation de cellules cancéreuses humaines [3], mais le poisson zèbre (*Danio rerio*) apparaît maintenant également comme un modèle expérimental particulièrement attractif.

Notamment, (i) le poisson zèbre possède une large descendance, environ 200 œufs par femelle et par semaine, qui offre une grande confiance dans les analyses statistiques ; (ii) l'homme et le poisson zèbre partagent de grandes similarités puisque 70 % des protéines humaines et 82 % des protéines impliquées dans des pathologies humaines ont un orthologue chez le poisson zèbre ; (iii) de nombreux processus cellulaires peuvent être observés chez le poisson zèbre en raison de la transparence des embryons ou chez des poissons adultes dépourvus de pigmentation ; (iv) les techniques de génétique inverse telles que la transgénèse ou l'édition du génome par les TALEN et le système CRISPR/Cas9 sont facilement applicables au poisson zèbre ; (v) le poisson zèbre absorbe les molécules qui sont dissoutes dans l'eau ; (vi) enfin, les coûts d'élevage du poisson zèbre sont très inférieurs à ceux des autres vertébrés modèles.

Le poisson zèbre est utilisé en recherche depuis le début des années 1980, principalement dans le domaine de la biologie du développement. Cependant, ses caractéristiques en font aussi un modèle émergeant particulièrement intéressant pour la recherche en cancérologie [4] et notamment pour l'étude du comportement des cellules cancéreuses après transplantation lors d'expériences d'allogreffe ou de xéno greffe.

### **Étude de la cancérogenèse par l'allogreffe de cellules tumorales chez le poisson zèbre**

Le poisson zèbre peut développer de nombreux types de cancers, tels que des leucémies à cellules B et T [5,6], des leucémies myéloïdes aiguës [7], des mélanomes [8], des neuroblastomes [9], des cancers hépatiques [10], des cancers pancréatiques [11], des tumeurs malignes des gaines des nerfs périphériques [12], des tumeurs de la lignée germinale [13] ou des sarcomes [14] suite à l'expression d'oncogènes par transgénèse ou à l'inactivation de gènes suppresseurs de tumeurs. Beaucoup des cancers du poisson zèbre obtenus par manipulation d'oncogènes ou de gènes suppresseurs de tumeurs altérés dans les cancers humains, sont similaires sur les plans histologique et moléculaire aux tumeurs humaines. Ainsi, il existe une conservation évolutive des mécanismes impliqués dans la tumorigenèse entre l'homme et le poisson zèbre [15]. L'allogreffe de ces cancers chez le poisson zèbre offre la possibilité d'étudier de nombreux aspects de la tumorigenèse comme la métastase, les cellules propagatrices de tumeurs, l'hétérogénéité tumorale ou l'écotaxie (homing) des cellules tumorales.

#### *Modèles de poisson zèbre utilisés en transplantation allogénique de cellules cancéreuses*

Dans les expériences de transplantations, le système immunitaire des vertébrés fait obstacle au développement des cellules cancéreuses injectées. Chez le poisson zèbre, le système immunitaire adaptatif commence à être pleinement fonctionnel vers les jours 21 à 28 [16,17] et les études de transplantations nécessitent donc que les poissons adultes hôtes soient immunosupprimés ou immunotolérants. Pour cela, plusieurs approches sont possibles. La suppression du système immunitaire peut être obtenue par traitements aux rayons g ou à la dexaméthasone. Une irradiation du poisson zèbre adulte par une dose de rayons g de 20–25 Gy induit la suppression du système immunitaire qui sera complètement restauré 20 jours plus tard [18]. Cependant, cette immunosuppression rend possible les études transitoires de transplantation [5,18,19]. En revanche, le traitement par la dexaméthasone permet des études de transplantation sur des périodes plus longues si le traitement est maintenu [20]. La dexaméthasone est responsable de l'ablation des cellules B et T et permet la transplantation de cellules de tumeurs solides mais pas de leucémies lymphoïdes [21,22]. Une approche alternative consiste à utiliser des poissons zèbres receveurs porteurs de mutations qui altèrent la fonction des cellules du système immunitaire. C'est ainsi que l'équipe de David Langenau a généré par édition du génome, une lignée de poissons zèbres porteurs d'une mutation dans le gène rag2 [23]. Cette mutation dans rag2 altère la recombinaison V(D)J et empêche la maturation des cellules B et T sans affecter la viabilité des poissons qui peuvent être transplantés par divers types de cellules cancéreuses [24]. Cependant, les poissons porteurs de la mutation rag2 présentent des taux variables de cellules B qui affectent les analyses de xenotransplantation et de plus, la lignée ne peut pas être maintenue à l'état homozygote. Pour pallier ces limitations, une lignée de poissons zèbres porteurs d'une mutation dans le gène prkdc a été générée par édition génomique [25]. Cette lignée ne possède pas de lymphocytes T et B matures, mais peut se reproduire à l'état homozygote et permet le suivi des cellules tumorales après transplantation.

L'immunosuppression par irradiation, par traitement chimique ou par manipulations génétiques ne permet pas d'étudier les relations entre les cellules tumorales et le système immunitaire. Aussi, la transplantation de cellules tumorales d'un poisson donneur vers un receveur génétiquement identique (clonal ou syngénique) est une alternative qui permet l'étude du comportement des cellules cancéreuses en présence d'un système immunitaire

pleinement fonctionnel [19,26,27]. Mais dans ce cas, seules les tumeurs générées dans les lignées syngéniques peuvent être transplantées, restreignant ainsi le champ d'application de cette méthodologie. Une limitation au suivi des cellules cancéreuses transplantées est également liée à la difficulté de visualisation de ces cellules dans le poisson zèbre receveur. Pour cela, une souche de poisson zèbre, casper (figure 1A), a été générée par le croisement entre deux souches mutantes, la souche nacre et la souche roy orbison [28]. La souche nacre possède une mutation dans le gène *mitfa* et ne possède pas de mélanocytes absorbant la lumière incidente, alors que la souche roy orbison n'a pas les iridophores qui réfléchissent la lumière. L'absence de mélanocytes et d'iridophores chez les poissons zèbres casper les rend relativement transparent et permet une bonne détection des cellules transplantées, et tout particulièrement si ces cellules cancéreuses expriment des marqueurs fluorescents. Les approches de transgénèse offrent par ailleurs de multiples possibilités d'expression de protéines fluorescentes chez les poissons donneurs qui permettront de suivre ces cellules cancéreuses après leur transplantation [5,14,24,28,29].

L'utilisation de poissons zèbres transgéniques en tant que receveurs dans les expériences de transplantation de cellules tumorales peut également permettre l'analyse d'aspects particuliers liés à la cancérogenèse. C'est ainsi par exemple, que la lignée transgénique Tg(*fli1* :GFP) qui exprime la protéine fluorescente verte GFP sous le contrôle du promoteur *fli1* dans les vaisseaux sanguins [30] favorise l'étude des phénomènes métastatiques. En effet, la transplantation de cellules de lymphomes qui expriment la protéine fluorescente rouge mCherry dans des poissons transgéniques Tg(*fli1* :GFP) de fond génétique casper permet de visualiser la manière dont les cellules tumorales s'intravasent dans le système vasculaire [31].

*Allotransplantation comme mesure du pouvoir tumorigène des cellules cancéreuses.*

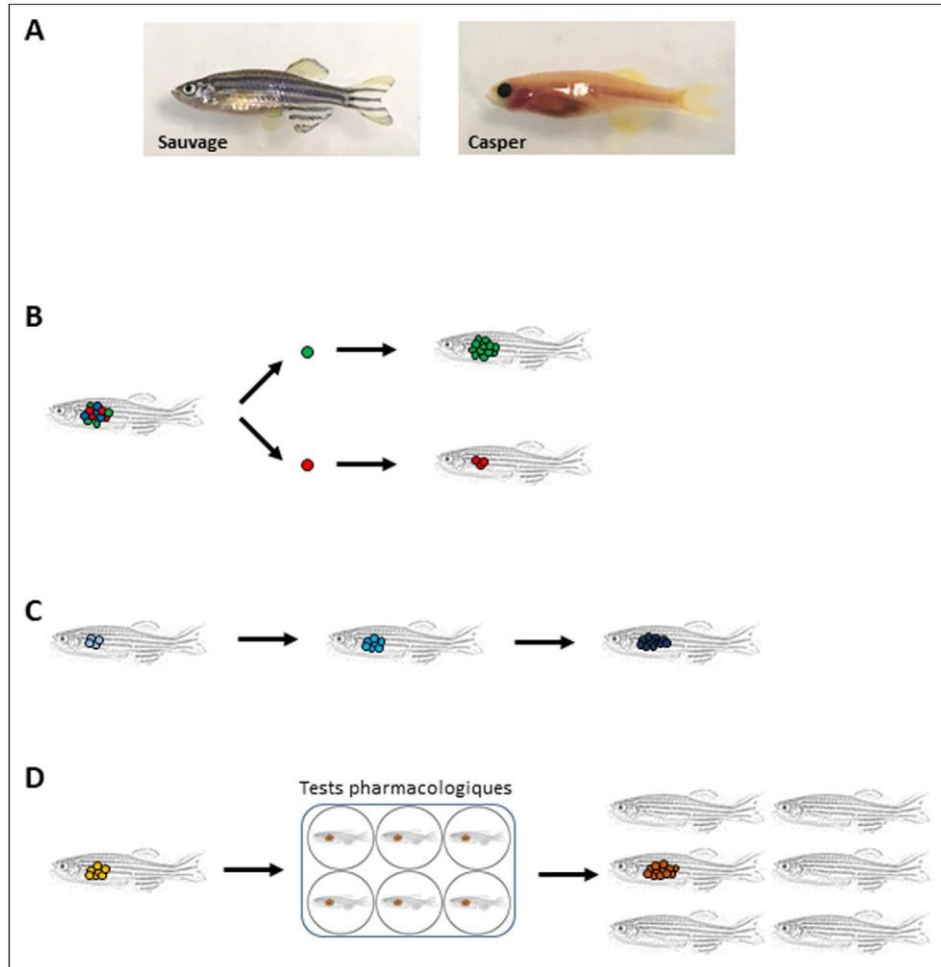
Les premières expériences d'allogreffes de cellules cancéreuses chez le poisson zèbre datent de 2003 [5].

Dans cette étude, l'oncogène murin *Mycest* fusionné au gène de la GFP (green fluorescent protein) et exprimé sous le contrôle du promoteur du gène *rag2* dans les lymphocytes T. Les cellules qui expriment + *Myc* (GFP) émergent du thymus, colonisent les muscles squelettiques et d'autres organes voisins pour provoquer la mort des poissons transgéniques avant l'âge de trois mois. Après transplantation dans la cavité péritonéale d'un poisson receveur, ces cellules conservent leur phénotype leucémique et présentent un phénomène d'écotaxie vers le thymus. De la même façon, les cellules tumorales qui proviennent de leucémies à cellules B, de mélanomes, de rhabdomyosarcomes, de neuroblastomes, de carcinomes hépatiques ou pancréatiques se développent après allotransplantation chez le poisson zèbre [6,24, 26,28,29]. En revanche, les cellules de syndromes myéloprolifératifs ou de lésions pré-malignes ne se maintiennent pas dans les poissons receveurs après transplantation [32,33]. Ainsi, l'allotransplantation permet d'évaluer le pouvoir tumorigène des cellules transplantées.

### *Allotransplantation pour caractériser l'hétérogénéité et la progression tumorale.*

Lors de la progression tumorale, la transformation d'une hyperplasie bénigne en tumeur cancéreuse est un processus stochastique qui implique des changements génomiques et épigénétiques et conduit l'apparition de clones génétiquement distincts au sein de la tumeur. Si l'allotransplantation chez le poisson zèbre permet d'imager la croissance tumorale au cours du temps [28], elle permet également d'étudier l'hétérogénéité cellulaire au sein des tumeurs [19,33]. Ainsi, des cellules uniques issues d'une même tumeur primaire chez un modèle de poisson zèbre transgénique qui développe une leucémie lymphoblastique T aiguë, forment après transplantation, des leucémies distinctes en termes de nombre de cellules propagatrices de leucémie, d'activation de voies de signalisation, de latence ou de réponse aux drogues. De même, les caractéristiques fonctionnelles des cellules tumorales peuvent changer suite à une évolution clonale au cours de transplantations successives. G12D Un modèle de poisson zèbre transgénique Tg(*rag2* :KRAS)<sup>Asp12</sup> qui exprime l'oncogène KRAS sous le contrôle du promoteur *rag2*, développe des rhabdomyosarcomes embryonnaires qui possèdent les mêmes caractéristiques histologiques et moléculaires que celles observées dans la pathologie humaine [14]. L'injection dans les poissons porteurs de rhabdomyosarcomes, de marqueurs fluorescents sous le contrôle de promoteurs dont l'activité dépend du stade de différenciation musculaire, permet la visualisation de populations cellulaires distinctes et de degré de différenciation différents, au sein de ces tumeurs [29]. Le tri et des dilutions successives de ces cellules tumorales avant transplantation montrent que ce sont les cellules les moins différenciées *myf5* qui possèdent le pouvoir de propagation tumorale. Ainsi, l'allotransplantation de cellules cancéreuses chez le poisson zèbre se révèle être un outil de choix pour la caractérisation des cellules propagatrices de tumeur (CPT) qui sont impliquées dans la récurrence des cancers. Cette approche combinée à l'utilisation de molécules chimiques permet d'identifier des

stratégies thérapeutiques de lutte contre les CPT [34]. +Si les CPT myf5 possède un fort pouvoir prolifératif, elles restent incapables d'entrer dans la circulation sanguine. À l'inverse, les cellules de rhabdomyosarcome plus différenciées qui expriment le marqueur myog (myogénine) ont une faible capacité proliférative mais un fort pouvoir d'intravasation [29].



**Figure 1** : L'allo-transplantation de cellules tumorales chez le poisson zèbre

A. La lignée de poisson zèbre casper dépourvue de mélanocytes et d'iridophores, permet le suivi des cellules cancéreuses transplantées ; B. Évaluation de l'hétérogénéité tumorale. Des cellules uniques isolées à partir d'une tumeur primaire hétérogène sont transplantées dans des poissons zèbres receveurs permettant ainsi l'évaluation du pouvoir tumorigène des clones tumoraux ; C. Suivi de l'évolution tumorale. Des transplantations successives permettent l'étude de l'évolution des tumeurs ; D. Test de molécules anticancéreuses. L'exposition de poissons zèbres transplantés permet d'évaluer l'action antitumorale de molécules et d'identifier des clones résistants ; (B), (C) et (D) sont adaptés de Blackburn et Langenau [87]

En utilisant l'expression de protéines fluorescentes sous le contrôle de promoteurs choisis comme marqueurs de différenciation dans des expériences d'allo-transplantation chez le poisson zèbre, il est donc possible d'étudier de nombreux aspects liés à la tumorigénèse, tels que l'investigation fonctionnelle de l'hétérogénéité tumorale, la caractérisation des cellules propagatrices de tumeurs, l'étude de l'évolution clonale ou de la résistance aux thérapies (figure 1B-D).

### **Xénotransplantation de cellules tumorales humaines chez l'embryon de poisson zèbre**

Les premiers essais de xénotransplantation de cellules tumorales humaines ont été publiés en 2005 par Lisa Lee *et al.* [35]. Entre une et cent cellules de mélanome métastatique humain qui présentent un phénotype dédifférencié, ont été transplantées dans des embryons de poisson zèbre au stade blastula (3,5– 4,5 heures après fécondation). Ces

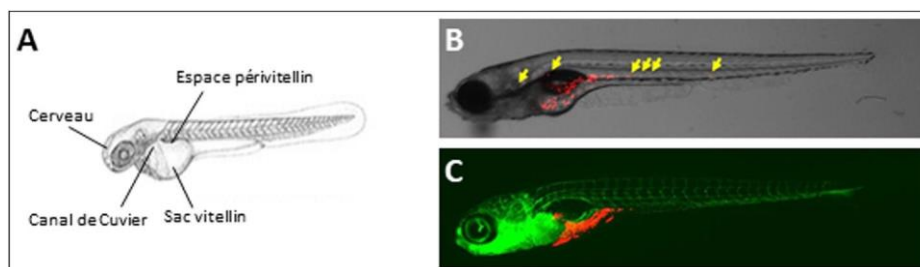
cellules survivent, présentent une importante motilité et se divisent dans l'hôte en maintenant leur état de dédifférenciation. Cette étude fut suivie par de nombreuses autres qui montraient que de nombreux types de cellules cancéreuses humaines peuvent être transplantés dans des embryons de poisson zèbre afin d'en étudier les propriétés [36,37].

### *Paramètres de la xénotransplantation*

Si les premières transplantations de cellules cancéreuses humaines ont été effectuées sur des embryons au stade blastula [35] et même si elles peuvent être pratiquées sur des poissons immunodéprimés d'un mois [38], la très grande majorité des expériences de xénotransplantations chez le poisson zèbre se fait en réalité sur des embryons âgés de deux jours. À ce stade, les organes et le système vasculaire du poisson zèbre sont développés et les chances que la transplantation des cellules cancéreuses induise des anomalies du développement restent faibles. En revanche, le système immunitaire adaptatif des poissons zèbres n'est pas encore développé à cet âge [17] évitant ainsi le rejet des cellules injectées.

Le site d'injection des cellules cancéreuses dans l'embryon de poisson zèbre de deux jours est un paramètre associé à la xénotransplantation (figure 2A). Les cellules de glioblastomes ou de rétinoblastomes peuvent être injectées dans le cerveau [39,40] ou la cavité oculaire [41,42] et les cellules leucémiques dans la veine caudale [43]. En revanche, la transplantation orthotopique n'est pas possible pour les cellules issues de cancers du sein, du poumon ou de la prostate. Aussi, la xénotransplantation de cellules tumorales humaines est souvent effectuée de manière hétérotopique. Le site d'injection le plus utilisé est alors le sac vitellin [36]. Le sac vitellin est une structure acellulaire suffisamment large qui permet d'envisager des expériences de xénotransplantation à relativement large échelle. Il fournit également un environnement riche en nutriments, favorable à la croissance et à la migration des cellules cancéreuses. Pour suivre les étapes plus tardives de la progression tumorale, la transplantation des cellules cancéreuses peut se faire dans l'espace périvitellin (l'espace entre la peau et la membrane externe du sac vitellin) [44]. Ce site permet l'injection d'une quantité de cellules relativement grande et se trouve à proximité de vaisseaux sanguins majeurs. L'injection directe dans la circulation au niveau du canal de Cuvier (veine cardinale commune) est également très utilisée pour les expériences de xénotransplantation [45–48]. Alors que l'injection de cellules dans l'espace périvitellin permet de suivre l'efficacité de l'intravasation, l'injection dans le canal de Cuvier permet l'étude de l'invasion et des micrométastases [45]. Dans le cas de poissons âgés de plus d'un mois, la xénotransplantation hétérotopique se fera dans la cavité péritonéale [38].

Une des problématiques liée à la xénotransplantation de cellules humaines chez le poisson zèbre est associée au fait que le poisson zèbre est un organisme poïkilotherme vivant à 26–28°C alors que les cellules humaines sont maintenues à 37°C. Toutefois, plusieurs équipes ont montré qu'il était possible de garder les embryons de poisson zèbre de deux jours à 35°C sans affecter leur développement tout en maintenant des conditions de température acceptable pour les cellules humaines transplantées [36,43,49].



**Figure 2** : La xénotransplantation de cellules tumorales dans des embryons de poissons zèbres de deux jours

A. Localisation des principaux sites de transplantation dans les expériences de xénogreffe ; B. Illustration du pouvoir invasif de cellules tumorales humaines (marquées en rouge) dans un embryon de poisson zèbre, 4 jours après transplantation ; C. Visualisation des cellules tumorales humaines (marquées en rouge) dans un embryon de poisson zèbre Tg (fli1:GFP) dont les vaisseaux sanguins apparaissent en vert.

Si les embryons et larves de poissons zèbres sont relativement transparents, quelques groupes de cellules pigmentées peuvent cependant altérer la visualisation des cellules tumorales transplantées. Aussi, le traitement des embryons par un inhibiteur de la mélanogénèse, la 1-phényl 2-thiourée (PTU) [50], ou l'utilisation de poissons de la souche casper [28] permettent d'obtenir des embryons receveurs parfaitement transparents. L'emploi de lignées transgéniques peut également faciliter les études d'interaction entre les cellules humaines et celles du poisson hôte dans les expériences de xénotransplantation. L'utilisation de la lignée Tg(fli1 :GFP) [30] permet ainsi d'étudier les relations entre les cellules tumorales et l'environnement vasculaire, alors que la lignée Tg(mpx :GFP) permet d'examiner la métastase induite par les neutrophiles [45]. D'autres lignées, telle que la lignée transgénique Tg(mpeg1 :GFP) qui marque les macrophages, pourront être utilisées pour suivre la réaction du système immunitaire inné de l'hôte en réponse à l'injection des cellules tumorales [51,52].

Le suivi des cellules humaines transplantées peut se faire en les marquant à l'aide de molécules lipophiles fluorescentes (figure 2B-C). Ces molécules se diluent au cours des divisions cellulaires et une alternative à cette approche est la transformation des cellules tumorales de façon à ce qu'elles expriment des protéines fluorescentes. La manipulation génétique des cellules avant leur transplantation peut, en outre, permettre d'étudier de nombreux aspects de la cancérogenèse.

### *Xénotransplantation chez le poisson zèbre et cancérogenèse*

De nombreuses études ont montré que la xénotransplantation de cellules tumorales de différentes origines dans le poisson zèbre permet l'analyse in vivo de divers phénomènes cancéreux comme l'invasion et la métastase [35,53,54], la néoangiogenèse [36,37,44,55] ou la réponse aux thérapies anticancéreuses [43,49,54,56,57].

L'utilité du modèle de xénotransplantation chez le poisson zèbre pour l'étude de l'invasion est clairement illustrée par les travaux de l'équipe de Christoph Bagowski qui a transplanté deux lignées de cellules de cancer pancréatique qui présentaient un fort (cellules PaTu8988-T n'exprimant pas la cadhérine E) ou un faible (cellules PaTu8988-S exprimant la cadhérine E) pouvoir invasif [54]. Deux jours après transplantation dans le sac vitellin d'embryons de deux jours, les cellules PaTu8988-T migraient dans la partie caudale de l'embryon alors que les cellules PaTu8988-S restaient au niveau du site d'injection. Le micro-ARN miR-10a apparaît comme un modulateur clé du comportement métastatique de ces cellules cancéreuses pancréatiques [57]. En effet, alors que les cellules PaTu8988-T à fort pouvoir invasif expriment le miR-10a, son expression n'est pas détectée dans les cellules PaTu8988-S qui migrent peu. L'inhibition de l'expression du miR-10a à l'aide d'un oligonucléotide complémentaire bloque le pouvoir invasif et métastatique des cellules PaTu8988-T transplantées dans l'embryon de poisson zèbre. De façon similaire, les antagonistes de l'acide rétinoïque, un activateur de l'expression de miR-10a, empêchent le développement des métastases. À l'inverse, la surexpression du miR10a augmente les capacités invasives des cellules PaTu8988-S dans ce modèle de xénotransplantation. Enfin, la dissémination des cellules tumorales dépend du système vasculaire puisque les cellules PaTu8988-T transplantées dans des embryons de poissons zèbres mutants cloche, dépourvues de système vasculaire fonctionnel, restent localisées à leur site d'injection malgré leur fort pouvoir invasif [54].

L'intravasation des cellules tumorales dans la circulation peut être visualisée et étudiée lors d'expériences de xénotransplantation. C'est ainsi que Stoletov *et al.* [38] ont injecté des cellules cancéreuses humaines dans la cavité péritonéale de poissons zèbres transgéniques Tg(fli1 :GFP) âgés d'un mois. Les cellules transplantées forment des microtumeurs responsables d'un remodelage et d'une perméabilisation vasculaire locale. Ce sont les cellules qui surexpriment le gène pro-métastatique RHOC qui intravasent préférentiellement au niveau de ces sites. Ce phénomène est la conséquence de la sécrétion de VEGF par les cellules qui expriment RHOC [38]. De la même manière, la xénotransplantation chez le poisson zèbre permet l'étude de l'extravasation. Ainsi, moins d'un jour après l'injection de cellules cancéreuses dans la circulation (canal de Cuvier) d'embryons de poissons zèbres transgéniques Tg(fli1 :GFP) à VEGFgés de deux jours, les cellules forment des embolies au niveau de l'artère caudale et sortent de la circulation. Alors que certaines cellules s'extravasent activement en s'insérant dans la paroi endothéliale, d'autres forment des amas cellulaires à l'intérieur des vaisseaux. Ces amas sont progressivement recouverts de cellules endothéliales puis sont exclus de la circulation vers les tissus environnants [58]. L'utilisation de méthodes d'inhibition de l'expression génique à l'aide de petits ARN interférents ou de surexpression dans les cellules tumorales humaines avant transplantation, a permis de montrer que l'intégrine b1 (ITGB1), le VEGF-A et le facteur pro-métastatique TWIST1 sont impliqués dans la locomotion et l'extravasation des cellules cancéreuses

[58]. Si les propriétés intrinsèques des cellules contrôlent le développement des métastases, les conditions environnementales comme les pressions mécaniques liées à la force du débit sanguin sont également des paramètres importants dans le développement des métastases. L'équipe de Jacky Goetz a suivi le mouvement et l'extravasation de cellules tumorales murines transplantées dans le canal de Cuvier d'embryons de poissons zèbres transgéniques Tg(fli1 :GFP) âgés de deux jours. Les mesures de vélocité de la circulation sanguine et sa modulation à l'aide de molécules qui altèrent l'activité cardiaque montrent que le flux sanguin contrôle l'arrêt des cellules tumorales circulantes dans les vaisseaux ainsi que leur extravasation [59].

De nombreuses données indiquent que les tumeurs contiennent une petite sous-population de cellules souches cancéreuses (CSC) capables de se renouveler, résister aux thérapies, former une tumeur primaire et des métastases [60]. Plusieurs études montrent que le modèle de xénotransplantation chez le poisson zèbre permet de caractériser les CSC de cancers du sein, de leucémies, de gliomes, de cancers du foie ou de la prostate [61–66]. En effet, après transplantation, les CSC forment des masses tumorales et migrent à une fréquence plus élevée que les cellules cancéreuses non-souches [61–65]. La xénotransplantation de CSC dans les embryons de poissons zèbres de deux jours constitue donc une excellente plateforme expérimentale pour la recherche de molécules ciblant spécifiquement ces cellules [62–64].

L'angiogenèse permettant la croissance de tumeurs primaires et la formation de tumeurs secondaires [67] peut être étudiée à l'aide du modèle de xénotransplantation dans les embryons de poissons zèbres de deux jours [36,44,55,68,69]. Les cellules cancéreuses transplantées se regroupent rapidement et recrutent les cellules endothéliales de l'hôte qui sont capables d'infiltrer ces masses cellulaires et former des vaisseaux [36]. Ces vaisseaux néoformés expriment les marqueurs endothéliaux comme fli1, kdrl (vegfr2) ou la VE-cadhérine (cdh5), et leur formation dépend de la production de FGF ou VEGF par les cellules transplantées [37]. Dans ce contexte, la manipulation génétique des cellules cancéreuses avant transplantation couplée à l'utilisation de poissons zèbres transgéniques Tg(fli1 :GFP) receveurs est une approche efficace pour l'étude des mécanismes moléculaires impliqués dans la néoangiogenèse induite par les cellules tumorales. Par exemple, l'inactivation à l'aide de petits ARN interférants, des kinases LIMK1 et LIMK2 impliquées dans la réorganisation du cytosquelette d'actine, réduit la néovascularisation induite par les cellules pancréatiques tumorales [55]. Ainsi, la xénotransplantation de cellules tumorales humaines dans le poisson zèbre permet d'aborder de nombreuses étapes de la cancérogenèse et montre également que les cellules humaines communiquent parfaitement avec les tissus du poisson receveur [70].

### *Xénotransplantation pour la recherche de molécules anticancéreuses*

Le modèle de transplantation de cellules tumorales chez le poisson zèbre est un outil qui possède un fort potentiel dans la recherche de molécules anticancéreuses [43,49,71,72]. L'environnement aquatique du poisson zèbre fait que beaucoup de molécules (selon leur solubilité) ajoutées dans l'eau seront absorbées par le poisson, évitant ainsi la pratique de l'injection à chaque animal. Par ailleurs, le poisson zèbre possède le complément complet des cytochromes P450 responsables du métabolisme des drogues [73,74]. Des études montrent également les grandes similitudes en termes de perméabilité de la barrière hémato-encéphalique entre poisson zèbre et vertébrés supérieurs [75]. Ainsi, le poisson zèbre apparaît comme un modèle pertinent pour l'identification et l'étude des propriétés pharmacologiques de molécules anticancéreuses. C'est à la fois, l'efficacité et l'éventuelle toxicité des molécules qui peut être testée chez ce modèle.

De nombreuses molécules ont été testées en utilisant le modèle de xénotransplantation de cellules tumorales chez le poisson zèbre de deux jours [43,56,64,72,76–79]. Par exemple, le traitement par divers agents anti-leucémiques (imatinib, cyclophosphamide, dasatinib ou parthénolide) réduit la charge leucémique sans affecter le développement d'embryons de deux jours transplantés par des cellules leucémiques [43,64]. Dans un modèle de xénotransplantation orthotopique, Welker *et al.* [79] ont même montré que le témozolomide pouvait affecter l'hétérogénéité tumorale de glioblastomes.

Bien que l'injection des cellules tumorales dans les embryons de poisson zèbre reste un processus manuel, l'automatisation et l'imagerie rapide à haute résolution des embryons injectés [80] ainsi que la taille des embryons qui peuvent être placés dans des plaques à 96 puits, permettent d'envisager le criblage de molécules chimiques afin d'identifier de nouveaux agents anticancéreux plus efficaces [63,72].

### *Xénotransplantation de cellules tumorales dérivées de patients pour une prise en charge personnalisée*



Le fait que la xénotransplantation chez l'embryon de poisson zèbre ne nécessite qu'un nombre limité de cellules tumorales (1 à 500) pour un résultat rapide (quelques jours) en termes d'évaluation du pouvoir tumorigène ou angiogénique, permet l'application de cette méthodologie aux cellules dérivées de patients (PDX, patient-derived xenograft) [43,54,57,71,81–84]. L'utilisation de tumeurs primaires dans les expériences de xénotransplantation chez le poisson zèbre s'applique à la fois aux leucémies dont les cellules sont plus faciles à obtenir et injecter [43,71] et aux tumeurs solides telles que les cancers du pancréas [54,57], les métastases osseuses [81], des carcinomes des glandes salivaires [83], les tumeurs neuroendocrines [84] ou de l'appareil digestif [54,82] dont les cellules doivent être dissociées avant injection. La transplantation des cellules issues de patients montre que ces dernières ont des capacités prolifératives et invasives dans le poisson zèbre receveur. Ainsi, cette approche montre que la xénotransplantation de cellules tumorales de patients peut être utilisée pour étudier leur comportement *in vivo*, mais plus important encore, l'approche permet également d'évaluer la réponse des cellules aux molécules anticancéreuses [43,57,71,82,83] dans le but d'identifier la thérapie individuelle la plus appropriée en préclinique. La pertinence du modèle de xénotransplantation chez le poisson zèbre en tant que plateforme préclinique pour une thérapie personnalisée a été démontrée par Bentley *et al.* [71]. Les cellules d'un patient de trois ans et demi atteint d'une leucémie lymphoblastique T aiguë ne répondant pas après transplantation dans le poisson zèbre, à la rapamycine, un inhibiteur de la voie mTOR, mais fortement à un inhibiteur de la g-sécrétase, a suggéré une altération dans la voie NOTCH. L'altération de cette voie est confirmée par l'identification d'une mutation ponctuelle dans le gène NOTCH1 [71]. De manière similaire, Fior *et al.* [82] ont étudié la réponse au FOLFOX (5-fluorouracile, oxaliplatine, acide folinique) de cellules tumorales de cancers colorectaux issues de patients après xénotransplantation chez le poisson zèbre. Dans la plupart des cas, la réponse des différents prélèvements au traitement était en corrélation avec les résultats de rechute observés en clinique 3 à 6 mois après chirurgie. De plus, les prélèvements de patients qui ne répondaient pas au cetuximab présentaient tout comme attendu, des mutations dans les oncogènes KRAS ou BRAF [82]. Ces différentes études sont des preuves de concept qui démontrent qu'il est possible de prédire la réponse des cellules tumorales de patients aux différents protocoles thérapeutiques dans une période d'une semaine suivant la biopsie, dans le cadre d'une médecine personnalisée.

### Conclusion

#### *Avantages et limites de l'utilisation du modèle poisson zèbre*

Le modèle poisson zèbre a émergé en tant qu'outil puissant pour l'étude génétique du développement chez les vertébrés. Les propriétés de ce modèle sont multiples et font que le poisson zèbre se place maintenant comme un acteur novateur dans la recherche sur le cancer (figure 3). Bien qu'elle en soit encore à ses débuts, l'approche de transplantation de cellules tumorales chez le poisson zèbre tire parti des nombreux avantages inhérents à ce modèle et offre la possibilité d'étudier la formation de la tumeur, son évolution, la formation de métastases et d'évaluer l'effet de molécules thérapeutiques anticancéreuses. Appliquée aux cellules dérivées de patients, la xénotransplantation chez l'embryon de poisson zèbre de deux jours pourrait éclairer les décisions relatives au choix du traitement optimal dans un délai relativement rapide.

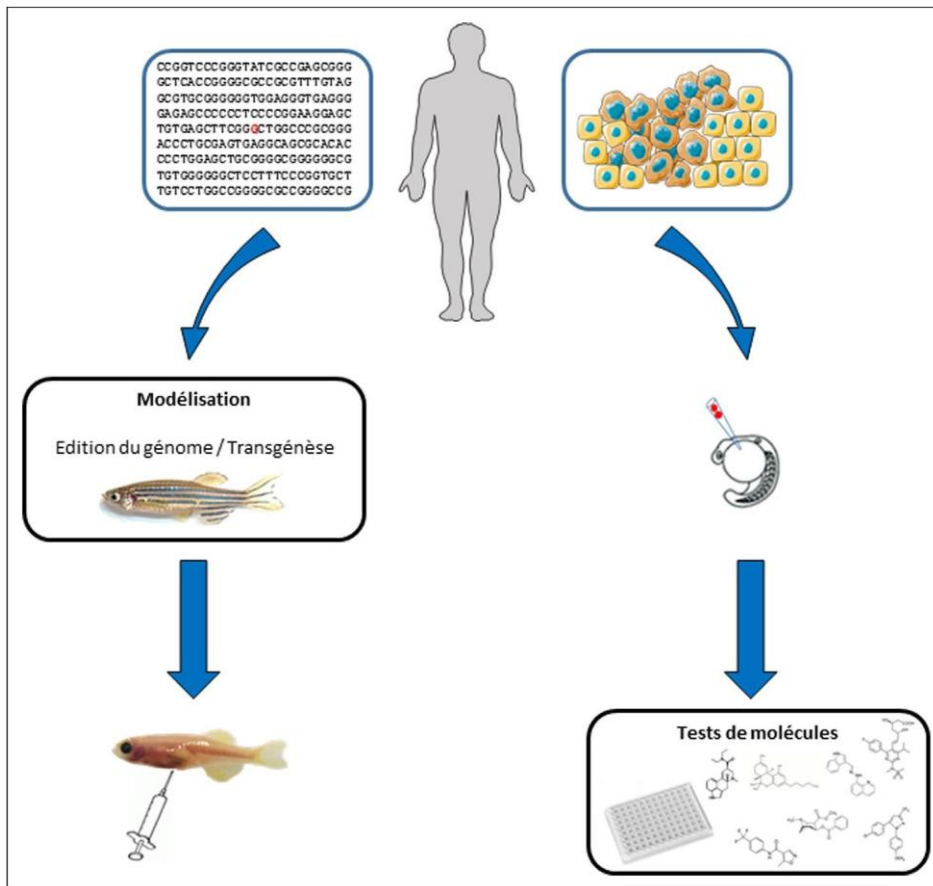


Figure 3 : Le modèle poisson zèbre dans la lutte contre le cancer

(Partie gauche) Les analyses génomiques permettent d'identifier les altérations génétiques associées aux cancers humains. Les approches par génétique inverse (édition du génome, transgénèse) permettent de modéliser ces altérations chez le poisson zèbre alors que l'allotransplantation des cellules tumorales permet d'étudier les caractéristiques de ces cancers. (Partie droite) La xénotransplantation de cellules tumorales issues de patients permet d'établir le choix des protocoles thérapeutiques adaptés dans le cadre d'une médecine personnalisée

Toutefois, comme tous les modèles expérimentaux, le modèle poisson zèbre possède un certain nombre de limitations. En particulier, les transplantations de cellules tumorales dérivées de patients sont effectuées dans des embryons de deux jours dépourvus du système immunitaire adaptatif, masquant ainsi en partie l'influence de la réaction immunitaire dans la progression tumorale. L'absence de tissus spécifiques comme le sein, les poumons ou la prostate fait que certaines cellules tumorales ne trouveront pas leur niche physiologique après xénotransplantation. Il existe également des différences génétiques notables entre le poisson zèbre et l'homme. Au cours de l'évolution, l'apparition des poissons téléostéens s'est accompagnée d'une duplication entière du génome suivie d'une perte massive de gènes qui fait qu'il existe environ 26 000 gènes codants chez le poisson zèbre pour environ 20 000 chez l'homme [85]. Cette duplication fait que certains gènes impliqués dans le cancer, comme NF1 ou PTEN, sont présents chez le poisson zèbre sous forme de deux orthologues, À l'inverse, la perte massive de gènes associée à cette duplication fait que d'autres gènes tels BRCA1, CDKN2A, LIF, OSM ou IL6, sont absents chez le poisson. En ce qui concerne le métabolisme des drogues dans le cadre de la recherche de molécules anticancéreuses, les analyses protéomiques et transcriptomiques montrent un fort degré de conservation des enzymes métaboliques entre l'homme et le poisson zèbre. Cependant, l'étude du métabolisme de la testostérone l'homme et le poisson zèbre adulte révèle que si les métabolites majeurs de la testostérone sont identiques, des différences sont tout de même détectées pour les métabolites mineurs [86]. Cependant, malgré ses limites, le modèle poisson zèbre a montré son utilité dans la recherche contre le cancer, offrant des opportunités uniques qui sont complémentaires aux approches réalisées à l'aide des modèles murins ou cellulaires. La demande clinique pour

une médecine personnalisée d'une part, et la volonté de l'industrie pharmaceutique à utiliser des outils de présélection moins onéreux pour identifier ou repositionner des molécules thérapeutiques d'autre part, font que la popularité du modèle de xénotransplantation chez le poisson zèbre va très vraisemblablement augmenter à l'avenir pour devenir un outil particulièrement puissant dans la lutte contre le cancer.

### Financements

Ce travail est supporté par l'Université de Lille, l'Inserm, la Ligue Contre le Cancer et le Cancéropôle Nord-Ouest.

### Références

- [1] Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018 : GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin* 2018;68:394–424.
- [2] Pascussi JM, Giraud J, Samalin E, Grillet F, Pannequin J. Bases fondamentales du processus métastatique. *Bull Cancer* 2016;103: S39–47.
- [3] Sharpless NE, Depinho RA. The mighty mouse: genetically engineered mouse models in cancer drug development. *Nat Rev Drug Discov* 2006;5:741–54.
- [4] Völkel P, Dupret B, Le Bourhis X, Angrand PO. Le modèle poisson zèbre dans la lutte contre le cancer. *Med Sci (Paris)* 2018;34:345–53.
- [5] Langenau DM, Traver D, Ferrando AA, Kutok JL, Aster JC, et al. Myc-induced T cell leukemia in transgenic zebrafish. *Science* 2003;299:887–90.
- [6] Sabaawy HE, Azuma M, Embree LJ, Tsai HJ, Starost MF, Hickstein DD, et al. TELAML1 transgenic zebrafish model of precursor B cell acute lymphoblastic leukemia. *Proc Natl Acad Sci USA* 2006;103:15166–71.
- [7] Zhuravleva J, Paggetti J, Martin L, Hammann A, Solary E, et al. MOZ/TIF2-induced acute myeloid leukaemia in transgenic fish. *Br J Haematol* 2008;143:378–82.
- [8] Patton EE, Widlund HR, Kutok JL, Kopani KR, Amatruda JF, et al. BRAF mutations are sufficient to promote nevi formation and cooperate with p53 in the genesis of melanoma. *Curr Biol* 2005;15:249–54.
- [9] Zhu S, Lee JS, Guo F, Shin J, Perez-Atayde AR, et al. Activated ALK collaborates with MYCN in neuroblastoma pathogenesis. *Cancer Cell* 2012;21:362–73.
- [10] Nguyen AT, Emelyanov A, Koh CH, Spitsbergen JM, Lam SH, et al. A high level of liverspecific expression of oncogenic Kras(V12) drives robust liver tumorigenesis in transgenic zebrafish. *Dis Model Mech* 2011;4:801–13.
- [11] Park SW, Davison JM, Rhee J, Hruban RH, Maitra A, Leach SD, et al. Oncogenic KRAS induces progenitor cell expansion and malignant transformation in zebrafish exocrine pancreas. *Gastroenterology* 2008;134:2080–90.
- [12] Berghmans S, Murphey RD, Wienholds E, Neubergh D, Kutok JL, et al. tp53 mutant zebrafish develop malignant peripheral nerve sheath tumors. *Proc Natl Acad Sci USA* 2005;102:407–12.
- [13] Gill JA, Lowe L, Nguyen J, Liu PP, Blake T, et al. Enforced expression of Simian virus 40 large T-antigen leads to testicular germ cell tumors in zebrafish. *Zebrafish* 2010;7:333–41.
- [14] Langenau DM, Keefe MD, Storer NY, Guyon JR, Kutok JL, et al. Effects of RAS on the genesis of embryonal rhabdomyosarcoma. *Genes Dev* 2007;21:1382–95.
- [15] Lam SH, Wu YL, Vega VB, Miller LD, Spitsbergen J, et al. Conservation of gene expression signatures between zebrafish and human liver tumors and tumor progression. *Nat Biotechnol* 2006;24:73–5.
- [16] Trede NS, Langenau DM, Traver D, Look AT, Zon LI. The use of zebrafish to understand immunity. *Immunity* 2004;20:367–79.
- [17] Lam SH, Chua HL, Gong Z, Lam TJ, Sin YM. Development and maturation of the immune system in zebrafish, *Danio rerio*: a gene expression profiling, in situ hybridization and immunological study. *Dev Comp Immunol* 2004;28:9–28.
- [18] Traver D, Winzeler A, Stern HM, Mayhall EA, Langenau DM, et al. Effects of lethal irradiation in zebrafish and rescue by hematopoietic cell transplantation. *Blood* 2004;104:1298–305.
- [19] Smith AC, Raimondi AR, Salthouse CD, Ignatius MS, Blackburn JS, et al. Highthroughput cell transplantation establishes that tumor-initiating cells are abundant in zebrafish T-cell acute lymphoblastic leukemia. *Blood* 2010;115:3296–303.
- [20] Moore JC, Langenau DM. Allograft Cancer Cell Transplantation in Zebrafish. *Adv Exp Med Biol* 2016;916:265–87.
- [21] Langenau DM, Ferrando AA, Traver D, Kutok JL, Hezel JP, et al. In vivo tracking of T cell development, ablation, and engraftment in transgenic zebrafish. *Proc Natl Acad Sci USA* 2004;101:7369–74.
- [22] Blackburn JS, Liu S, Wilder JL, Dobrinski KP, Lobbardi R, Moore FE, et al. Clonal evolution enhances leukemia-propagating cell frequency in T cell acute lymphoblastic leukemia through Akt/mTORC1 pathway activation. *Cancer Cell* 2014;25:366–78.

- [23] Tang Q, Abdelfattah NS, Blackburn JS, Moore JC, Martinez SA, et al. Optimized cell transplantation using adult rag2 mutant zebrafish. *Nat Methods* 2014;11:821–4.
- [24] Tang Q, Moore JC, Ignatius MS, Tenente IM, Hayes MN, et al. Imaging tumour cell heterogeneity following cell transplantation into optically clear immune-deficient zebrafish. *Nat Commun* 2016;7:10358.
- [25] Moore JC, Tang Q, Yordán NT, Moore FE, Garcia EG, et al. Single-cell imaging of normal and malignant cell engraftment into optically clear prkdc-null SCID zebrafish. *J Exp Med* 2016;213:2575–89.
- [26] Mizgirev IV, Revskoy SY. Transplantable tumor lines generated in clonal zebrafish. *Cancer Res* 2006;66:3120–5.
- [27] Mizgirev I, Revskoy S. Generation of clonal zebrafish lines and transplantable hepatic tumors. *Nat Protoc* 2010;5:383–94.
- [28] White RM, Sessa A, Burke C, Bowman T, LeBlanc J, et al. Transparent adult zebrafish as a tool for in vivo transplantation analysis. *Cell Stem Cell* 2008;2:183–9.
- [29] Ignatius MS, Chen E, Elpek NM, Fuller AZ, Tenente IM, et al. In vivo imaging of tumorpropagating cells, regional tumor heterogeneity, and dynamic cell movements in embryonal rhabdomyosarcoma. *Cancer Cell* 2012;21:680–93.
- [30] Lawson ND, Weinstein BM. In vivo imaging of embryonic vascular development using transgenic zebrafish. *Dev Biol* 2002;248:307–18.
- [31] Feng H, Stachura DL, White RM, Gutierrez A, Zhang L, et al. T-lymphoblastic lymphoma cells express high levels of BCL2, S1P1, and ICAM1, leading to a blockade of tumor cell intravasation. *Cancer Cell* 2010;18:353–66.
- [32] Le X, Langenau DM, Keefe MD, Kutok JL, Neuberg DS, Zon LI, et al. Heat shock-inducible Cre/Lox approaches to induce diverse types of tumors and hyperplasia in transgenic zebrafish. *Proc Natl Acad Sci USA* 2007;104:9410–5.
- [33] Blackburn JS, Liu S, Raiser DM, Martinez SA, Feng H, et al. Notch signaling expands a premalignant pool of T-cell acute lymphoblastic leukemia clones without affecting leukemiapropagating cell frequency. *Leukemia* 2012;26:2069–78.
- [34] Chen EY, DeRan MT, Ignatius MS, Grandinetti KB, Clagg R, et al. Glycogen synthase kinase 3 inhibitors induce the canonical WNT/bcatenin pathway to suppress growth and self-renewal in embryonal rhabdomyosarcoma. *Proc Natl Acad Sci USA* 2014;111:5349–54.
- [35] Lee LM, Seftor EA, Bonde G, Cornell RA, Hendrix MJ. The fate of human malignant melanoma cells transplanted into zebrafish embryos: assessment of migration and cell division in the absence of tumor formation. *Dev Dyn* 2005;233:1560–70.
- [36] Haldi M, Ton C, Seng WL, McGrath P. Human melanoma cells transplanted into zebrafish proliferate, migrate, produce melanin, form masses and stimulate angiogenesis in zebrafish. *Angiogenesis* 2006;9:139–51.
- [37] Nicoli S, Ribatti D, Cotelli F, Presta M. Mammalian tumor xenografts induce neovascularization in zebrafish embryos. *Cancer Res* 2007;67:2927–31.
- [38] Stoletov K, Montel V, Lester RD, Gonias SL, Klemke R, et al. High-resolution imaging of the dynamic tumor cell vascular interface in transparent zebrafish. *Proc Natl Acad Sci USA* 2007;104:17406–11.
- [39] Wehmas LC, Tanguay RL, Punnoose A, Greenwood JA. Developing a Novel Embryo-Larval Zebrafish Xenograft Assay to Prioritize Human Glioblastoma Therapeutics. *Zebrafish* 2016;13:317–29.
- [40] Welker AM, Jaros BD, Puduvali VK, Imitola J, Kaur B, Beattie CE. Standardized orthotopic xenografts in zebrafish reveal glioma cellline-specific characteristics and tumor cell heterogeneity. *Dis Model Mech* 2016;9:199–210.
- [41] Jo DH, Son D, Na Y, Jang M, Choi JH, et al. Orthotopic transplantation of retinoblastoma cells into vitreous cavity of zebrafish for screening of anticancer drugs. *Mol Cancer* 2013;12:71.
- [42] Chen X, Wang J, Cao Z, Hosaka K, Jensen L, et al. Invasiveness and metastasis of retinoblastoma in an orthotopic zebrafish tumor model. *Sci Rep* 2015;5:10351.
- [43] Pruvot B, Jacquél A, Droin N, Auberger P, Bouscary D, et al. Leukemic cell xenograft in zebrafish embryo for investigating drug efficacy. *Haematologica* 2011;96:612–6.
- [44] Nicoli S, Presta M. The zebrafish/tumor xenograft angiogenesis assay. *Nat Protoc* 2007;2:2918–23.
- [45] He S, Lamers GE, Beenakker JW, Cui C, Ghotra VP, et al. Neutrophil-mediated experimental metastasis is enhanced by VEGFR inhibition in a zebrafish xenograft model. *J Pathol* 2012;227:431–45.
- [46] Drabsch Y, He S, Zhang L, Snaar-Jagalska BE, ten Dijke P. Transforming growth factor- $\beta$  signalling controls human breast cancer metastasis in a zebrafish xenograft model. *Breast Cancer Res* 2013;15:R106.
- [47] Naber HP, Drabsch Y, Snaar-Jagalska BE, ten Dijke P, van Laar T. Snail and Slug, key regulators of TGF- $\beta$ -induced EMT, are sufficient for the induction of single-cell invasion. *Biochem Biophys Res Commun* 2013;435:58–63.
- [48] de Boeck M, Cui C, Mulder AA, Jost CR, Ikeno S, Ten Dijke P. Smad6 determines BMP-regulated invasive behaviour of breast cancer cells in a zebrafish xenograft model. *Sci Rep* 2016;6:24968.
- [49] Corkery DP, Delleire G, Berman JN. Leukaemia xenotransplantation in zebrafish–chemotherapy response assay in vivo. *Br J Haematol* 2011;153:786–9.
- [50] Karlsson J, von Hofsten J, Olsson PE. Generating transparent zebrafish: a refined method to improve detection of gene expression during embryonic development. *Mar Biotechnol (NY)* 2001;3:522–7.
- [51] Hamilton L, Astell KR, Velikova G, Sieger D. A zebrafish live imaging model reveals differential responses of microglia toward glioblastoma cells in vivo. *Zebrafish* 2016;13:523–34.
- [52] Britto DD, Wyroba B, Chen W, Lockwood RA, Tran KB, et al. Macrophages enhance Vegf-driven angiogenesis in an embryonic zebrafish tumour xenograft model. *Dis Model Mech* 2018;11 [pii: dmm035998].

- [53] Zhao H, Tang C, Cui K, Ang BT, Wong ST. A screening platform for glioma growth and invasion using bioluminescence imaging. Laboratory investigation. *J Neurosurg* 2009;111:238–46.
- [54] Marques IJ, Weiss FU, Vlecken DH, Nitsche C, Bakkers J, et al. Metastatic behaviour of primary human tumours in a zebrafish xenotransplantation model. *BMC Cancer* 2009;9:128.
- [55] Vlecken DH, Bagowski CP. LIMK1 and LIMK2 are important for metastatic behavior and tumor cell-induced angiogenesis of pancreatic cancer cells. *Zebrafish* 2009;6:433–9.
- [56] Harfouche R, Basu S, Soni S, Hentschel DM, Mashelkar RA, Sengupta S. Nanoparticle-mediated targeting of phosphatidylinositol3-kinase signaling inhibits angiogenesis. *Angiogenesis* 2009;12:325–38.
- [57] Weiss FU, Marques IJ, Woltering JM, Vlecken DH, Aghdassi A, et al. Retinoic acid receptor antagonists inhibit miR-10a expression and block metastatic behavior of pancreatic cancer. *Gastroenterology* 2009;137:2136–45.
- [58] Stoletov K, Kato H, Zardoujian E, Kelber J, Yang J, et al. Visualizing extravasation dynamics of metastatic tumor cells. *J Cell Sci* 2010;123:2332–41.
- [59] Follain G, Osmani N, Azevedo AS, Allio G, Mercier L, et al. Hemodynamic forces tune the arrest, adhesion, and extravasation of circulating tumor cells. *Dev Cell* 2018;45:33–52.
- [60] Turdo A, Veschi V, Gaggianesi M, Chinnici A, Bianca P, et al. Meeting the challenge of targeting cancer stem cells. *Front Cell Dev Biol* 2019;7:16.
- [61] Eguiara A, Holgado O, Belouqui I, Abalde L, Sanchez Y, et al. Xenografts in zebrafish embryos as a rapid functional assay for breast cancer stem-like cell identification. *Cell Cycle* 2011;10:3751–7.
- [62] Yang XJ, Cui W, Gu A, Xu C, Yu SC, et al. A novel zebrafish xenotransplantation model for study of glioma stem cell invasion. *PLoS One* 2013;8:e61801.
- [63] Zhang B, Shimada Y, Kuroyanagi J, Nishimura Y, Umemoto N, et al. Zebrafish xenotransplantation model for cancer stem-like cell study and high-throughput screening of inhibitors. *Tumour Biol* 2014;35:11861–69.
- [64] Zhang B, Shimada Y, Kuroyanagi J, Umemoto N, Nishimura Y, Tanaka T. Quantitative phenotyping-based in vivo chemical screening in a zebrafish model of leukemia stem cell xenotransplantation. *PLoS One* 2014;9: e85439.
- [65] Zoni E, van der Horst G, van de Merbel AF, Chen L, Rane JK, et al. miR-25 Modulates Invasiveness and Dissemination of Human Prostate Cancer Cells via Regulation of alpha5 and alpha6-Integrin Expression. *Cancer Res* 2015;75:2326–36.
- [66] Wenger A, Larsson S, Danielsson A, Elbæk KJ, Kettunen P, et al. Stem cell cultures derived from pediatric brain tumors accurately model the originating tumors. *Oncotarget* 2017;8:18626–39.
- [67] Mittal K, Ebos J, Rini B. Angiogenesis and the tumor microenvironment: vascular endothelial growth factor and beyond. *Semin Oncol* 2014;41:235–51.
- [68] Moshal KS, Ferri-Lagneau KF, Haider J, Pardhanani P, Leung T. Discriminating different cancer cells using a zebrafish in vivo assay. *Cancers (Basel)* 2011;3:4102–13.
- [69] Li D, Li XP, Wang HX, Shen QY, Li XP, et al. VEGF induces angiogenesis in a zebrafish embryo glioma model established by transplantation of human glioma cells. *Oncol Rep* 2012;28:937–42.
- [70] Tulotta C, Stefanescu C, Beletkaia E, Bussmann J, Tarbashevich K, et al. Inhibition of signaling between human CXCR4 and zebrafish ligands by the small molecule IT1t impairs the formation of triple-negative breast cancer early metastases in a zebrafish xenograft model. *Dis Model Mech* 2016;9:141–53.
- [71] Bentley VL, Veinotte CJ, Corkery DP, Pinder JB, LeBlanc MA, et al. Focused chemical genomics using zebrafish xenotransplantation as a pre-clinical therapeutic platform for T-cell acute lymphoblastic leukemia. *Haematologica* 2015;100:70–6.
- [72] Jung DW, Oh ES, Park SH, Chang YT, Kim CH, et al. A novel zebrafish human tumor xenograft model validated for anti-cancer drug screening. *Mol Biosyst* 2012;8:1930–9.
- [73] Alderton W, Berghmans S, Butler P, Chassaing H, Fleming A, et al. Accumulation and metabolism of drugs and CYP probe substrates in zebrafish larvae. *Xenobiotica* 2010;40:547–57.
- [74] Goldstone JV, McArthur AG, Kubota A, Zquette J, Parente T, et al. Identification and developmental expression of the full complement of Cytochrome P450 genes in Zebrafish. *BMC Genomics* 2010;11:643.
- [75] Jeong JY, Kwon HB, Ahn JC, Kang D, Kwon SH, et al. Functional and developmental analysis of the blood-brain barrier in zebrafish. *Brain Res Bull* 2008;75:619–28.
- [76] van der Ent W, Burrello C, Teunisse AF, Ksander BR, van der Velden PA, et al. Modeling of human uveal melanoma in zebrafish xenograft embryos. *Invest Ophthalmol Vis Sci* 2014;55:6612–22.
- [77] Zhang S, Cao Z, Tian H, Shen G, Ma Y, et al. SKLB1002, a novel potent inhibitor of VEGF receptor 2 signaling, inhibits angiogenesis and tumor growth in vivo. *Clin Cancer Res* 2011;17:4439–50.
- [78] Ghotra VP, He S, van der Horst G, Nijhoff S, de Bont H, et al. SYK is a candidate kinase target for the treatment of advanced prostate cancer. *Cancer Res* 2015;75:230–40.
- [79] Welker AM, Jaros BD, An M, Beattie CE.

- Changes in tumor cell heterogeneity after chemotherapy treatment in a xenograft model of glioblastoma. *Neuroscience* 2017;356:35–43.
- [80] Ghotra VP, He S, de Bont H, van der Ent W, Spaink HP, et al. Automated whole animal bio-imaging assay for human cancer dissemination. *PLoS One* 2012;7:e31281.
- [81] Mercatali L, La Manna F, Groenewoud A, Casadei R, Recine F, et al. Development of a Patient-Derived Xenograft (PDX) of Breast Cancer Bone Metastasis in a Zebrafish Model. *Int J Mol Sci* 2016;17:1375.
- [82] Fior R, Póvoa V, Mendes RV, Carvalho T, Gomes A, et al. Single-cell functional and chemosensitive profiling of combinatorial colorectal therapy in zebrafish xenografts. *Proc Natl Acad Sci USA* 2017;114:E8234–43.
- [83] Chen C, Choudhury S, Wangsa D, Lescott CJ, Wilkins DJ, et al. A multiplex preclinical model for adenoid cystic carcinoma of the salivary gland identifies regorafenib as a potential therapeutic drug. *Sci Rep* 2017;7:11410.
- [84] Gaudenzi G, Albertelli M, Dicitore A, Würth R, Gatto F, et al. Patient-derived xenograft in zebrafish embryos: a new platform for translational research in neuroendocrine tumors. *Endocrine* 2017;57:214–9.
- [85] Howe K, Clark MD, Torroja CF, Torrance J, Berthelot C, et al. The zebrafish reference genome sequence and its relationship to the human genome. *Nature* 2013;496:498–503.
- [86] Chng HT, Ho HK, Yap CW, Lam SH, Chan EC. An investigation of the bioactivation potential and metabolism profile of Zebrafish versus human. *J Biomol Screen* 2012;17:974–86.
- [87] Blackburn JS, Langenau DM. Zebrafish as a model to assess cancer heterogeneity, progression and relapse. *Dis Model Mech* 2014;7:755–62.

### 1.3 Conclusion

Le poisson zèbre est devenu un modèle de choix en cancérologie de par ses nombreux avantages. La conservation de nombreux gènes impliqués dans le développement tumoral entre l'Homme et le poisson permet d'extrapoler les résultats obtenus sur ce modèle. Les structures tissulaires des organes ainsi que des tumeurs sont très similaires. En identifiant et en prélevant les tumeurs développées chez un individu, il a été possible de réaliser des allogreffes afin notamment suivre la potentialité des cellules à former de nouvelles tumeurs et métastases et ainsi interroger leur clonogénicité. Le modèle poisson zèbre permet ainsi d'approfondir nos connaissances fondamentales sur les différents processus de cancérogénèse. Les xénogreffes de cellules cancéreuses d'origine humaines sur des embryons permettent aujourd'hui d'évaluer leurs capacités métastatiques à large échelle. Plus récemment, de nouveaux protocoles ont été mis en place. Les cellules cancéreuses directement issues des patients sont greffées sur les embryons. Ces tests permettent de prédire les réponses aux traitements des patients et pourraient devenir des tests de premiers plans pour orienter les choix thérapeutiques (Costa *et al.*, 2020 ; 2022).

## 2. Enhancement of Breast Cancer Cell Aggressiveness by lncRNA H19 and Its Mir-675 Derivative: Insight into Shared and Different Actions

### 2.1 Contexte scientifique

En France, le type de cancer le plus représenté chez les femmes est le cancer du sein. Malgré des campagnes de dépistages massives qui permettent la détection des tumeurs à un stade précoce dans plus de 60% des cas, le cancer du sein reste la 1<sup>ère</sup> cause de décès par cancer dans la population féminine française (*Panorama des cancers*, INCA, 2022). Plusieurs sous-types de cancers mammaires ont été définis en fonction de leur caractéristiques histologiques et moléculaires (Weigelt *et al.*, 2010 ; Eroles *et al.*, 2012). La majorité des cancers du sein sont dues à une hyperactivation des récepteurs hormonaux ER (estrogen receptor), PR (progesteron receptor) ou HER2 (human epidermal growth factor receptor-2). Une faible proportion de cancers du sein, dite tumeurs triple-négatives, n'est positive pour aucun de ces récepteurs. Ces tumeurs sont associées au pronostic de survie le plus sombre (Johannsson *et al.*, 2021).

Le pourcentage du génome qui code pour des protéines est très faible en comparaison de la proportion génomique transcrite (Qu *et al.*, 2013) et ce qui était auparavant appelé « ADN poubelle » représente de nombreux ARNs non codants, transcrits et fonctionnels (Kapranov *et al.* 2007). Ils sont regroupés en plusieurs classes en fonction de leur tailles et fonctions, comme les microARNs, les ARNs de transferts (ARNt) ou encore les longs ARN non codants (lncARNs). Le premier long ARN qui fut découvert est *H19* (Brannan *et al.*, 1990). Le lncARN *H19* possède de nombreux rôles biologiques. Il est exprimé durant le développement embryonnaire, puis au stade adulte dans quelques tissus spécifiques comme la glande mammaire (Adrienssens *et al.*, 1999). De façon intéressante, *H19* est surexprimé dans une grande majorité des tumeurs mammaires et augmente le pouvoir tumorigène des cellules cancéreuses (Lottin *et al.*, 2002). L'action de *H19* est médiée par des interactions avec de nombreuses protéines et notamment les protéines du PRC2 (Angrand *et al.*, 2015), mais il est également le précurseur de deux miARNs, miR-675-5p et miR-675-3p (Cai *et al.*, 2006). Le rôle oncogénique de miR-675-5p a été démontré dans plusieurs cancers (Lecerf *et al.*, 2019). Cette étude a donc pour but de comprendre les implications de *H19* et de miR-675-5p dans l'agressivité des cellules cancéreuses mammaires, et notamment au travers d'une expérience de xénotransplantation sur embryons de poissons zèbres.



## 2.2 Article

### Enhancement of Breast Cancer Cell Aggressiveness by lncRNA *H19* and Its Mir-675 Derivative: Insight into Shared and Different Actions

Evodie Peperstraete<sup>3,†</sup>, Clément Lecerf<sup>1,†</sup>, Jordan Collette<sup>1</sup>, Constance Vennin<sup>1</sup>, Ludivine Raby<sup>1</sup>, Pamela Völkel<sup>1</sup>, Pierre-Olivier Angrand<sup>1</sup>, Marie Winter<sup>1</sup>, François Bertucci<sup>2</sup>, Pascal Finetti<sup>2</sup>, Chann Lagadec<sup>1</sup>, Samuel Meignan<sup>1,3</sup>, Roland P. Bourette<sup>1</sup>, Xuefen Le Bourhis<sup>1</sup> and Eric Adriaenssens<sup>1,\*</sup>

University Lille, CNRS, INSERM, CHU Lille, Centre Oscar Lambret, UMR 9020–UMR 1277–Canther–Cancer Heterogeneity, Plasticity and Resistance to Therapies, F-59000 Lille, France; evodie.peperstraete.etu@univ-lille.fr (E.P.); clement.lecerf@univ-lille.fr (C.L.); jorkan62@gmail.com (J.C.); constance.vennin@gmail.com (C.V.); ludivine.raby@gmail.com (L.R.); pamela.voelkel@univ-lille.fr (P.V.); pierre-olivier.angrand@univ-lille.fr (P.-O.A.); marie.winter.etu@univ-lille.fr (M.W.); chann.lagadec@inserm.fr (C.L.); samuel.meignan@inserm.fr (S.M.); Roland.BOURETTE@ibl.cnrs.fr (R.P.B.); xuefen.le-bourhis@univ-lille.fr (X.L.B.)

<sup>1</sup> Laboratoire d'Oncologie Prédictive, CRCM, Institut Paoli-Calmettes, INSERM UMR1068, CNRS UMR7258, Aix-Marseille Université, Département d'Oncologie Médicale, Institut Paoli-Calmettes, 13009 Marseille, France; BERTUCCIF@ipc.unicancer.fr (F.B.); FINETTIP@ipc.unicancer.fr (P.F.)

<sup>2</sup> Tumorigenesis and Resistance to Treatment Unit, Centre Oscar Lambret, F-59000 Lille, France

\* Correspondence: eric.adriaenssens@univ-lille.fr; Tel.: +33-(0)3-20-33-64-06 † The two first authors contributed equally to this work.

Received: 29 May 2020; Accepted: 28 June 2020; Published: 29 June 2020

#### Abstract

Breast cancer is a major public health problem and the leading world cause of women death by cancer. Both the recurrence and mortality of breast cancer are mainly caused by the formation of metastasis. The long non-coding RNA *H19*, the precursor of miR-675, is involved in breast cancer development. The aim of this work was to determine the implication but, also, the relative contribution of *H19* and miR-675 to the enhancement of breast cancer metastatic potential. We showed that both *H19* and miR-675 increase the invasive capacities of breast cancer cells in xenografted transgenic zebrafish models. In vitro, *H19* and miR-675 enhance the cell migration and invasion, as well as colony formation. *H19* seems to induce the epithelial-to-mesenchymal transition (EMT), with a decreased expression of epithelial markers and an increased expression of mesenchymal markers. Interestingly, miR-675 simultaneously increases the expression of both epithelial and mesenchymal markers, suggesting the induction of a hybrid phenotype or mesenchymal-to-epithelial transition (MET). Finally, we demonstrated for the first time that miR-675, like its precursor *H19*, increases the stemness properties of breast cancer cells. Altogether, our data suggest that *H19* and miR-675 could enhance the aggressiveness of breast cancer cells through both common and different mechanisms.

### Introduction

Long non-coding RNAs (lncRNAs) (>200 nt) are essential in cell biology, and their dysfunction plays a critical role in cancer development and progression. Indeed, lncRNAs are involved in diverse cellular processes such as cell proliferation, apoptosis, differentiation and pluripotency, but their mechanisms of action remain largely undeciphered [1]. Among these lncRNAs, *H19*, the first lncRNA discovered located in the *H19/IGF2* locus, is a subject of interest.

*H19* is submitted to genomic imprinting [2] and is expressed during embryonic development. Its expression is repressed after birth, except in a few tissues like the mammary gland, renal gland and uterus [3–6]. Many studies have shown that *H19* promotes tumor phenotypes and induces metastasis in various cancers like gastric, colorectal, bladder, renal, lung and breast cancers but, also, in glioblastoma [7–13]. We have previously shown that *H19* is overexpressed in 70% of breast cancer and promotes the tumorigenic properties of cancer cells [3,14]. The *H19* gene is upregulated by transcription factors such as E2F1 to enhance the cell cycle progression and cell invasion [12].

*H19* can exert its protumorigenic function through diverse molecular mechanisms like the targeting of transcriptional factors or chromatin modifier complexes such as PRC2 (polycomb repressive complex 2) [1]. *H19* binds and recruits the histone methyltransferase EZH2 to the promoter of the proapoptotic gene *BIK* (BCL-2 interacting killer), inducing a reshaping of the chromatin (by trimethylation of the lysine 27 of histone H3) and an inhibition of the *BIK* transcription [15]. *H19* also interacts with microRNAs (regulatory small non-coding RNAs) to serve as a sponge by sequestering them and inhibiting their actions. For instance, *H19* sponges miR-let7 to maintain the breast cancer stem cells status [16]. Moreover, *H19* increases the expression of DNMT1, a DNA methyltransferase, by sponging miR-152, thus inducing the growth and invasion of breast cancer cells [17].

In addition, *H19* could generate two mature miRNAs, miR-675-5p (miR-675) and miR-675-3p (miR-675\*) [18]. These miRNAs mainly act as posttranscriptional repressors by interacting with the mRNA target [19]. *H19*-derived miR-675 favors the tumor progression by repressing several well-known tumor suppressor genes, such as *Rb* [20], *Twist1* [21] or *RUNX1* [22]. We have identified c-Cbl and Cbl-b, two ubiquitin ligase E3, as specific targets of miR-675-5p in breast cancer cells [23].

We have already demonstrated the oncogenic role of the *H19* gene in breast tumorigenesis [14], and *H19*-derived miR-675 has been reported to promote the tumorigenesis of several cancers [20,22,23]. In this study, we examined the involvement and the relative contribution of *H19* and miR-675 in promoting breast cancer cell aggressiveness. Our results indicate that *H19* and miR-675 are in favor of cell migration, invasion and stemness through both common and different mechanisms.

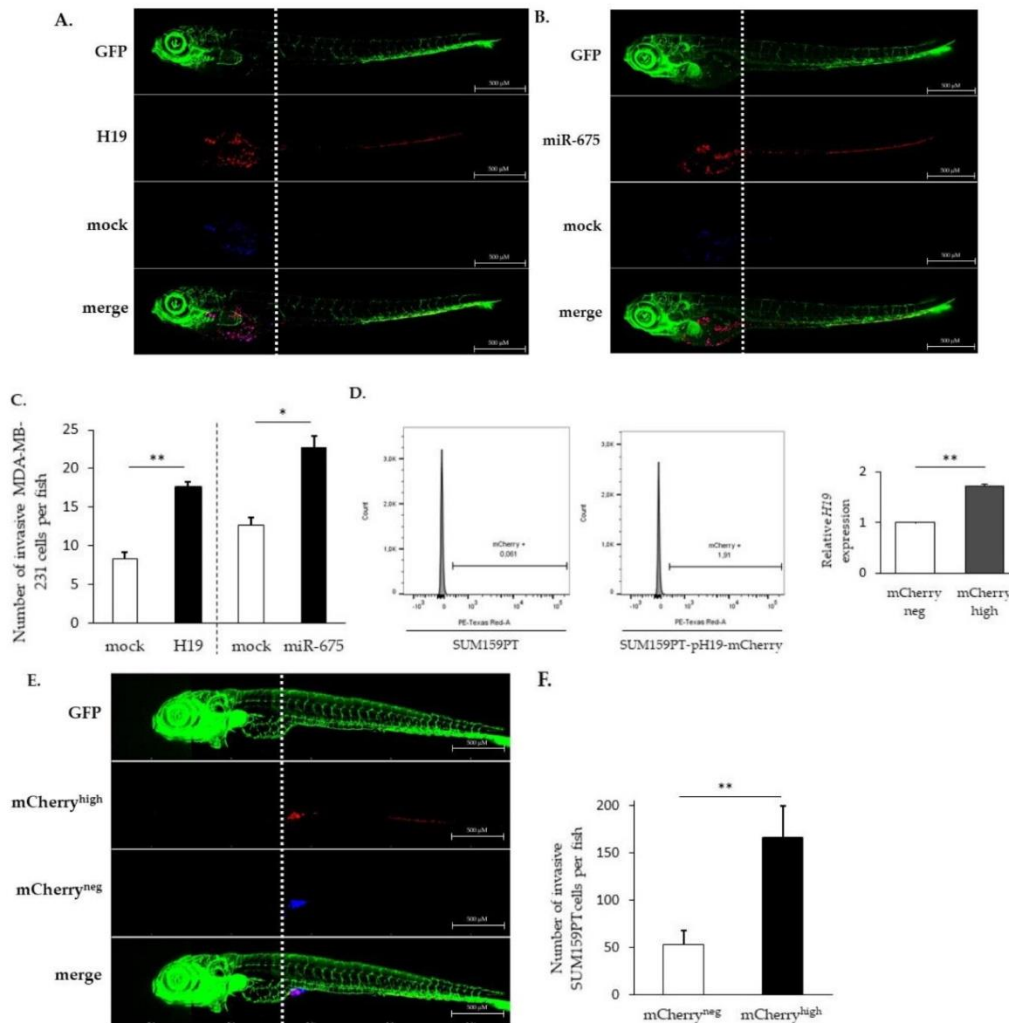
### Results

#### *LncRNA H19 and miR-675 Promote Breast Cancer Cell Invasion in Zebrafish Xenograft Model*

A tumor cell transplantation in zebrafish embryos represents a simple and rapid approach to study a tumor cell invasion and metastasis. The optical transparency of the embryos offers the advantage to monitor cancer cell behavior within a few days after the transplantation [24]. In order to investigate the relative contribution of *H19* and miR-675 in the metastatic process in vivo, breast cancer cells, stained with liposoluble fluorophores, were injected into the yolk sac of transparent transgenic zebrafish embryos having their entire vascular system labeled with green fluorescence, and the invasion of the injected cells was evaluated three days post-injection, as described in Materials and Methods. An increased invasion was observed for MDA-MB-231 breast cancer cells stably overexpressing *H19* or miR-675 compared to the control cells (Figure 1A–C).

In order to highlight the role of the *H19* gene without ectopic overexpression, the SUM159PT breast cancer cell line was transfected with a vector expressing the stable red fluorescent protein mCherry under the control of

the *H19* promoter. This reporter system allowed us to select cells overexpressing *H19* within the total population without modulating its expression.



**Figure 1.** *H19* and miR-675 both promote cancer cell invasion in vivo. (A) Invasive capacities of MDA-MB-231 stably overexpressing *H19* and the control, stained with lipophilic tracers, in transgenic zebrafish. Fluorescent pictures were captured using automated image acquisition software. (B) Invasive capacities of MDA-MB-231 stably overexpressing miR-675 and the control, stained with lipophilic tracers, in transgenic zebrafish. Fluorescent pictures were captured using automated image acquisition software. (C) Quantification of invasive cells per zebrafish. (D) mCherry protein fluorescence in SUM159PT transfected or not with pH19-mCherry plasmid. Fluorescence intensity is categorized in mCherry<sup>neg</sup> and mCherry<sup>high</sup> cellular subpopulations. Relative *H19* expression in those subpopulations is figured. (E) Invasive capacities of mCherry<sup>neg</sup> and mCherry<sup>high</sup> cellular subpopulations, stained with lipophilic tracers, in transgenic zebrafish. Fluorescent pictures were captured using automated image acquisition software. (F) Quantification of invasive cells per zebrafish. For each experiment, forty embryos were used. \* $p < 0.05$  and \*\* $p < 0.01$ .

As expected, the mCherry<sup>high</sup> cell fraction displayed a higher *H19* expression when compared to the mCherry<sup>neg</sup> cell fraction (Figure 1D), validating the established cellular model. Of note, the level of *H19* in mCherry<sup>high</sup> cells may vary according to the experiments (2 to 14 times higher than mCherry<sup>neg</sup> cells). Thus, we were able to isolate cells overexpressing lncRNA *H19* from a heterogeneous cellular population, and we chose a population overexpressing two-fold the *H19* gene to conduct further experiments.

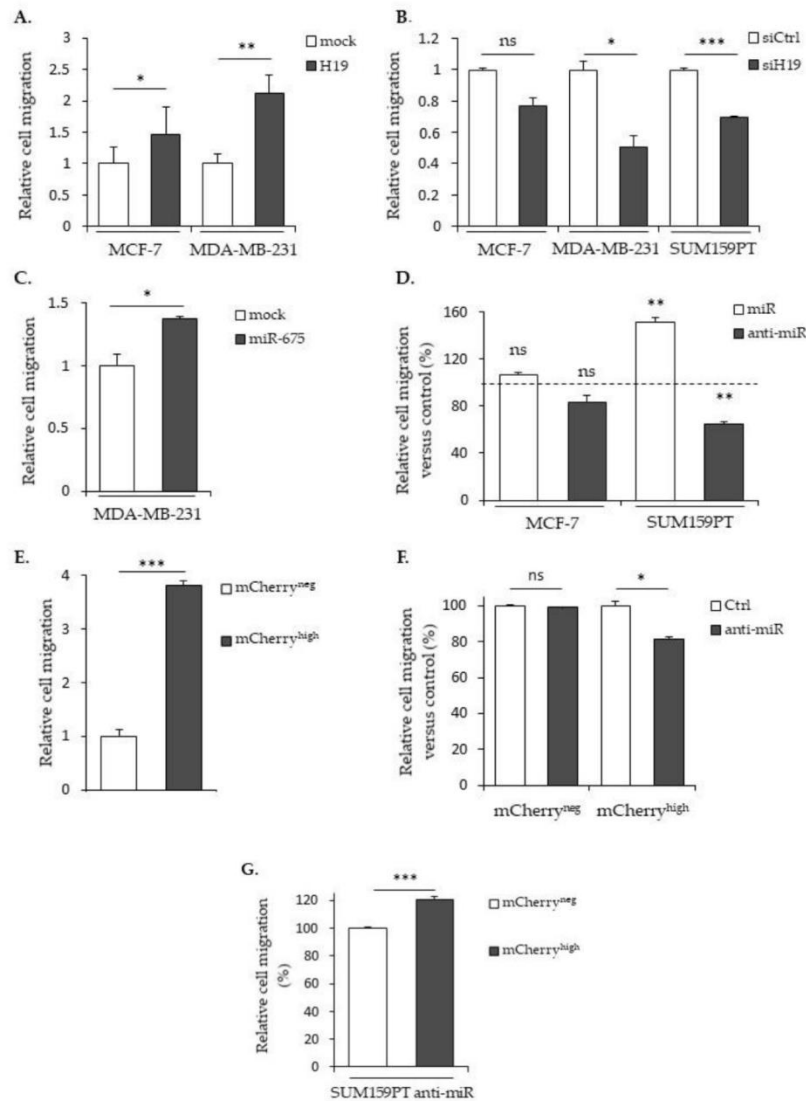
SUM159PT pH19-mCherry<sup>neg/high</sup> cells, stained with liposoluble fluorophores, were coinjected into the yolk sac of transgenic zebrafishes. We showed that mCherry<sup>high</sup> cells move further towards the caudal end through the

dorsal aorta, caudal artery and then caudal vein compared to mCherry<sup>neg</sup> cells (Figure 1E). As shown in Figure 1E,F, mCherry<sup>high</sup> cells were three times more invasive than mCherry<sup>neg</sup> cells, further confirming that cells expressing higher levels of *H19* were more invasive.

The *H19*- and miR-675-enhanced breast cancer cell invasions in vivo prompted us to determine if and how *H19* and miR-675 could affect breast cancer cells in vitro in terms of migration, invasion, epithelial-to-mesenchymal transition (EMT) and stemness, as all these processes are known to be involved in tumor invasion.

***LncRNA H19 and miR-675 Enhance Breast Cancer Cell Migration and Invasion in Collagen Gel***

The migratory capacities of breast cancer cells overexpressing or not *H19* or miR-675 were determined by using Transwell assays. In MCF-7 or MDA-MB-231 cells stably overexpressing *H19* mRNA, the cell migration was significantly increased (Figure 2A). On the contrary, in parental MDA-MB-231 and SUM159PT cells treated with siRNA targeting *H19*, the cell migration was significantly decreased (Figure 2B). In MDA-MB-231 cells stably overexpressing miR-675, the cell migration was significantly increased (Figure 2C). Similarly, parental SUM159PT cells transfected with miR-675 anti-miR exhibited decreased migration (Figure 2D).



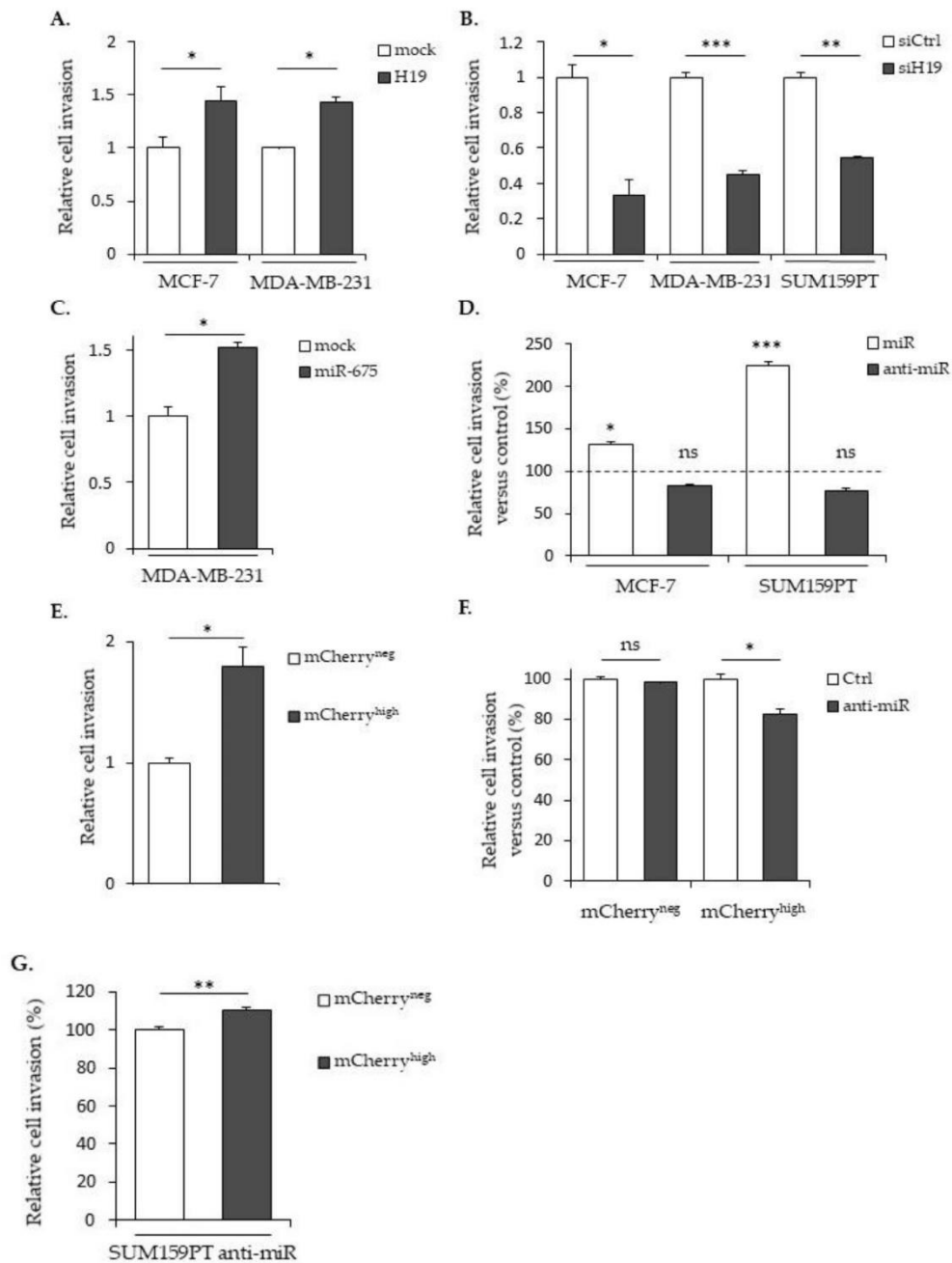
**Figure 2.** *H19* and miR-675 both promote breast cancer cell migrations. (A) Control (mock) or *H19*-stably overexpressing cells (H19) were cultured in Transwell for 24 h. Migrated cells were then incubated with 1-mM Hoescht 33258 and counted. (B) Migratory capacities of the control (siCtrl) and *H19*-knockdown cells (siH19) determined by Transwell assay. (C) MDA-MB-231 control (mock) or miR-675-stably overexpressing cells (miR-675) were cultured in Transwell for 24 h. Migrated cells were then incubated

with 1-mM Hoescht 33258 and counted. (D) Migratory capacities of miR-675-transfected cells (miR) or miR-675-specific inhibitor transfected cells (anti-miR) determined by Transwell assay. Results are presented as a percentage of the control. (E) Relative migratory capacities of SUM159PT-pH19-mCherry<sup>neg</sup> versus SUM159PT-pH19-mCherry<sup>high</sup> determined by Transwell assay. (F) Migratory capacities of miR-675-specific inhibitor transfected cells (anti-miR). Results are presented as a percentage of the control. (G) Migratory capacities of miR-675-specific inhibitor transfected cells (anti-miR). Results are presented as a percentage of the mCherry<sup>neg</sup> condition. \*  $p < 0.05$ ; \*\*  $p < 0.01$ ; \*\*\*  $p < 0.001$ ; ns: not significant.

Then, we checked the migration abilities in the SUM159PT pH19-Cherry model. We observed a three-fold increase of the migration with the mCherry<sup>high</sup> fraction compared to the mCherry<sup>neg</sup> fraction (Figure 2E). As expected, the use of a miR-675 inhibitor reduced 25% of the migration in the mCherry<sup>high</sup> fraction compared to the mCherry<sup>neg</sup> control cells (Figure 2F). However, when we compared the cell migration with the miR-675 inhibitor in the mCherry<sup>neg</sup> and mCherry<sup>high</sup> fractions, we observed an increase of the migration in the presence of the miR-675 inhibitor in mCherry<sup>high</sup> (Figure 2G). This suggests that, with a two-fold gene expression, the lncRNA *H19* per se is able to increase the migration independently of miR-675.

We next determined the role of *H19* and miR-675 in the cell invasion by using Transwell previously coated with collagen. In MCF-7 or MDA-MB-231 cells stably overexpressing *H19* mRNA, the cell invasion was significantly increased compared to the mock cells (Figure 3A). Accordingly, *H19* knockdown using siRNA reduced the invasive capacities of parental MCF-7, MDA-MB-231 or SUM159PT cells (Figure 3B). The miR-675 expression significantly increased the invasion, as observed in MDA-MB-231 cells stably overexpressing miR-675 (Figure 3C) or in MCF-7 and SUM159PT cells transfected with miR-675 mimics (Figure 3D). The invasion of mCherry<sup>high</sup> cells was also increased when compared to mCherry<sup>neg</sup> cells (Figure 3E). The miR-675 inhibitor diminished the invasion of mCherry<sup>high</sup> cells compared to the control cells (Figure 3F), highlighting the involvement of miR-675 in the *H19*-induced invasion. However, in the same way as for the cell migration, we observed an increase of the invasion in the presence of the miR-675 inhibitor in mCherry<sup>high</sup> (Figure 3G), suggesting that the lncRNA *H19*-induced invasion does not require miR-675 action.

All together, these data show that both *H19* and miR-675 are able to promote the migration and invasion of breast cancer cells.



**Figure 3.** *H19* and miR-675 participate in the breast cancer cells invasion. (A) Control (mock) or *H19*-stably overexpressing cells (H19) were cultured in Transwell for 24 h. Invasive cells were then incubated with 1-mM Hoescht 33258 and counted. (B) Invasive capacities of the control (siCtrl) and *H19*-knockdowned cells (siH19) determined by Transwell assay. (C) MDA-MB-231 control (mock) or miR-675-stably overexpressing cells (miR-675) were cultured in Transwell for 24 h. Invasive cells were then incubated with 1-mM Hoescht 33258 and counted. (D) Invasive capacities of miR-675-transfected cells (miR) or miR-675-specific inhibitor transfected cells (anti-miR) determined by Transwell assay. Results are presented as the percentage of the control. (E) Relative invasive capacities of SUM159PT-pH19-mCherry<sup>neg</sup> versus SUM159PT-pH19-mCherry<sup>high</sup> determined by Transwell assay. (F) Invasive capacities of miR-675-specific inhibitor transfected cells (anti-miR). Results are presented as the percentage of the control. (G) Invasive capacities of miR-675-specific inhibitor transfected cells (anti-miR). Results are presented as a percentage of the mCherry<sup>neg</sup> condition. \*  $p < 0.05$ ; \*\*  $p < 0.01$ ; \*\*\*  $p < 0.001$ ; ns: not significant.

*LncRNA H19 and miR-675 Differentially Regulate the Expression of Epithelial and Mesenchymal Markers*

Enhanced mobility and invasive abilities are required for cancer cells to invade the surrounding tissues and promote metastatic development. These properties are also part of the phenotype changes due to the epithelial-to-mesenchymal transition (EMT). The EMT is a process in which cells lose their epithelial characteristics to acquire new properties of mesenchymal cells. During the EMT, epithelial cells undergo dramatic molecular modifications with the downregulation of tight- and adherent-junction proteins such as ZO-1 and E-cadherin and the upregulation of specific markers of mesenchymal cells, including N-cadherin, vimentin and fibronectin. The expression of these markers is regulated by numerous transcriptional factors such as TCF8/ZEB1 and Snail. In a first approach to study the EMT process, we measured, by Western blot, the expression of molecular markers of epithelial or mesenchymal phenotypes (Figure 4).

ZO-1 is a protein that forms tight junctions. It binds transmembrane proteins to the actin cytoskeleton. E-cadherin participates in intramembrane junctions to modify the cell morphology by modulating the cytoskeleton. The TCF8/ZEB1 transcription factor inhibits adhesion junctions by suppressing the expression of E-cadherin, whereas the Snail transcription factor suppresses ZO-1 expression. Vimentin is of mesenchymal origin; it modulates the structural dynamics and reorganizes the intermediate filaments. In MCF-7 cells displaying an epithelial phenotype, the expression of mesenchymal markers is not detected in whatever *H19* level. However, ZO-1 and E-cadherin proteins (epithelial markers) are less expressed in MCF-7 cells overexpressing *H19* compared to control cells (Figure 4A). In SUM159PT cells presenting less epithelial phenotypes than MCF-7 cells, a decrease of the ZO-1 expression and an increase of N-cadherin, vimentin, and Snail protein levels were observed in the mCherry<sup>high</sup> cells compared to the mCherry<sup>neg</sup> cells (Figure 4B). Interestingly, the expressions of mesenchymal markers (TCF8/ZEB1, N-cadherin, vimentin and Snail) were further enhanced in MDA-MB-231 cells stably overexpressing *H19* (Figure 4C). Together, these data indicate that, whatever the initial phenotypes of the cells, high levels of *H19* promote the EMT by decreasing the expression of epithelial markers and/or increasing the expression of mesenchymal markers.

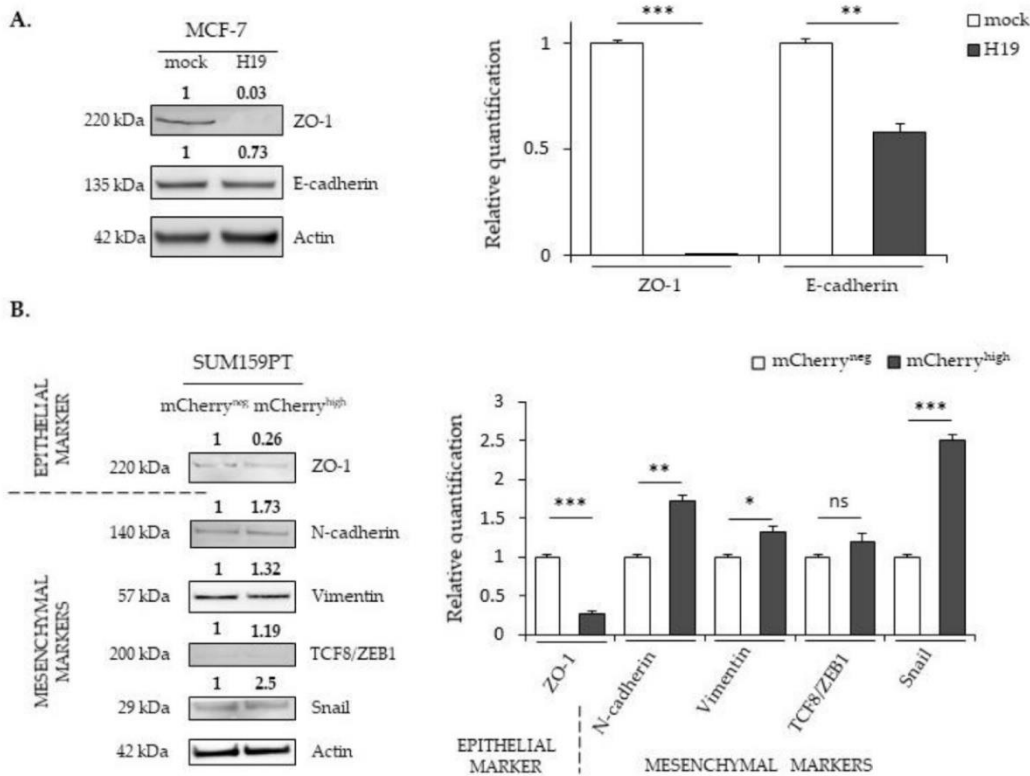
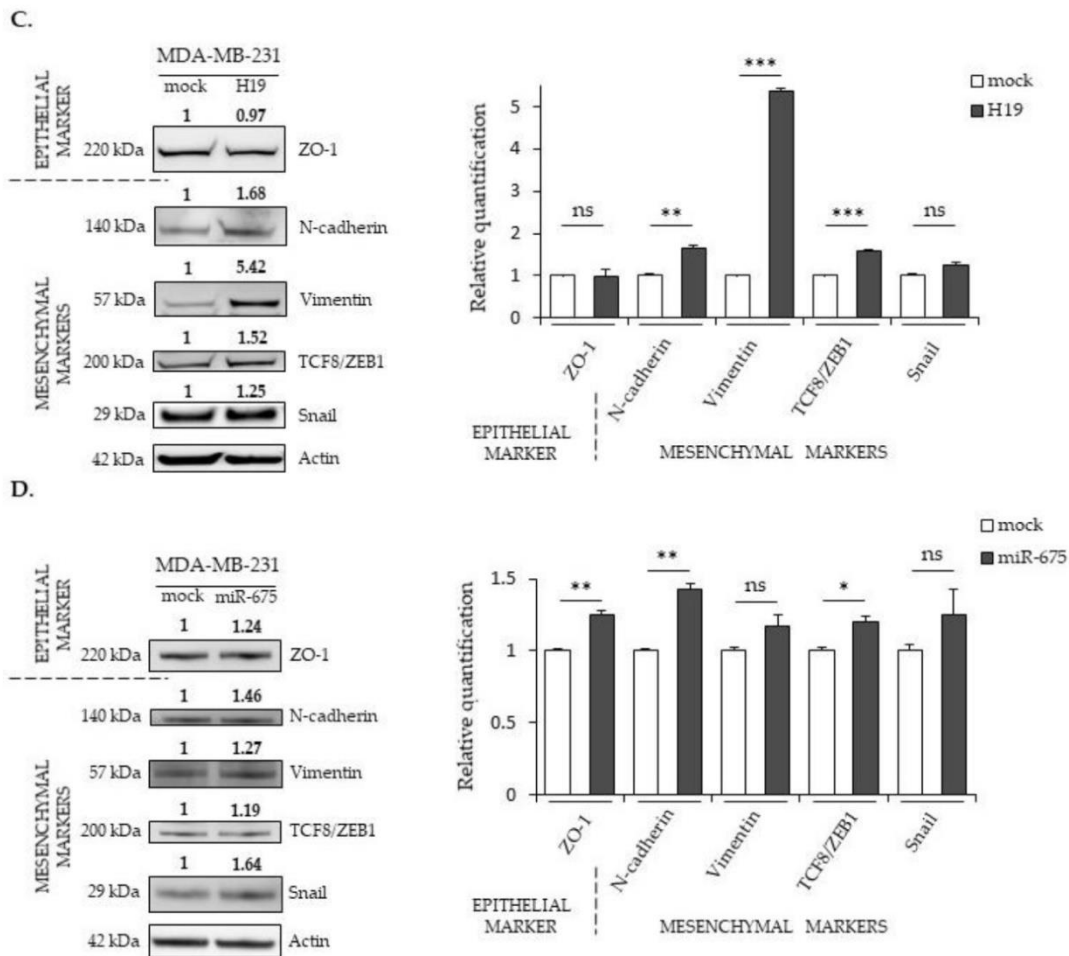


Figure 4. Cont.





**Figure 4.** The effects of *H19* and miR-675 on the expressions of the epithelial-to-mesenchymal transition (EMT) markers. (A) EMT protein expressions in the MCF-7 control (mock) or *H19*-stably overexpressing cells (*H19*), determined by Western blot analysis. (B) EMT protein expressions in SUM159PT-pH19-mCherry<sup>neg</sup> (mCherry<sup>neg</sup>) or SUM159PT-pH19-mCherry<sup>high</sup> (mCherry<sup>high</sup>), determined by Western blot analysis. (C) EMT protein expressions in the MDA-MB-231 control (mock) or *H19*-stably overexpressing cells (*H19*), determined by Western blot analysis. (D) EMT protein expressions in the MDA-MB-231 control (mock) or miR-675-stably overexpressing cells (miR-675), determined by Western blot analysis. For each panel, actin was used as the equi-loading control. The relative signal intensities were quantified by ImageJ and shown above the protein bands for the representative experiment figure. The quantification of the triplicate is figured in the graph beside. \*  $p < 0.05$ ; \*\*  $p < 0.01$ ; \*\*\*  $p < 0.001$ ; ns: not significant. Uncropped blots are shown in Figures S1–S22.

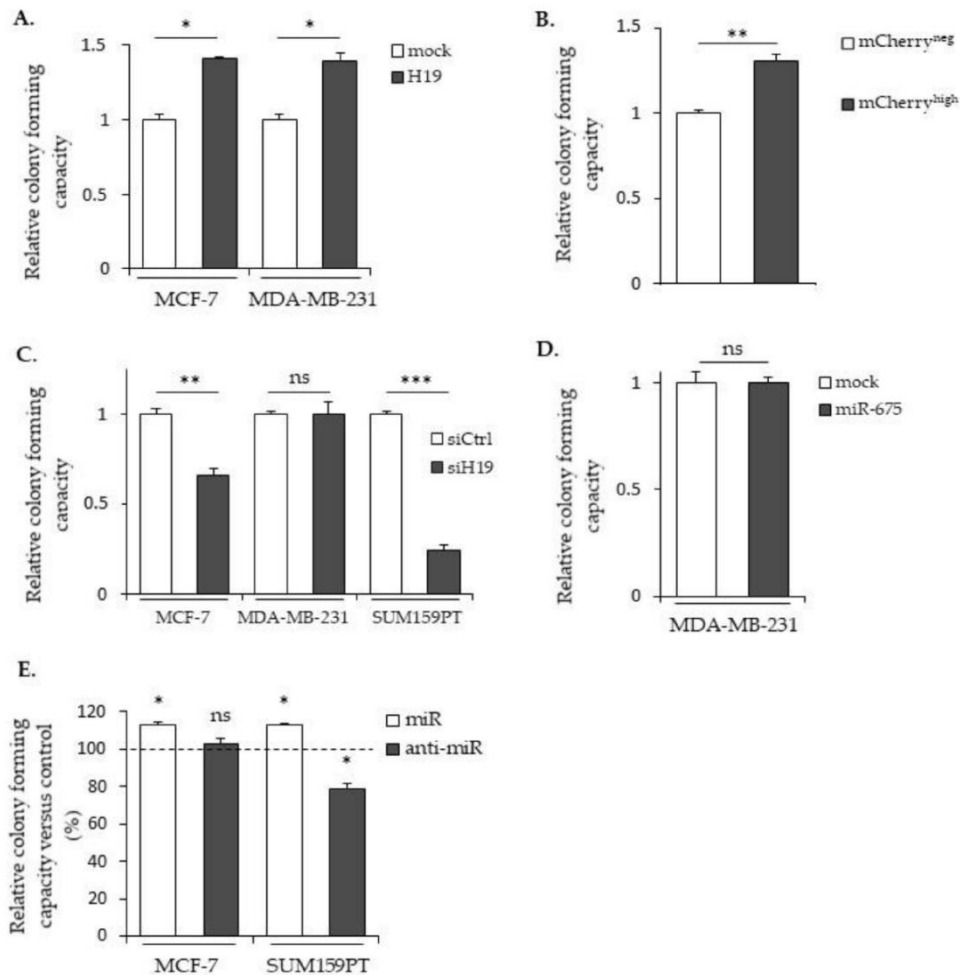
The expression profile of the EMT markers in MDA-MB-231 cells stably overexpressing miR-675 was quite surprising, with an increased expression for both the epithelial marker ZO-1 and the mesenchymal markers TCF8/ZEB1 and N-cadherin (Figure 4D). In addition, vimentin and Snail expressions were not significantly modified (Figure 4D). These modifications suggest a hybrid phenotype induced by miR-675.

#### *LncRNA H19 and miR-675 Promote Breast Cancer Colony Formation*

The clonogenic assay is an *in vitro* cell survival assay based on the ability of a single cell to survive and grow into a colony, which is indicative of cancer cell aggressiveness. We thus performed clonogenicity assays with cells expressing different levels of *H19* and miR-675 (Figure 5). As shown in Figure 5A, MCF-7 and MDA-MB-231 cells stably overexpressing *H19* formed more colonies compared to the controls. Similar results were observed in mCherry<sup>high</sup> compared to mCherry<sup>neg</sup> SUM159PT cells (Figure 5B). Furthermore, *H19* knockdowns using siRNA induced a decrease in colony formation of native MCF-7 and of SUM159PT cell lines (Figure 5C). These results



indicate that *H19* is able to enhance the clonogenicity of breast cancer cells. The effect of miR-675 in colony formations varied according to the cell lines used: MDA-MB-231 cells stably overexpressing miR-675 displayed a similar colony formation ability when compared to the control cells (Figure 5D). In contrast, miR-675 mimics increased colony formations in MCF-7 and SUM159PT cells, while anti-miR-675 decreased colony formations only in SUM159PT cells (Figure 5E).



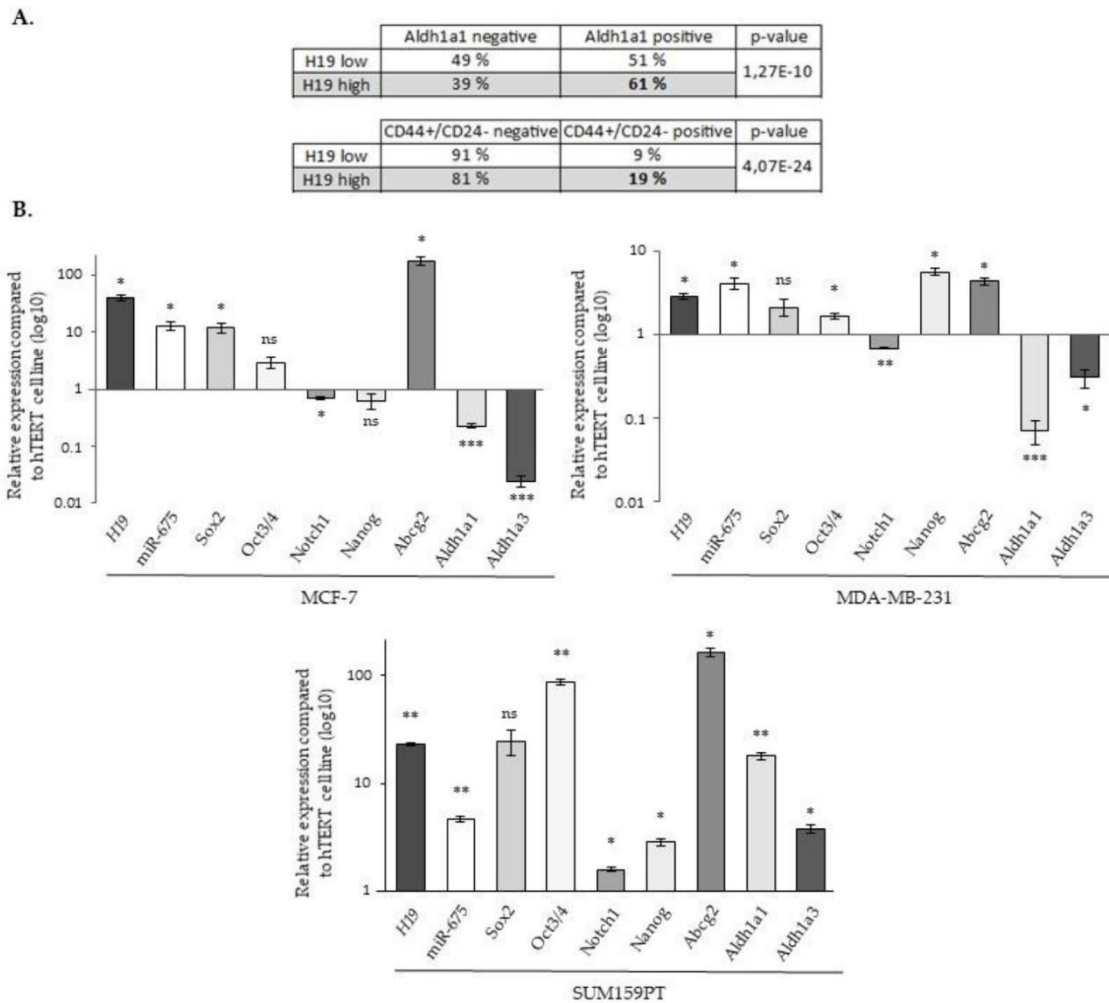
**Figure 5.** Effects of *H19* and miR-675 in colony formations. (A) The clonogenic capacities of the control (mock) or *H19*-overexpressing cells (*H19*). (B) The relative clonogenic capacities of SUM159PT-pH19-mCherry<sup>neg</sup> versus SUM159PT-pH19-mCherry<sup>high</sup>. (C) The clonogenic capacities of the control (siCtrl) and *H19*-knockdown cells (siH19). (D) The clonogenic capacities of the MDA-MB-231 control (mock) or miR-675-overexpressing cells (miR-675). (E) The clonogenic capacities of the miR-675-transfected cells (miR) or miR-675-specific inhibitor transfected cells (anti-miR). Results are presented as a percentage of the control. \*  $p < 0.05$ ; \*\*  $p < 0.01$ ; \*\*\*  $p < 0.001$ ; ns: not significant.

### *LncRNA H19 and miR-675 Enhance the Stemness Properties of Breast Cancer Cells*

To further investigate the involvement of lncRNA *H19* and miR-675 in breast cancer cell aggressiveness, we then investigated if lncRNA *H19* and miR-675 could endow the cells with stemness properties. For this, a meta-analysis on the gene expression data of more than 5000 breast tumor samples [25] was first carried out to explore the correlation of *H19* expression and that of the stem cell gene signature. Two published gene signatures were used, one obtained with cells expressing the ALDH1a1 isoform [26] and the other with cells harboring a combination of the overexpression of CD44 and a low expression of CD24 (CD44+/CD24-) [27]. Tumors were classified as *H19*<sup>high</sup> and *H19*<sup>low</sup> according to the expression of *H19* higher or lower, respectively, to the median expression value. For

## PARTIE II : LE POISSON ZEBRE EN CANCEROLOGIE

each stem cell signature tested, we observed a higher expression of *H19* in tumors expressing the signature, suggesting that lncRNA *H19* may play a role in breast cancer stem cells (Figure 6A).



**Figure 6.** The expression of *H19*, *miR-675* and different stem cell markers. (A) The expression of the *H19* gene in mammary tumors expressing or not stem cell signatures. The table represents the *H19* gene expression dependent on the gene signature of the tumors. Two gene signatures were used, one obtained in cell expressing ALDH1A1 and the other in CD44+/CD24-. (B) The relative expression of *H19*, *Sox2*, *Oct3/4*, *Notch1*, *Nanog*, *Abcg2*, *Aldh1a1* and *Aldh1a3* genes and *miR-675* in MCF-7, MDA-MB-231 and SUM159PT cells. The expression levels were related to the expression levels in hTERT cells indexed to 1. For panel A, a Fisher's exact test was performed. \*  $p < 0.05$ ; \*\*  $p < 0.01$ ; \*\*\*  $p < 0.001$ ; ns: not significant.

We then analyzed by RT-qPCR the expression of the different markers involved in the cancer stem cells' biology (*Sox2*, *Oct3/4*, *Notch1*, *Nanog*, *Abcg2*, *Aldh1a1* and *Aldh1a3*). As expected, a correlation was found between *H19* and *miR-675*. Of note, although the expressions of *H19* and *miR-675* were generally correlated with that of cancer stem cell markers, some differences existed depending on the considered cell line. For example, diminutions of *Notch1*, *Aldh1a1* and *Aldh1a3* expressions were observed in the MCF-7 and MDA-MB-231 cell lines, contrary to SUM159PT cells. Similarly, the *Nanog* expression was increased in MDA-MB-231 and in SUM159PT cells but not in MCF-7 cells (Figure 6B).

The sphere-forming capacity of cells is used as the gold standard in an in vitro functional assay to analyze the ability of cancer stem cells to proliferate under anchorage-independent conditions in a defined medium [28]. As shown in Figure 7A, the MCF-7 and MDA-MB-231 cells stably overexpressing *H19* formed more spheres compared to the controls. Similarly, the ability of cells to form spheres was also increased in mCherry<sup>high</sup> SUM159PT

cells compared to mCherry<sup>neg</sup> ones (Figure 7B). A decrease of the sphere-forming capacity was observed in the native three cell lines when *H19* was knocked down with siRNA (Figure 7C), confirming that *H19* was able to enhance the sphere-forming capacity of breast cancer cells. On the other hand, MDA-MB-231 cells stably overexpressing miR-675 also formed more mammospheres (Figure 7D). Similarly, the miR-675 mimic increased and the miR-675 inhibitor decreased the sphere-forming capacity of native MCF-7 and SUM159PT cells (Figure 7E), indicating that miR-675 was also involved in the sphere-forming capacity of breast cancer cells. Interestingly, the miR-675 mimic did not further increase the sphere formation in MCF-7 and MDA-MB-231 cells overexpressing *H19*. However, the anti-miR transient expression induced a decrease of the sphere formation by 70% in MCF-7 cells (Figure 7F) and by 60% in MDA-MB-231 cells *H19* (Figure 7G). Complementarily, we performed an ALDEFUOR assay on the intact and viable cells after the siRNA *H19*, miR-675 mimic and miR-675 inhibitor transient expressions. The siRNA *H19* induced a diminution of the ALDH<sup>high</sup> subpopulation compared to the control cells, suggesting that *H19* enriched for breast cancer stem cells. Similar trends, though not significant, were observed in cells that were transfected with the miR-675 mimic or inhibitor (Figure 7H).

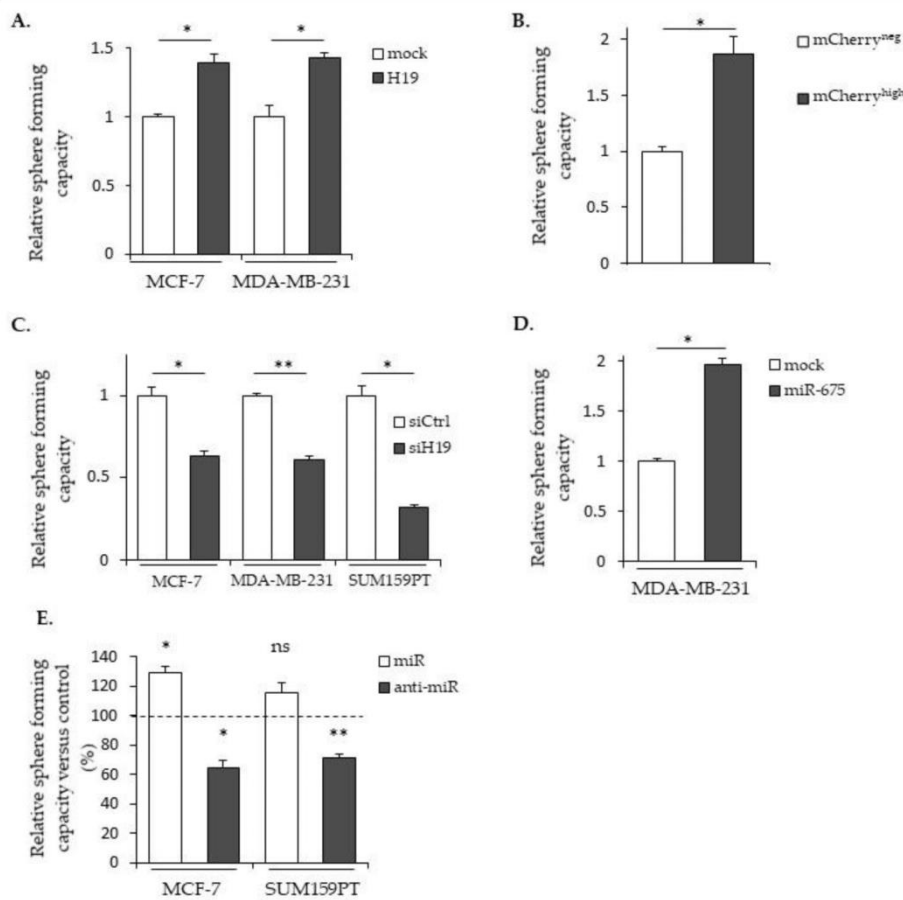
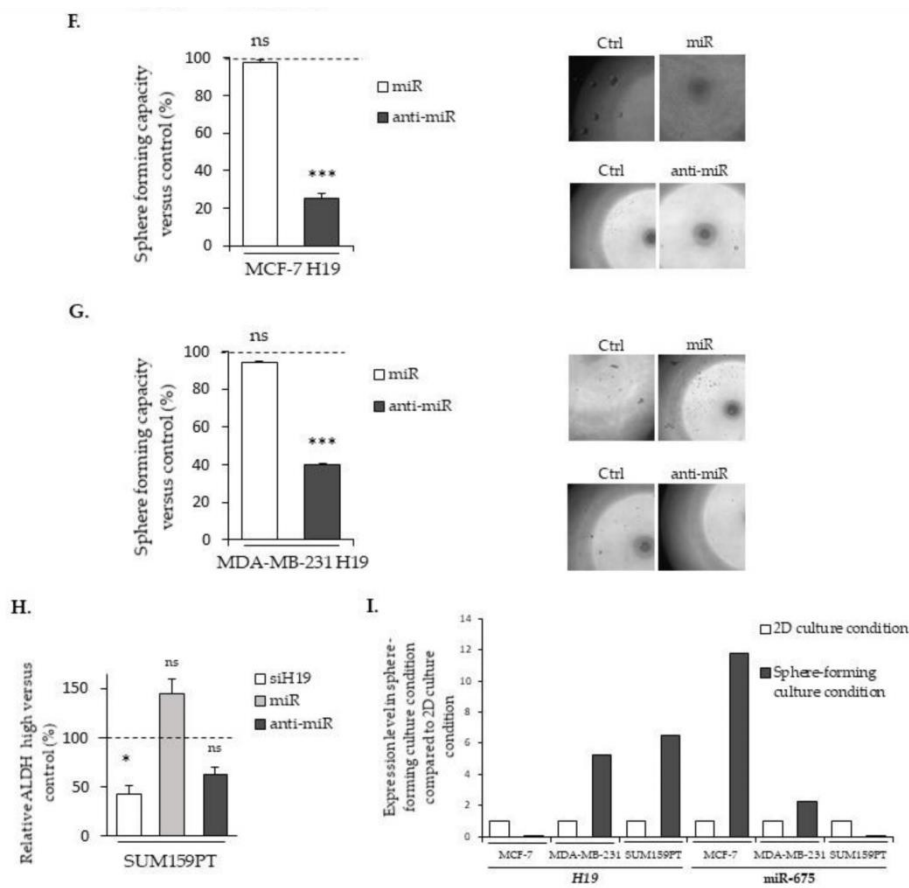


Figure 7. Cont.



**Figure 7.** *H19* and miR-675 enhance the sphere formation of breast cancer cells. (A) Sphere-forming capacities of the control (mock) or *H19*-stably overexpressing cells (H19). (B) The relative sphere-forming capacities of SUM159PT-pH19-mCherry<sup>neg</sup> versus SUM159PT-pH19-mCherry<sup>high</sup>. (C) The sphere-forming capacities of the control (siCtrl) and *H19*-knockdown cells (siH19). (D) The sphere-forming capacities of the MDA-MB-231 control (mock) or miR-675-stably overexpressing cells (miR-675). (E) The sphere-forming capacities of miR-675-transfected cells (miR) or miR-675-specific inhibitor transfected cells (anti-miR). Results are presented as a percentage of the control. (F) The sphere-forming capacities of MCF-7 *H19*-stably overexpressing cells transfected with miR-675 (miR) or the miR-675-specific inhibitor (anti-miR). Results are presented as a percentage of the control. Representative pictures for each condition are shown. (G) The sphere-forming capacities of MDA-MB-231 *H19*-stably overexpressing cells transfected with miR-675 (miR) or the miR-675-specific inhibitor (anti-miR). Results are presented as a percentage of the control. Representative pictures of each condition are shown. (H) ALDEFLUOR-positive subpopulations defined by the ALDEFLUOR assay in *H19*-knockdown cells (siH19), miR-675-transfected cells (miR) or miR-675-knockdown cells (anti-miR). Results are presented as a percentage of the ALDEFLUOR-positive subpopulation in the native cells. (I) The expression levels of *H19* and miR-675 in the cell lines cultured in sphere-forming conditions versus the same cell lines cultured in 2D conditions. \*  $p < 0.05$ ; \*\*  $p < 0.01$ ; \*\*\*  $p < 0.001$ ; ns: not significant.

Enhanced sphere formations by *H19* and miR-675 prompted us to determine their expressions in native cells under culture conditions favoring stem cell proliferations. We observed an increased expression of *H19* in MDA-MB-231 and SUM159PT and an increased expression of miR-675 in MCF-7 in these conditions when compared to a monolayer culture (Figure 7I).

Collectively, these results demonstrate the involvement of *H19* and miR-675 in the enrichment of cancer stem cells.

## Discussion

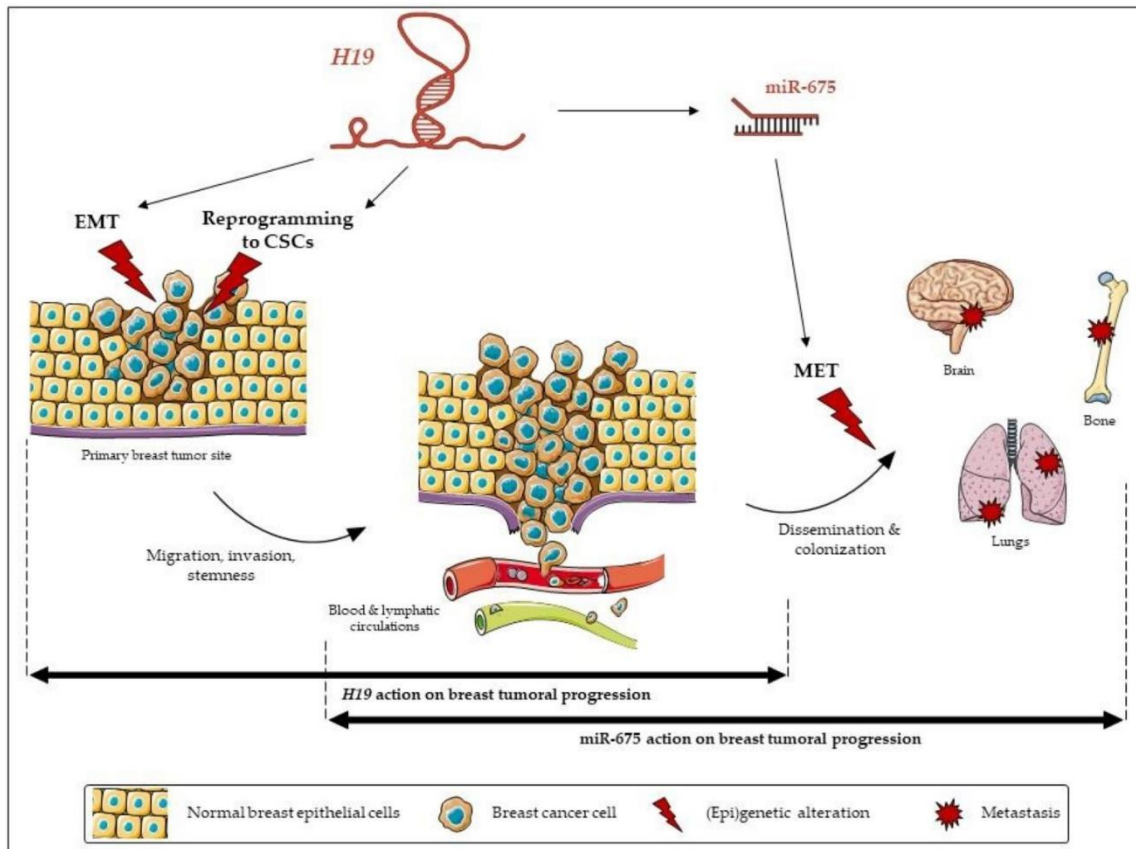
The long non-coding RNA *H19* has been shown to intervene at multiple steps of tumorigenesis, such as cell proliferation, migration, invasion, induction of the epithelial-to-mesenchymal transition and metastasis [13].

However, several data indicate that *H19*-derived miR-675 could also be implicated in these oncogenic cellular processes [20,22,23]. Our work focuses on the roles of both *H19* and miR-675 in metastasis-related phenotypes such as migration, invasion, the EMT, colony formation and stemness properties.

We assessed the breast cancer invasion by the transplantation of cells in zebrafish embryos. Indeed, cell lines from a variety of human cancers, including metastatic melanoma, pancreatic, ovarian, breast, glioma and colorectal cell lines, are capable of proliferating, invading and forming tumor masses in the zebrafish embryo during xenotransplantation studies [29,30]. The Tg(*fli1*:GFP) transgenic line used in this study allowed us to visualize blood vessels in the living host and to allow for the analysis of cancers cells after their intravasation into the blood vessels. We demonstrated that both *H19* and miR-675 overexpressing cells were able to more efficiently colonize zebrafish.

In vitro, *H19* and miR-675, alone or together, participate in cellular migration and invasion; migration and invasion are enhanced by both the miR-675 mimic transient expression and miR-675 overexpression but, also, by *H19* overexpression. On the contrary, decreases in the migration and invasion are observed upon *H19*-silencing or miR-675-inhibitor transient expressions. However, we observed that the miR-675 inhibitor has different effects according to the cell types, probably reflecting their basal levels of *H19* expressions or their molecular classifications. Indeed, the MCF-7 cell line is a luminal-like subtype, whereas the MDA-MB-231 cell line is a triple-negative subtype. To reinforce this hypothesis, it has been shown that *H19* is associated with a poor prognosis in triple-negative breast cancer patients [31].

Both *H19* and miR-675 are able to modulate the expressions of the EMT markers, although their actions differ according to the cell types. In the MCF-7 cell line, *H19*-overexpressing cells displayed decreased expressions of epithelial markers. By contrast, in MDA-MB-231 and in SUM159PT pH19-mCherry<sup>high</sup> cells, high levels of *H19* were associated with increased expressions of mesenchymal markers, and, very interestingly, SUM159PT pH19-mCherry<sup>high</sup> cells also displayed decreased expressions of the epithelial marker ZO-1. Our results are in agreement with the work of Liao *et al.*, who demonstrated that the lncRNA *H19* induces the proliferation and invasion of lung cancer cells via the overexpression of N-cadherin and vimentin and the decrease of E-cadherin [32]. In MDA-MB-231 cells, miR-675 promotes ZO-1 and N-cadherin expressions. These results highlight an interesting difference of action between *H19* and miR-675. In the same cell line, *H19* favors the expression of EMT markers, whereas the action of miR-675 seems ambiguous by favoring the expression of both mesenchymal and epithelial markers, suggesting the role of miR-675 in both the EMT and mesenchymal-to-epithelial transition (MET). Indeed, the miR-675 alone does not induce the same molecular determinants as *H19*. These data reveal that *H19* and miR-675 participate in both the migration and invasion, yet *H19* and miR-675 do not exert the same role during the molecular events of the EMT. The MET is essential to colonize and proliferate in the different stages of metastasis [33]. Thus, *H19* and miR-675 could have sequential effects leading to metastasis; *H19* would favor the departure of cells from the primary tumor by promoting the EMT, while miR-675 would also favor a metastatic colonization and the development of secondary tumors by inducing the MET (Figure 8). Alternately, an intermediate status between the epithelial and mesenchymal phenotypes, i.e., the hybrid epithelial-mesenchymal state, is increasingly described to be involved in migration and invasion. The hybrid phenotype endows cancer cells with a more plastic status to adapt the stressful environment for the metastasis formation [34]. Indeed, the hybrid phenotype may contribute to the cancer collective cell migration and, in fine, to metastasis [35]. Our data suggested that miR-675 could favor the acquisition of this hybrid phenotype to favor metastases.



**Figure 8.** The relative contribution of long non-coding (lnc)RNA *H19* and its miR-675 in breast cancer progression.

The epithelial-to-mesenchymal transition is an important cellular mechanism corresponding to a differentiation from an epithelial cell to a mesenchymal-like cell. It involves multiple molecular pathways [36]. The EMT appears essential for embryonic development and tissue repair but, also, for the progression of cancer and metastases [37]. The role of miR-675 in the EMT is not described, and the role of *H19* in the EMT appears contradictory. Numerous data indicate that *H19* promotes the EMT [38–40] in various tissues, including the breast [41], but Zhang and collaborators demonstrated that *H19* is implied in the metastasis suppression of hepatocellular carcinoma [42]. Our results showing the contrary effects of *H19* and miR-675 could explain these discrepancies.

As previously indicated, the lncRNA *H19* can act at multiple levels of regulation, more particularly by sponging microRNAs. For instance, Lv *et al.* demonstrated that the upregulation of *H19* promotes the migration and invasion in bladder cancer by sponging miR-29b-3p. However, the knockdown of *H19* allowed miR-29b-3p to facilitate the MET [43]. It is also known that the miR-200 family (miR-200a, miR-200b, miR-200c, miR-141 and miR-429) is involved in the EMT via the ZEB1/E-cadherin pathway [44,45]. Indeed, these microRNAs regulate ZEB1 and ZEB2 proteins through the repressed E-cadherin expression. ZEB1 and ZEB2 are transcriptional factors implicated in metastasis. When the miR-200 family is inhibited, the EMT is induced, and, conversely, the expression of these microRNAs leads to the MET [46]. Many studies show that *H19* inhibits members of miR-200 family [47]. Indeed, *H19* sponges miR-200a or miR-200b/c to promote cancer metastasis through ZEB1 and ZEB2 upregulations [48,49].

Complementarily, we explored the role of *H19* and miR-675 in the resistance to anoikis. Anoikis is a cell death induced by the detachment of a cell from the extracellular matrix. This cellular phenotype is involved in the metastatic development; indeed, during invasion, cells have to break their adhesions so that they can reach the secondary tumor site. To evaluate the role of *H19* and miR-675 in this process, we studied the resistance to anoikis in breast cancer cell lines overexpressing *H19* or miR-675. The anoikis resistance assay in MCF-7 cells stably overexpressing or not *H19* does not show any long-term statistical differences. In the same way, the viability of



MDA-MB-231 cells overexpressing *H19* or miR-675 does not show any significant variation compared to the control.

In connection with the phenotypes linked to the EMT, we further investigated the involvement of *H19* in the regulation of breast cancer stem cells. Indeed, *H19* has already been associated with stemness in breast cancer [16,31,50], in part by favoring symmetric division. In this report, we highlight a correlation between the expression of the *H19* gene and the presence of stem cell markers in a cohort of more than 5000 breast cancer clinical samples. A gene signature of cells expressing two stemness markers, CD44+/CD24- and ALDH1A1 [51,52], is correlated with the *H19* expression. In addition, our results indicate that the overexpression of *H19* and miR-675 observed in cancer cells is accompanied overall with an overexpression of different stem cell markers such as *Sox2*, *Oct3/4* and *Abcg2* in breast cancer cell lines. We observed an increase of *Nanog* only in MDA-MB-231 and SUM159PT triple-negative breast cancer cells and an increase of *Notch1*, *Aldh1a1* and *Aldh1a3* only in the SUM159PT cell line. In these latter cells, the activity of ALDH is decreased with siRNA-targeting *H19* and with the miR-675 inhibitor but is enhanced by the transient expression of the miR-675 mimic. To our knowledge, it is the first time that the role of miR-675 on cancer cells' stemness has been described. *Nanog*, *Sox2* and *Oct3/4* are transcription factors largely described in the literature for their involvement in the maintenance of self-renewal and pluripotency of embryonic stem cells [53]. Their implication in cancer stem cells' regulation has also been demonstrated [54]. *Abcg2* is an efflux pump associated with the drug resistance of cancer stem cells [54]. A transcriptional regulatory network involving *Sox2*, *Oct3/4* and c-Myc ensures the maintenance of the stem cell pluripotency [53]. Interestingly, c-Myc has been shown to bind alleles specifically upstream of the promoter of the *H19* gene in order to promote its transcription [55]. Similarly, *Sox2* and *Oct3/4* are able to bind upstream of the *H19* gene [56]. This fixation prevents the promoter methylation on the maternal allele and, therefore, contributes to the gene expression. The control of the *H19* gene by these three factors suggests that *H19* may be a factor involved in the stem cell regulation.

It is often described that miRNAs have a similar role to the host genes that produce it. However, in several cases, miRNAs have antagonistic actions against its host gene [57]. For instance, in breast cancer, miR-301 derived from *SKA2* gene, which has an oncogenic action, negatively regulates tumor suppressors such as PTEN [58]. Conversely, miR-483-5p, derived from the *IGF2* gene, inhibits angiogenesis, unlike its host gene [59]. In our work, we show a similar role and/or antagonist actions of miR-675 and its host gene *H19*, depending on the phenotype studied. Indeed, *H19* and miR-675 have similar functions during the cellular migration and invasion and seem to regulate stemness in the same manner. By contrast, their role would be antagonistic during the epithelial-to-mesenchymal transition and colony formation.

We show that miR-675 modulates the phenotypes associated with the occurrence of metastases in breast cancer cells. However, miR-675 generates two mature miRNAs (miR-675-5p and miR-675-3p), both of which have different targets [60]. A logical continuation of this work will therefore be to determine which mature miRNA is associated with these phenotypes.

In continuation of this investigation, it would be interesting to study the molecular targets involved in the occurrence and the development of breast cancer metastases. Despite the similarity of the phenotypes observed, lncRNA *H19* and miR-675 may not have the same molecular determinants and, therefore, use distinct signaling pathways. lncRNA *H19* induces the cell migration and invasion by different mechanisms, depending on the cancers. Indeed, *H19* promotes the migration and invasion of colon cancer cells via the MAPK signaling pathway [61] and those of human osteosarcoma through the NF- $\kappa$ B pathway [62]. It would be relevant to investigate if these signaling pathways are similarly activated in breast cancer cells, as a better knowledge of the molecular determinants associated with the *H19* and miR-675 phenotypes in breast cancer progression would provide new therapeutic opportunities.

### Materials and Methods

#### *Cell Culture*

The breast cancer cell lines MCF-7 estrogen-sensitive, MDA-MB-231 and SUM159PT estrogen-insensitive were maintained routinely in Dulbecco's modified Eagle's medium (DMEM, Gibco), Roswell Park Memorial Institute medium (RPMI, Gibco) and Ham's F-12 Nutrient Mix (Gibco), respectively, containing 10% fetal bovine serum (FBS) and 0.01% ZellShield (Clinisciences). F-12 medium was supplemented with 500  $\mu$ L of insulin (10 mg/mL), 27.8  $\mu$ L of hydrocortisone (100 mg/mL and 5 mL of HEPES 1M. Each cell lines were obtained from the American Type Culture Collection and cultured at 37 °C with 5% CO<sub>2</sub> and 95% of air-humidified atmosphere. The MCF-7 and MDA-MB-231 cells arose from a pleural effusion of patients with metastatic breast adenocarcinoma. The SUM159PT cells came from a pleural effusion of patients with pleomorphic breast carcinoma. The immortalized cell line, hTERT, was cultured in MEGM (Lonza, Levallois, France) supplemented with 1% penicillin/streptomycin and served as the controls for the relative expression of the pluripotency genes.

#### *Establishment of H19 and miR-675 Overexpressing Cell Lines*

To establish the cell lines overexpressing *H19*, the SUM159PT cell line was transfected with 1  $\mu$ g of pH19-mCherry using Nucleofector Amaxa (Lonza), and cells were allowed to recover for 48 h. Cells were then selected in the presence of hygromycin B at 0.6 mg/mL for at least one month before their use for in vitro experiments. The *H19* overexpressing MCF-7 and MDA-MB-231 cell lines and the miR-675 overexpressing MDA-MB-231 cell line were previously described [23].

#### *siRNA and miRNA Transfection*

For siRNA transfection,  $1 \times 10^5$  cells were plated in 6-well plates. After 24h, cells were transfected with negative control (SR-CL000-005; Eurogentec) or *H19* siRNA (Table S1) using jetPRIME® Transfection Reagent, according to the manufacturer's guidelines (Polyplus Transfection®). Cells were lysed 48 h after transfection.

For miRNA transfection,  $1 \times 10^5$  cells were plated in 6-well plates. After 24 h, cells were transfected with Hsa-miR-675-5p mimic or its hairpin inhibitor with DharmaFECT-Duo, according to the manufacturer's guidelines (Thermo Fischer Scientific Dharmacon). Cells were lysed 48 h after transfection.

#### *RNA Extraction, Reverse Transcription and Real-time RT-PCR*

RNA extraction and qRT-PCR were performed as previously described [63]. Primers used for qRT-PCR are described in Table S2. miRNA extraction and qRT-PCR were performed as previously described [23].

#### *Migration and Invasion Assays*

Cell migration and invasion were determined by Transwell assay. For migration assay,  $1 \times 10^4$  cells were seeded on a 0.045-mg/mL collagen (Millipore)-coated insert (0.8  $\mu$ M; BD Biosciences) of 12-well plates. After 8 h, cells migrating to the other side of the filter were stained with Hoescht 33258 1 mM and counted. For the invasion assay,  $4 \times 10^4$  cells were seeded on a 3-mg/mL collagen (Millipore)-coated insert (0.8  $\mu$ M; BD Biosciences) of 12-well plates. After 24 h, cells invading to the other side of the filter were stained with Hoescht 33258 1 mM and counted.

#### *Colony-forming and Sphere-forming Capacities*

In the clonogenic assay, 250 cells were seeded in 100-mm<sup>2</sup> dishes. After 15 days, wells were fixed with 4% paraformaldehyde and stained with 0.5% crystal violet and counted.

For the sphere-forming capacity, a ranging from 1024 cells to 1 cell were cultured in a sphere medium consisting of phenol red-free DMEM-F12 (Gibco), 0.4% bovine serum albumin (Sigma-Aldrich, St Quentin Fallavier, France), 10 mL of B27 additive (Invitrogen, Illkirch, France), 5 mg/mL of insulin



(Sigma-Aldrich), 4 µg/mL of heparin and 20 ng/mL of epidermal growth factor (EGF) and fibroblast growth factor (FGF) (Biotechne, Abingdon, OX, United Kingdom). Cells were seeded in 96-well low-adhesion plates. The number of spheres per well were counted 4 days later.

### *Aldefluor Assay*

To measure the ALDH1 enzyme activity,  $1 \times 10^6$  cells were suspended in a ALDEFUOR assay buffer containing ALDH1 substrate (BAAA, 1 µmol/L; StemCell) and incubated for 30 min at 37 °C. In each experiment, a sample of cells was incubated, under identical conditions, with 50 mmol/L of diethylaminobenzaldehyde as a negative control. Only the 30% most-negative cells were collected as ADLH<sup>-</sup> cells. Flow cytometry data were acquired on a CyAn ADP cytometer (Beckman Coulter, Villepinte, France) with Summit software. All analyses were performed with FlowJo software.

### *Western Blot Analysis*

Cell lysis, electrophoresis, protein transfer, immunoblotting and signal revelation were performed as previously described [23]. Primary antibodies used were anti-E-cadherin, anti-N-cadherin, anti-Snail, anti-TCF8/ZEB1, anti-vimentin, anti-ZO-1 (#9782; Cell Signaling) and anti-actin (A2066-2ML; Sigma, St Quentin Fallavier, France) as a loading control.

### *Transgenic Zebrafish Xenograft*

SUM159PT pH19-mCherry<sup>neg/high</sup> and MDA-MB-231 cells were incubated for 40 min in an atmosphere at 37 °C and 5% CO<sub>2</sub>, with a solution of lipophilic tracers (Vybrant DiI labeling solution for mCherry<sup>high</sup> and for MDA-MB-231 cells overexpressing *H19* and miR-675, Vybrant DiD labeling solution for mCherry<sup>neg</sup> and, for each, a MDA-MB-231 control; Invitrogen) prepared in 1 mL of culture medium without FBS. Then, cells were dissociated with trypsin/0.25% EDTA (Gibco) and counted. One million cells of each condition were centrifuged at 100 g for 5 min. The pellet was resuspended with 100 µL of PBS/EDTA 5 mM. From the laying, transgenic zebrafish (*fli1:GFP*) were placed in 0.2-mM 1-phenyl-2-thio-urea (Sigma) to prevent pigmentation up to 48 h postfertilization. Fish were anaesthetized with 0.04-mg/mL MS-222 (Sigma, A5040), and SUM159PT pH19-mCherry<sup>neg</sup> (DiD) and pH19-mCherry<sup>high</sup> (DiI) cells or MDA-MB-231 control cells (DiD) and MDA-MB-231-overexpressed *H19* or MDA-MB-321 control cells (DiD) and MDA-MB-231-overexpressed miR-675 cells (DiI) were coinjected using a stereomicroscope (LEICA M125) and a microinjector (FemtoJET; Eppendorf) in a yolk sac. After injection, zebrafish were placed in a solution of phenylthiourea and penicillin-streptomycin renewed every 2 days at 28 °C in the dark. During 48 h, the incubation temperature of the zebrafish varied (30 °C at 24 h and 32 °C at 48 h up to 3 days postinjection) to promote cell development. Fish were fixed with 4% paraformaldehyde solution and placed in Petri dishes with glass bottoms. Fluorescent pictures were captured using automated image acquisition software from ZEISS with a LSM 880 microscope. Transgenic Tg(*fli1:EGFP*)<sup>y1</sup> zebrafish [64] were maintained in compliance with the French and European Union guidelines for the handling of laboratory animals (Directive 2010/63/EU of the European Parliament and of the Council of 22 September 2010 on the protection of animals used for scientific purposes). The experimental procedures carried out on zebrafish were reviewed and approved by the local ethics committee, CEEA 75 Nord Pas-de-Calais and the French Ministry of Higher Education and Research (APAFiS approval number 13527-2018011722529804\_v3).

### *Statistical Analysis*

Data are expressed as the mean values ± standard error of the mean of at least 3 independent experiments. The statistical analysis was done by using a Student's *t*-test, and a *p*-value < 0.05 was considered significant. \* *p* < 0.05, \*\* *p* < 0.01 and \*\*\* *p* < 0.001.

### Conclusions

In conclusion, this work highlights the involvement and relative contribution of lncRNA *H19* and miR-675 in the occurrence and development of breast cancer metastasis. In most cases, *H19* and miR-675 exert similar effects. Indeed, we show that both *H19* and miR-675 participate in the migration, invasion and stemness of breast cancer cells. Yet, miR-675 does not have the same effect as the lncRNA *H19* during the EMT. Their actions appear to be complementary and contribute to the metastasis development. These results highlight the interactions between lncRNA and miRNA to influence the EMT and, after further validation, might be taken into account in the therapy of breast cancer.

**Supplementary Materials:** The following are available online at <http://www.mdpi.com/2072-6694/12/7/1730/s1>, Figure S1: Uncropped western blot membranes (visible condition) corresponding to Figure S2, Figure S2: Revelation of uncropped western blot membranes showing the detection of ZO-1 protein expression in MCF-7 control (mock) or *H19*-stably overexpressing cells (*H19*), and in MDA-MB-231 control (mock), *H19*-stably overexpressing cells (*H19*) or miR-675-stably overexpressing cells (miR-675). Actin was used as an equiloading control, Figure S3: Uncropped western blot membranes (visible condition) corresponding to Figure S4, Figure S4: Revelation of uncropped western blot membranes showing the detection of E-cadherin protein expression in MCF-7 control (mock) or *H19*-stably overexpressing cells (*H19*). Actin was used as an equiloading control. Figure S5: Uncropped western blot membranes (visible condition) corresponding to Figure S6. Figure S6: Revelation of uncropped western blot membranes showing the detection of TCF8/ZEB1 protein expression in MDA-MB-231 control (mock), *H19*-stably overexpressing cells (*H19*) or miR-675-stably overexpressing cells (miR-675). Actin was used as an equiloading control, Figure S7: Uncropped western blot membranes (visible condition) corresponding to Figure S8, Figure S8: Revelation of uncropped western blot membranes showing the detection of N-cadherin protein expression in MDA-MB-231 control (mock), *H19*-stably overexpressing cells (*H19*) or miR-675-stably overexpressing cells (miR-675). Actin was used as an equiloading control, Figure S9: Uncropped western blot membranes (visible condition) corresponding to Figure S10, Figure S10: Revelation of uncropped western blot membranes showing the detection of vimentin protein expression in MDA-MB-231 control (mock), *H19*-stably overexpressing cells (*H19*) or miR-675-stably overexpressing cells (miR-675). Actin was used as an equiloading control, Figure S11: Uncropped western blot membranes (visible condition) corresponding to Figure S12, Figure S12: Revelation of uncropped western blot membranes showing the detection of snail protein expression in MDA-MB-231 control (mock), *H19*-stably overexpressing cells (*H19*) or miR-675-stably overexpressing cells (miR-675). Actin was used as an equiloading control, Figure S13: Uncropped western blot membranes (visible condition) corresponding to Figure S14, Figure S14: Revelation of uncropped western blot membranes showing the detection of ZO-1 protein expression in SUM159PT-pH19-mCherryneg (mCherryneg) or SUM159PT-pH19-mCherryhigh (mCherryhigh) cells. Actin was used as an equiloading control, Figure S15: Uncropped western blot membranes (visible condition) corresponding to Figure S16, Figure S16: Revelation of uncropped western blot membranes showing the detection of TCF8/ZEB1 protein expression in SUM159PT-pH19-mCherryneg (mCherryneg) or SUM159PT-pH19-mCherryhigh (mCherryhigh) cells. Actin was used as an equiloading control, Figure S17: Uncropped western blot membranes (visible condition) corresponding to Figure S18, Figure S18: Revelation of uncropped western blot membranes showing the detection of N-cadherin protein expression in SUM159PT-pH19-mCherryneg (mCherryneg) or SUM159PT-pH19-mCherryhigh (mCherryhigh) cells. Actin was used as an equiloading control, Figure S19: Uncropped western blot membranes (visible condition) corresponding to Figure S20, Figure S20: Revelation of uncropped western blot membranes showing the detection of vimentin protein expression in SUM159PT-pH19-mCherryneg (mCherryneg) or SUM159PT-pH19-mCherryhigh (mCherryhigh) cells. Actin was used as an equiloading control, Figure S21: Uncropped western blot membranes (visible condition) corresponding to Figure S22, Figure S22: Revelation of uncropped western blot membranes showing the detection of snail protein expression in SUM159PT-pH19-mCherryneg (mCherryneg) or SUM159PT-pH19-mCherryhigh (mCherryhigh) cells. Actin was used as an equiloading control, Table S1: References of *H19* siRNA transient transfection, Table S2. Primer used for qRT-PCR.

**Author Contributions:** Conceptualization, E.A. and X.L.B.; Methodology, C.L. (Clément Lecerf), E.P. and E.A.; Investigation and data acquisition, E.P., C.L. (Clément Lecerf), J.C., C.V., C.L. (Chann Lagadec), R.P.B., P.F., F.B., L.R., M.W., P.V., S.M. and P.-O.A.; Writing—Original draft preparation, C.L. (Clément Lecerf), E.P. and E.A.; Writing—Review and editing, E.A. and X.L.B.; Project administration, E.A. and funding acquisition, E.A. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by INSERM and a grant from INCA (PLBio 2010-180), “Ligue contre le cancer” and “Cancéropôle Nord-Ouest/projets émergents 2017”. C.L. (Clément Lecerf), E.P. and J.C. were supported by

doctoral fellowships from the University of Lille. C.V. was supported by doctoral fellowships from the University of Lille 2 and by ARC Foundation fellowships.

**Acknowledgments:** The authors thank members of the laboratory for helpful discussions and administrative and technical support.

**Conflicts of Interest:** The authors declare no conflicts of interest.

### References

1. Angrand, P.O.; Vennin, C.; Le Bourhis, X.; Adriaenssens, E. The role of long non-coding RNAs in genome formatting and expression. *Front Genet.* **2015**, *6*, e165. [[CrossRef](#)]
2. Brannan, C.I.; Dees, E.C.; Ingram, R.S.; Tilghman, S.M. The product of the H19 gene may function as an RNA. *Mol Cell Biol.* **1990**, *10*, 28–36. [[CrossRef](#)] [[PubMed](#)]
3. Adriaenssens, E.; Dumont, L.; Lottin, S.; Bolle, D.; Leprêtre, A.; Delobelle, A.; Bouali, F.; Dugimont, T.; Coll, J.; Curgy, J.J. H19 overexpression in breast adenocarcinoma stromal cells is associated with tumor values and steroid receptor status but independent of p53 and Ki-67 expression. *Am. J. Pathol.* **1998**, *153*, 1597–1607. [[CrossRef](#)]
4. Adriaenssens, E.; Lottin, S.; Dugimont, T.; Fauquette, W.; Coll, J.; Dupouy, J.P.; Boilly, B.; Curgy, J.J. Steroid hormones modulate H19 gene expression in both mammary gland and uterus. *Oncogene* **1999**, *18*, 4460–4473. [[CrossRef](#)] [[PubMed](#)]
5. Liu, J.; Kahri, A.I.; Heikkilä, P.; Ilvesmäki, V.; Voutilainen, R. H19 and insulin-like growth factor-II gene expression in adrenal tumors and cultured adrenal cells. *J. Clin. Endocrinol. Metab.* **1995**, *80*, 492–496. [[CrossRef](#)] [[PubMed](#)]
6. Ariel, I.; Weinstein, D.; Voutilainen, R.; Schneider, T.; Lustig-Yariv, O.; de Groot, N.; Hochberg, A. Genomic imprinting and the endometrial cycle. *The expression of the imprinted gene H19 in the human female reproductive organs. Diagn Mol Pathol Am. J. Surg. Pathol. Part B* **1997**, *6*, 17–25. [[CrossRef](#)]
7. Liu, G.; Xiang, T.; Wu, Q.F.; Wang, W.X. *Long Noncoding RNA H19-Derived miR-675 Enhances Proliferation and Invasion via RUNX1 in Gastric Cancer Cells*; Cognizant Communication Corporation: Putnam Valley, NY, USA, 2016; Volume 23, pp. 99–107. [[CrossRef](#)]
8. Liang, W.C.; Fu, W.M.; Wong, C.W.; Wang, Y.; Wang, W.M.; Hu, G.X.; Zhang, L.; Xiao, L.J.; Wan, D.C.; Zhang, J.F.; *et al.* The lncRNA H19 promotes epithelial to mesenchymal transition by functioning as miRNA sponges in colorectal cancer. *Oncotarget* **2015**, *6*, 22513–22525. [[CrossRef](#)]
9. Liu, C.; Chen, Z.; Fang, J.; Xu, A.; Zhang, W.; Wang, Z. H19-derived miR-675 contributes to bladder cancer cell proliferation by regulating p53 activation. *Tumor Biol.* **2016**, *37*, 263–270. [[CrossRef](#)]
10. Wang, L.; Cai, Y.; Zhao, X.; Jia, X.; Zhang, J.; Liu, J.; Zhen, H.; Wang, T.; Tang, X.; Liu, Y.; *et al.* Down-regulated long non-coding RNA H19 inhibits carcinogenesis of renal cell carcinoma. *Neoplasma* **2015**, *62*, 412–418. [[CrossRef](#)]
11. Shi, Y.; Wang, Y.; Luan, W.; Wang, P.; Tao, T.; Zhang, J.; Qian, J.; Liu, N.; You, Y. Long Non-Coding RNA H19 Promotes Glioma Cell Invasion by Deriving miR-675. *PLoS ONE* **2014**, *9*, e86295. [[CrossRef](#)]
12. Berteaux, N.; Lottin, S.; Monté, D.; Pinte, S.; Quatannens, B.; Coll, J.; Hondermarck, H.; Curgy, J.J.; Dugimont, T.; Adriaenssens, E. H19 mRNA-like noncoding RNA promotes breast cancer cell proliferation through positive control by E2F1. *J. Biol. Chem.* **2005**, *280*, 29625–29636. [[CrossRef](#)] [[PubMed](#)]
13. Lecerf, C.; Le Bourhis, X.; Adriaenssens, E. The long non-coding RNA H19: An active player with multiple facets to sustain the hallmarks of cancer. *Cell Mol. Life Sci.* **2019**, *76*, 4673–4687. [[CrossRef](#)] [[PubMed](#)]
14. Lottin, S.; Adriaenssens, E.; Dupressoir, T.; Berteaux, N.; Montpellier, C.; Coll, J.; Dugimont, T.; Curgy, J.J. Overexpression of an ectopic H19 gene enhances the tumorigenic properties of breast cancer cells. *Carcinogenesis* **2002**, *23*, 1885–1895. [[CrossRef](#)] [[PubMed](#)]

15. Si, X.; Zang, R.; Zhang, E.; Liu, Y.; Shi, X.; Zhang, E.; Shao, L.; Li, A.; Yang, N.; Han, X.; *et al.* LncRNA H19 confers chemoresistance in ER $\alpha$ -positive breast cancer through epigenetic silencing of the pro-apoptotic gene BIK. *Oncotarget* **2016**, *7*, 81452–81462. [[CrossRef](#)] [[PubMed](#)]
16. Peng, F.; Li, T.T.; Wang, K.L.; Xiao, G.Q.; Wang, J.H.; Zhao, H.D.; Kang, Z.J.; Fan, W.J.; Zhu, L.L.; Li, M.; *et al.* H19/let-7/LIN28 reciprocal negative regulatory circuit promotes breast cancer stem cell maintenance. *Cell Death Dis.* **2017**, *8*, e2569. [[CrossRef](#)]
17. Li, Z.; Li, Y.; Li, Y.; Ren, K.; Li, X.; Han, X.; Wang, J. Long non-coding RNA H19 promotes the proliferation and invasion of breast cancer through upregulating DNMT1 expression by sponging miR-152. *J. Biochem. Mol. Toxicol.* **2017**, *31*, e21933. [[CrossRef](#)]
18. Cai, X.; Cullen, B.R. The imprinted H19 noncoding RNA is a primary microRNA precursor. *RNA* **2007**, *13*, 313–316. [[CrossRef](#)]
19. Garzon, R.; Fabbri, M.; Cimmino, A.; Calin, G.A.; Croce, C.M. MicroRNA expression and function in cancer. *Trends Mol. Med.* **2006**, *12*, 580–587. [[CrossRef](#)]
20. Tsang, W.P.; Ng, E.K.O.; Ng, S.S.M.; Jin, H.; Yu, J.; Sung, J.J.Y.; Kwok, T.T. Oncofetal H19-derived miR-675 regulates tumor suppressor RB in human colorectal cancer. *Carcinogenesis* **2010**, *31*, 350–358. [[CrossRef](#)]
21. Hernandez, J.M.; Elahi, A.; Clark, C.W.; Wang, J.; Humphries, L.A.; Centeno, B.; Bloom, G.; Fuchs, B.C.; Yeatman, T.; Shibata, D. miR-675 Mediates Downregulation of Twist1 and Rb in AFP-Secreting Hepatocellular Carcinoma. *Ann Surg. Oncol.* **2013**, *20*, 625–635. [[CrossRef](#)]
22. Zhuang, M.; Gao, W.; Xu, J.; Wang, P.; Shu, Y. The long non-coding RNA H19-derived miR-675 modulates human gastric cancer cell proliferation by targeting tumor suppressor RUNX1. *Biochem. Biophys. Res. Commun.* **2014**, *448*, 315–322. [[CrossRef](#)] [[PubMed](#)]
23. Vennin, C.; Spruyt, N.; Dahmani, F.; Julien, S.; Bertucci, F.; Finetti, P.; Chassat, T.; Bourette, R.P.; Le Bourhis, X.; Adriaenssens, E. H19 non coding RNA-derived miR-675 enhances tumorigenesis and metastasis of breast cancer cells by downregulating c-Cbl and Cbl-b. *Oncotarget* **2015**, *6*, 29209–29223. [[CrossRef](#)] [[PubMed](#)]
24. Hill, D.; Chen, L.; Snaar-Jagalska, E.; Chaudhry, B. Embryonic zebrafish xenograft assay of human cancer metastasis. *F1000Research* **2018**, *7*, e1682. [[CrossRef](#)]
25. Daulat, A.M.; Bertucci, F.; Audebert, S.; Sergé, A.; Finetti, P.; Josselin, E.; Castellano, R.; Birnbaum, D.; Angers, S.; Borg, J.P. PRICKLE1 Contributes to Cancer Cell Dissemination through Its Interaction with mTORC2. *Dev. Cell* **2016**, *37*, 311–325. [[CrossRef](#)]
26. Charafe-Jauffret, E.; Ginestier, C.; Iovino, F.; Wicinski, J.; Cervera, N.; Finetti, P.; Hur, M.J.; Diebel, M.E.; Monville, F.; Dutcher, J.; *et al.* Breast cancer cell lines contain functional cancer stem cells with metastatic capacity and a distinct molecular signature. *Cancer Res.* **2009**, *69*, 1302–1313. [[CrossRef](#)]
27. Creighton, C.J.; Li, X.; Landis, M.; Dixon, J.M.; Neumeister, V.M.; Sjolund, A.; Rimm, D.L.; Wong, H.; Rodriguez, A.; Herschkowitz, J.I.; *et al.* Residual breast cancers after conventional therapy display mesenchymal as well as tumor-initiating features. *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 13820–13825. [[CrossRef](#)]
28. Ishiguro, T.; Ohata, H.; Sato, A.; Yamawaki, K.; Enomoto, T.; Okamoto, K. Tumor-derived spheroids: Relevance to cancer stem cells and clinical applications. *Cancer Sci.* **2017**, *108*, 283–289. [[CrossRef](#)]
29. Grolez, G.P.; Hammadi, M.; Barras, A.; Gordienko, D.; Slomianny, C.; Völkel, P.; Angrand, P.O.; Pinault, M.; Guimaraes, C.; Potier-Cartereau, M.; *et al.* Encapsulation of a TRPM8 Agonist, WS12, in Lipid Nanocapsules Potentiates PC3 Prostate Cancer Cell Migration Inhibition through Channel Activation. *Sci. Rep.* **2019**, *9*, e7926. [[CrossRef](#)]
30. Raby, L.; Völkel, P.; Le Bourhis, X.; Angrand, P.O. La transplantation de cellules tumorales chez le poisson zèbre: De la recherche translationnelle à la médecine personnalisée. *Bull Cancer* **2020**, *107*, 30–40. [[CrossRef](#)]

31. Shima, H.; Kida, K.; Adachi, S.; Yamada, A.; Sugae, S.; Narui, K.; Miyagi, Y.; Nishi, M.; Ryo, A.; Murata, S.; *et al.* Lnc RNA H19 is associated with poor prognosis in breast cancer patients and promotes cancer stemness. *Breast Cancer Res. Treat.* **2018**, *170*, 507–516. [[CrossRef](#)]
32. Liao, S.; Yu, C.; Liu, H.; Zhang, C.; Li, Y.; Zhong, X. Long non-coding RNA H19 promotes the proliferation and invasion of lung cancer cells and regulates the expression of E-cadherin, N-cadherin, and vimentin. *OncoTargets Ther.* **2019**, *12*, 4099–4107. [[CrossRef](#)] [[PubMed](#)]
33. Matouk, I.J.; Halle, D.; Raveh, E.; Gilon, M.; Sorin, V.; Hochberg, A. The role of the oncofetal H19 lncRNA in tumor metastasis: Orchestrating the EMT-MET decision. *Oncotarget* **2016**, *7*, 3748–3765. [[CrossRef](#)] [[PubMed](#)]
34. He, P.; Qiu, K.; Jia, Y. Modeling of mesenchymal hybrid epithelial state and phenotypic transitions in EMT and MET processes of cancer cells. *Sci. Rep.* **2018**, *8*, e14323. [[CrossRef](#)] [[PubMed](#)]
35. Mayor, R.; Etienne-Manneville, S. The front and rear of collective cell migration. *Nat. Rev. Mol. Cell Biol.* **2016**, *17*, 97–109. [[CrossRef](#)]
36. Satoh, K.; Hamada, S.; Shimosegawa, T. Involvement of epithelial to mesenchymal transition in the development of pancreatic ductal adenocarcinoma. *J. Gastroenterol.* **2015**, *50*, 140–146. [[CrossRef](#)]
37. Liu, F.; Gu, L.N.; Shan, B.E.; Geng, C.Z.; Sang, M.X. Biomarkers for EMT and MET in breast cancer: An update. *Oncol. Lett.* **2016**, *12*, 4869–4876. [[CrossRef](#)]
38. Chen, M.J.; Deng, J.; Chen, C.; Hu, W.; Yuan, Y.C.; Xia, Z.K. LncRNA H19 promotes epithelial mesenchymal transition and metastasis of esophageal cancer via STAT3/EZH2 axis. *Int. J. Biochem. Cell Biol.* **2019**, *113*, 27–36. [[CrossRef](#)]
39. Liu, L.; Liu, L.; Lu, S. LncRNA H19 promotes viability and epithelial-mesenchymal transition of lung adenocarcinoma cells by targeting miR-29b-3p and modifying STAT3. *Int. J. Oncol.* **2019**, *54*, 929–941. [[CrossRef](#)]
40. Liang, W.Q.; Zeng, D.; Chen, C.F.; Sun, S.M.; Lu, X.F.; Peng, C.; Lin, H.Y. Long noncoding RNA H19 is a critical oncogenic driver and contributes to epithelial-mesenchymal transition in papillary thyroid carcinoma. *Cancer Manag. Res.* **2019**, *11*, 2059–2072. [[CrossRef](#)]
41. Xiong, H.; Shen, J.; Chen, Z.; Yang, J.; Xie, B.; Jia, Y.; Jayasinghe, U.; Wang, J.; Zhao, W.; Xie, S.; *et al.* H19/let-7/Lin28 ceRNA network mediates autophagy inhibiting epithelial-mesenchymal transition in breast cancer. *Int. J. Oncol.* **2020**, *56*, 794–806. [[CrossRef](#)]
42. Zhang, L.; Yang, F.; Yuan, J.; Yuan, S.; Zhou, W.; Huo, X.; Xu, D.; Bi, H.S.; Wang, F.; Sun, S.H. Epigenetic activation of the MiR-200 family contributes to H19-mediated metastasis suppression in hepatocellular carcinoma. *Carcinogenesis* **2013**, *34*, 577–586. [[CrossRef](#)] [[PubMed](#)]
43. Lv, M.; Zhong, Z.; Huang, M.; Tian, Q.; Jiang, R.; Chen, J. LncRNA H19 regulates epithelial–mesenchymal transition and metastasis of bladder cancer by miR-29b-3p as competing endogenous RNA. *Biochim Biophys Acta BBA. Mol. Cell. Res.* **2017**, *1864*, 1887–1899. [[CrossRef](#)]
44. O'Brien, S.J.; Carter, J.V.; Burton, J.F.; Oxford, B.G.; Schmidt, M.N.; Hallion, J.C.; Galandiuk, S. The role of the miR-200 family in epithelial–mesenchymal transition in colorectal cancer: A systematic review. *Int. J. Cancer* **2018**, *142*, 2501–2511. [[CrossRef](#)] [[PubMed](#)]
45. Title, A.C.; Hong, S.J.; Pires, N.D.; Hasenöhr, L.; Godbersen, S.; Stokar-Regenscheit, N.; Bartel, D.P.; Stoffel, M. Genetic dissection of the miR-200-Zeb1 axis reveals its importance in tumor differentiation and invasion. *Nat. Commun.* **2018**, *9*, e4671. [[CrossRef](#)] [[PubMed](#)]
46. Gregory, P.A.; Bert, A.G.; Paterson, E.L.; Barry, S.C.; Tsykin, A.; Farshid, G.; Vadas, M.A.; Khew-Goodall, Y.; Goodall, G.J. The miR-200 family and miR-205 regulate epithelial to mesenchymal transition by targeting ZEB1 and SIP1. *Nat. Cell Biol.* **2008**, *10*, 593–601. [[CrossRef](#)]
47. Li, M.; Chen, H.; Zhao, Y.; Gao, S.; Cheng, C. H19 Functions as a ceRNA in Promoting Metastasis Through Decreasing miR-200s Activity in Osteosarcoma. *DNA Cell Biol.* **2016**, *35*, 235–240. [[CrossRef](#)]



48. Zhao, Y.; Feng, C.; Li, Y.; Ma, Y.; Cai, R. LncRNA H19 promotes lung cancer proliferation and metastasis by inhibiting miR-200a function. *Mol. Cell Biochem.* **2019**, *460*, 1–8. [[CrossRef](#)]
49. Zhou, S.; He, Y.; Yang, S.; Hu, J.; Zhang, Q.; Chen, W.; Xu, H.; Zhang, H.; Zhong, S.; Zhao, J.; *et al.* The regulatory roles of lncRNAs in the process of breast cancer invasion and metastasis. *Biosci. Rep.* **2018**, *38*, 1–2. [[CrossRef](#)]
50. Wang, M.; Li, Y.; Xiao, G.D.; Zheng, X.Q.; Wang, J.C.; Xu, C.W.; Qin, S.; Ren, H.; Tang, S.H.; Sun, X. H19 regulation of oestrogen induction of symmetric division is achieved by antagonizing Let-7c in breast cancer stem-like cells. *Cell Prolif.* **2019**, *52*, e12534. [[CrossRef](#)]
51. Al-Hajj, M.; Wicha, M.S.; Benito-Hernandez, A.; Morrison, S.J.; Clarke, M.F. Prospective identification of tumorigenic breast cancer cells. *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 3983–3988. [[CrossRef](#)]
52. Tomita, H.; Tanaka, K.; Tanaka, T.; Hara, A. Aldehyde dehydrogenase 1A1 in stem cells and cancer. *Oncotarget* **2016**, *7*, 11018–11032. [[CrossRef](#)] [[PubMed](#)]
53. Kim, J.; Chu, J.; Shen, X.; Wang, J.; Orkin, S.H. An extended transcriptional network for pluripotency of embryonic stem cells. *Cell* **2008**, *132*, 1049–1061. [[CrossRef](#)] [[PubMed](#)]
54. Hadjimichael, C.; Chanoumidou, K.; Papadopoulou, N.; Arampatzi, P.; Papamatheakis, J.; Kretsovali, A. Common stemness regulators of embryonic and cancer stem cells. *World J. Stem Cells* **2015**, *7*, 1150–1184. [[CrossRef](#)] [[PubMed](#)]
55. Barsyte-Lovejoy, D.; Lau, S.K.; Boutros, P.C.; Khosravi, F.; Jurisica, I.; Andrulis, I.L.; Tsao, M.S.; Penn, L.Z. The c-Myc oncogene directly induces the H19 noncoding RNA by allele-specific binding to potentiate tumorigenesis. *Cancer Res.* **2006**, *66*, 5330–5337. [[CrossRef](#)] [[PubMed](#)]
56. Abi Habib, W.; Azzi, S.; Brioude, F.; Steunou, V.; Thibaud, N.; Das Neves, C.; Le Jule, M.; Chantot-Bastarud, S.; Keren, B.; Lyonnet, S.; *et al.* Extensive investigation of the IGF2/H19 imprinting control region reveals novel OCT4/SOX2 binding site defects associated with specific methylation patterns in Beckwith-Wiedemann syndrome. *Hum Mol. Genet.* **2014**, *23*, 5763–5773. [[CrossRef](#)] [[PubMed](#)]
57. Gao, X.; Qiao, Y.; Han, D.; Zhang, Y.; Ma, N. Enemy or partner: Relationship between intronic micrnas and their host genes. *IUBMB Life* **2012**, *64*, 835–840. [[CrossRef](#)]
58. Shi, W.; Gerster, K.; Alajez, N.M.; Tsang, J.; Waldron, L.; Pintilie, M.; Hui, A.B.; Sykes, J.; P'ng, C.; Miller, N.; *et al.* MicroRNA-301 mediates proliferation and invasion in human breast cancer. *Cancer Res.* **2011**, *71*, 2926–2937. [[CrossRef](#)]
59. Qiao, Y.; Ma, N.; Wang, X.; Hui, Y.; Li, F.; Xiang, Y.; Zhou, J.; Zou, C.; Jin, J.; Lv, G.; *et al.* MiR-483-5p controls angiogenesis in vitro and targets serum response factor. *FEBS Lett.* **2011**, *585*, 3095–3100. [[CrossRef](#)]
60. Collette, J.; Le Bourhis, X.; Adriaenssens, E. Regulation of Human Breast Cancer by the Long Non-Coding RNA H19. *Int. J. Mol. Sci.* **2017**, *18*, e2319. [[CrossRef](#)]
61. Yang, W.; Redpath, R.E.; Zhang, C.; Ning, N. Long non-coding RNA H19 promotes the migration and invasion of colon cancer cells via MAPK signaling pathway. *Oncol. Lett.* **2018**, *16*, 3365–3372. [[CrossRef](#)]
62. Zhao, J.; Ma, S.T. Downregulation of lncRNA H19 inhibits migration and invasion of human osteosarcoma through the NF- $\kappa$ B pathway. *Mol. Med. Rep.* **2018**, *17*, 7388–7394. [[CrossRef](#)] [[PubMed](#)]
63. Berteaux, N.; Aptel, N.; Cathala, G.; Genton, C.; Coll, J.; Daccache, A.; Spruyt, N.; Hondermarck, H.; Dugimont, T.; Curgy, J.J.; *et al.* A novel H19 antisense RNA overexpressed in breast cancer contributes to paternal IGF2 expression. *Mol. Cell Biol.* **2008**, *28*, 6731–6745. [[CrossRef](#)] [[PubMed](#)]
64. Lawson, N.D.; Weinstein, B.M. In vivo imaging of embryonic vascular development using transgenic zebrafish. *Dev. Biol.* **2002**, *248*, 307–318. [[CrossRef](#)] [[PubMed](#)]

### 2.3 Conclusion

Les ARNnc *H19* et miR-675 promeuvent la tumorigenèse et le caractère agressif des cellules cancéreuses mammaires. L'association du phénotype des cellules cancéreuses et de la surexpression de ces ARNnc apparaissent par plusieurs mécanismes : la transition épithélio-mésenchymateuse, la reprogrammation des cellules en cellules souches cancéreuses, la résistance à l'Anoïkis ou encore la capacité migratoire. La transplantation des cellules cancéreuses mammaires surexprimant *H19* ou miR-675 montre que ces cellules présentent des capacités invasives supérieure aux cellules contrôles. Ces résultats démontrent que ces deux ARNnc augmentent le potentiel migratoire des cellules cancéreuses. Ce potentiel passe par des modifications morphologiques et moléculaires des cellules et ces modifications ont été déterminées par la suite *in vitro*. La xénogreffe de cellules cancéreuses sur des embryons de poisson zèbre est donc un excellent modèle en recherche pré-clinique en cancérologie.

### 3. Genetic Engineering of Zebrafish in Cancer Research

#### 3.1 Contexte scientifique

L'apparition de maladies néoplasiques n'est pas une caractéristique humaine, et affecte un nombre très important d'espèces (Pesavento *et al.*, 2018), notamment les poissons (Harshbarger *et al.*, 2001). Chez les poissons, des tumeurs sont décrites dès le début du 20<sup>ème</sup> siècle (Plehn, 1906). Les recherches en oncologie se sont intéressées plus particulièrement au développement de mélanome dans le genre *Xiphophorus*. Le croisement hybride de deux espèces, communément appelées platy et xipho, conduit à l'apparition de mélanomes dans la descendance. Les platys possèdent un profil de cellules pigmentées, les macromélanophores, qui est absent des xiphos. Au sein de la descendance, ces macromélanophores deviennent hautement prolifératifs et indifférenciés, ce qui conduit à l'apparition de néoplasie. La notion d'hérédité et ainsi de génétique du cancer apparaît donc dès les années 1920 (Patton *et al.*, 2010). De nombreuses espèces de poissons développent spontanément et sporadiquement des tumeurs avec des caractéristiques histologiques similaires à ce qui est observé chez l'homme (Harshbarger *et al.*, 2001). Afin d'obtenir des modèles de carcinogénèses, des protocoles ont été mis en place pour induire la tumorigenèse chez le poisson zèbre (Stanton, 1965). A l'époque, le poisson zèbre est un modèle animal qui se développe dans les laboratoires. L'entretien de cette espèce est relativement facile, il est possible d'élever 30 individus dans un bac de 10 L en maintenant une température autour de 25°C. Les molécules sont diluées dans l'eau de l'aquarium et absorbées par les poissons, et permet une exposition d'un groupe d'individu large à une même concentration d'une molécule mutagène. Des lignées mutantes de poissons zèbres ont été générées pour un projet de mutagenèse à large échelle (Haffter *et al.*, 1996 ; Driever *et al.*, 1996). L'objectif de ce projet était initialement d'obtenir une meilleure connaissance des gènes impliqués dans le développement embryonnaire. Néanmoins, les processus de différenciation cellulaire et de développement sont souvent proches des mécanismes à l'origine du développement tumoral. La banque de mutants est un atout précieux en cancérologie. Enfin, le séquençage du génome entier du poisson zèbre (Howe *et al.*, 2013), associé au développement des outils de biologie moléculaire et d'édition du génome comme les ZFNs (Urnoy *et al.*, 2010), les TALENs (Boch *et al.*, 2009 ; Moscou & Bogdanove, 2009), ou les CRISPRs (Doudna & Charpentier, 2014) permettent des modifications précises d'oncogènes ou de gènes suppresseurs de tumeur, et placent le poisson zèbre comme un modèle de premier plan en cancérologie. De ce fait, cette revue a pour but de rassembler toutes les approches génétiques sur le poisson zèbre dans le domaine de la recherche sur le cancer.



## 3.2 Article

### Genetic Engineering of Zebrafish in Cancer Research

Ludivine Raby, Pamela Völkel, Xuefen Le Bourhis<sup>ORCID</sup> and Pierre-Olivier Angrand \*

Univ. Lille, CNRS, Inserm, CHU Lille, UMR9020-U1277-CANTHER-Cancer Heterogeneity Plasticity and Resistance to Therapies, F-59000 Lille, France; ludivine.raby.etu@univ-lille.fr (L.R.); pamel.voelkel@univ-lille.fr (P.V.); xuefen.le-bourhis@univ-lille.fr (X.L.B.)

Received: 6 July 2020; Accepted: 31 July 2020; Published: 4 August 2020

#### Abstract

Zebrafish (*Danio rerio*) is an excellent model to study a wide diversity of human cancers. In this review, we provide an overview of the genetic and reverse genetic toolbox allowing the generation of zebrafish lines that develop tumors. The large spectrum of genetic tools enables the engineering of zebrafish lines harboring precise genetic alterations found in human patients, the generation of zebrafish carrying somatic or germline inheritable mutations or zebrafish showing conditional expression of the oncogenic mutations. Comparative transcriptomics demonstrate that many of the zebrafish tumors share molecular signatures similar to those found in human cancers. Thus, zebrafish cancer models provide a unique in vivo platform to investigate cancer initiation and progression at the molecular and cellular levels, to identify novel genes involved in tumorigenesis as well as to contemplate new therapeutic strategies.

#### Introduction

As early as 1902, Marianne Plehn at the Versuchsstation für Fischrei (Munich, Germany) pioneered the description of cancer in fish and recorded various cases of cancers in different salmonoids and cyprinoids [1]. Among early studies, in the 1920s and beginning of 1930s Curt Kosswig (Universität Münster, Germany), Georg Häussler (Institut für Experimentelle Krebsforschung, Heidelberg, Germany) and Myron Gordon (Cornell University, NY, USA) independently observed that interspecies hybrids between strains of the southern platyfish *Xiphophorus maculatus* and the green swordtail *Xiphophorus helleri* spontaneously develop malignant melanomas [2–5]. This groundbreaking observation conceptualizes the genetic basis and the heredity of certain cancers, about 40 years before the discovery of the first oncogenes. Indeed, these *Xiphophorus* neoplasms originate from the macromelanophores as the result of the interaction of two genetic traits. In southern platyfish, a tumor suppressor (*cdkn2x*) conceals the effects of an oncogene (*xmrk*). When green swordtails are mated by artificial insemination with the interspecies hybrids, the segregation of the tumor suppressor gene from the oncogene is responsible for melanoma formation according to Mendelian laws [6,7]. Over the years, spontaneous and induced neoplasms from a spectrum of tissues were observed in more than 200 fish species farmed or raised in aquaria and belonging to a large number of families including

Poeciliidae (livebearers), Cyprinidae (carps and minnows), Cichlidae (cichlids), Cyprinodontidae (killifish), Characidae (tetras), Adrianichthyidae (medakas), Aplocheilidae (rivulins), Anguillidae (anguillas), Percidae (perches) or Salmonidae (salmons and trouts) [8]. These cancers originate from all tissues and are highly similar to human tumors at the histological and cellular levels raising the idea that fish could be used as cancer models and within the various fish models, zebrafish has progressively become central to cancer research.

The zebrafish (*Danio rerio*) is a small (3 to 4 cm long) freshwater fish belonging the Cyprinidae family that lives in rivers and rice paddies in India, Nepal and Bangladesh [9,10]. It emerged as a model to study early development

of the embryo in the 1930s, but rapidly expanded into a larger area of research [11]. The success of the zebrafish in experimental biology is mainly due to its attributes that include: (i) the large number of progeny (100–200 embryos per clutch) offering high confidence in statistical analysis; (ii) the production of optically clear embryos that undergo rapid development ex utero; (iii) the manipulability of its embryology [12]; (iv) comparison of the zebrafish genome to the human reference genome shows that about 70% of human genes have at least one zebrafish orthologue (when considering disease-causing human genes the orthology is up to 80%) [13]; (v) the ability of zebrafish to absorb molecules that are dissolved in water allowing drug assays and screens [14,15]; and (vi) the reduced zebrafish husbandry expenses. In addition, transparent mutant lines like Casper and Crystal have been developed to facilitate in vivo imaging throughout all stages of zebrafish including adults [16,17]. Similarly, a plethora of transgenic zebrafish lines such as Tg(fli1:EGFP) [18] and Tg(flk1:EGFP) [19] having their entire vascular system labeled with green fluorescence or Tg(gata1:DsRed) [20] with red fluorescent blood cells allows to follow the development and location of different organs and cells in vivo.

This review aims to provide an overview of the genetic approaches making the zebrafish an excellent model in cancer research, as first proposed in 2002 by Leonard Zon and colleagues (Children’s Hospital, Boston, MA, USA) [21].

### Chemical Carcinogenesis

Zebrafish spontaneously develop, although rarely, almost any type of tumors[22–24]. These tumors usually occur after the age of 1 or 2 years. Zebrafish was also the first fish species subjected to chemical exposure for cancer research. In the 1960s, Mearle Stanton (National Institute of Cancer, Bethesda, MD, USA) exposed zebrafish to the carcinogen diethylnitrosamine (N-nitroso-diethylamine, DEN), and observed that they developed hepatic neoplasias [25]. A number of chemical compounds known to be carcinogenic in mammals also induce tumor formation in zebrafish [26–31] (Table 1).

**Table 1.** Zebrafish cancers induced by chemical exposure.

Chemicals	Cancer Types	References
NDMA	Hepatocarcinoma	[26]
DEN	Liver and pancreas carcinomas	[27]
DMBA	Liver, gill, pancreas, gastrointestinal, pancreas, testis, cartilage, muscle, blood vessel, connective and lymphoid tissues and neural tumors	[28,29]
DBP	Hepatocarcinoma	[29]
MNNG	Liver, gill, pancreas, gastrointestinal, pancreas, testis, kidney, muscle tumors and hemangio(sarco)mas	[30]
ENU	Epidermal papillomas	[31]

DBP, dibenzo(a,l)pyrene; DEN, diethylnitrosamine (N-nitrosodimethylamine); DMBA, 7,12-dimethylbenz[a]anthracene; ENU, N-ethyl-N-nitrosourea; MNNG, N-methyl-N'-nitro-N-nitrosoguanidine; NDMA, N-nitrosodimethylamine.

Systematic studies showed that N-nitrosodimethylamine (NDMA) induces only liver tumors [26] and N-nitrosodiethylamine (DEN) primarily induces liver and pancreas carcinomas [27],

whereas 7,12-dimethylbenz[a]anthracene (DMBA) and N-methyl-N'-nitro-N-nitrosoguanidine (MNNG) can induce a broad tumor spectrum, including epithelial, mesenchymal and neural tumors [28,30]. In contrast, ethylnitrosourea (ENU) exposure is more selective, with most treated fish developing epidermal papillomas but no invasive skin cancers [31]. These fish also exhibit cavernous hemangiomas and malignant peripheral nerve

sheath tumors at low incidence. Taken together, these data demonstrate the capacity of the zebrafish to develop a diverse range of cancers that pathologically resemble the human tumor types.

DMBA is a carcinogen metabolized in the liver and able to induce in zebrafish a spectrum of liver lesions corresponding to the different stages of neoplastic progression in human, from dysplastic nodules to high graded hepatocellular carcinomas. Using oligonucleotide microarrays, gene expression signatures of zebrafish liver tumors were compared to normal zebrafish liver and human liver cancers [29,32]. The study revealed a marked dysregulation in hallmark cancer genes including genes associated with cell cycle, cytoskeletal organization, metastasis, RNA processing and protein synthesis. Furthermore, the study shows that there are strong similarities at the transcriptomic level between zebrafish and human liver tumors that extend to tumor progression [29]. Thus, zebrafish neoplasms appear closely related to human cancers in terms of histopathology and in terms of alteration of gene expression programs.

Chemical carcinogenesis in zebrafish easily generates models that resemble to human cancers. However, the approach has limits linked to late tumorigenesis onset, low incidence of tumor development, or genetic and location heterogeneities of the induced tumors. Furthermore, zebrafish chemical carcinogenesis models cannot be propagated as lines. Thus, the full potential of the zebrafish model in cancer research would require the generation of fish lines harboring genetic alteration in cancer genes.

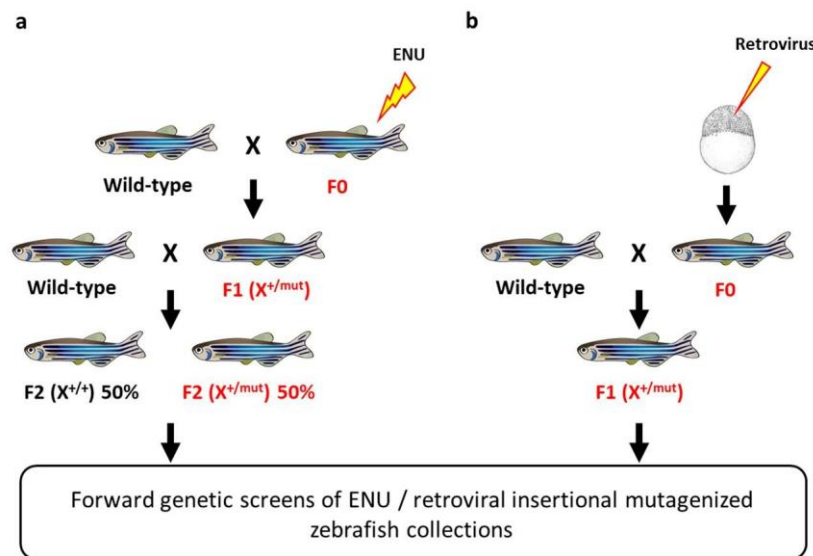
### Forward Genetics and Zebrafish Mutagenesis Screens

N-ethyl-N-nitrosourea (ENU) is an alkylating agent preferentially responsible for A to T transversions and AT to GC transitions and therefore used as a potent point mutagen in *Drosophila* [33,34], mouse [35], as well as in zebrafish [12,36]. In the 1990s, two large-scale ENU-based mutagenesis screens in zebrafish were simultaneously conducted by Wolfgang Driever (Massachusetts General Hospital, Charlestown, MA, USA), Christiane Nüsslein-Volhard (Max-Planck-Institut für Entwicklungsbiologie, Tübingen, Germany) and their coworkers and led to the identification of thousands of mutants showing developmental defects [37,38] (Figure 1a). In addition to the enormous impact they have in the field of developmental genetics, such large-scale mutagenesis screens occasionally resulted in the identification of mutant zebrafish lines with an increased incidence of spontaneous neoplasia or a higher sensitivity chemical carcinogenesis (Table 2). For instance, in a forward genetic screen for mutations affecting ciliary motility in a large-scale ENU mutagenesis zebrafish library, Freek van Eeden and colleagues (Hubrecht Laboratory, Utrecht, The Netherlands) isolated a zebrafish mutant line carrying a mutation in the *dnaaf1* (*lrrc50*) gene [39]. While the homozygous mutants died during larval development due to severe ciliopathy phenotypes, heterozygous mutants develop to adulthood without apparent defects. However, a high seminoma prevalence is observed in the male population during the second and third year of life, with a penetrance exceeding 90% [40]. Subsequently, the authors showed that mutations in *DNAAF1* are also associated with human seminoma. Similarly, works from James Amatruda's lab

(University of Texas Southwestern Medical Center, Dallas, TX, USA) showed that nonsense mutations in the bone morphogenetic protein receptor gene *bmpr1bb* (*alk6b*) is responsible for testicular germ cell tumors [41]. These examples nicely demonstrate that forward genetic screens in zebrafish allows the identification of novel tumor suppressor genes involved in cancer.

The suitability of the zebrafish model in forward genetic screens was further demonstrated by the Nancy Hopkins lab (Massachusetts Institute of Technology, Cambridge, MA, USA) using retroviral insertional mutagenesis [42] (Figure 1b). Amsterdam and collaborators noticed that several heterozygous lines from this large-scale insertional mutagenesis displayed a high mortality rate, sometimes exceeding 50%, with the presence of tumors by the age of 2 years [43].

Out of the 12 heterozygous mutant zebrafish lines developing malignant peripheral nerves sheath tumors (MPNSTs), one of them harbored a mutation in the neurofibromatosis type 2 (*nf2a*) gene, an orthologue of the *NF2* tumor suppressor gene, while all the others carried retroviral insertions within ribosomal protein genes, linking protein biogenesis to cancer.



**Figure 1.** Strategies for forward genetic screens in zebrafish. (a) Chemical mutagenesis. Adult male zebrafish are treated with N-ethyl-N-nitrosourea (ENU) mutagen. Mutations induced in the germ cells can be propagated to the next generation through breeding wild-type females. Each F1 offspring contains a unique set of mutations; (b) Retroviral insertional mutagenesis. High-titer retroviruses are injected into 1K-stage embryos generating F0 mosaic zebrafish. Germ-line mutations are propagated in F1 through mating with wild-type fish.

Retroviral insertion causes gene inactivation by disruption of coding or splicing sequences, but may also be responsible for gene activation due to the enhancer activity of the viral long terminal repeats (LTRs). In this respect, the retroviral insertional screen conducted by Nancy Hopkin and co-workers led to the discovery that retroviral insertions at the *fbxw4* locus are responsible for the overexpression of the *fgf8* neighboring gene that is in turn involved in neuroblastoma tumorigenesis [44].

Due to the late onset and low incidence of the disease, large-scale genetic screens for tumor development in adult zebrafish remains somehow laborious in terms of fish housing because a high number of fish should be hosted during large periods of time. For this reason, several laboratories set up various strategies to screen for cancer-related phenotypes in zebrafish embryo or larvae. The laboratory of Leonard Zon searched for mutations increasing cell proliferation by whole-mount immunohistochemistry. The authors looked for an increase in phosphohistone H3 (H3S10ph) a marker of M-phase cells, in haploid embryos at 36 hours post-fertilization (hpf) and identified the mutants *crash&burn* (*crb*) and *cease&desist* (*cds*) [45,46]. The *crb* mutant harbors a *mybl2* (*bmyb*) loss-of-function mutation responsible for defects in cell cycle progression and for genomic instability [45]. The *cds* mutant carries a mutation in the *separase* (*espl1*) gene coding for a protease cleaving the cohesin complex at the metaphase-anaphase transition and thus involved in the segregation of sister chromatids [46]. Both *mybl2* and *espl1* mutations slightly increase (2 to 2.5-fold, respectively) cancer incidence when zebrafish embryos are treated with MNNG [45,46].

Marco Koudijs and coworkers (Hubrecht Laboratory, Utrecht, The Netherlands) used a similar proliferation screen on a large-scale ENU-mutagenesis mutant library [47]. Whole-mount in situ hybridization on mutant zebrafish embryos at 40 hpf with a proliferating cell nuclear antigen (PCNA) probe was used to identify mutations that disrupt the genes *sufu* (suppressor of fused), *hhip* (hedgehog interacting protein) and *ptach1* (patched 1). These three genes code for repressors of Hedgehog signaling, a pathway which is activated in a number of solid cancers [48,49]. Following a different approach, the laboratory of Joseph Yost (University of Utah School of Medicine, Salt Lake City, UT, USA) described an embryonic screen aiming at the identification of mutants deficient in ionizing gamma irradiation-induced apoptosis and found a mutation I166T in the tumor suppressor Tp53 [50]. Adult zebrafish carrying this mutation are also predisposed to cancer development, predominantly sarcomas.

**Table 2.** Cancer genes identified in zebrafish mutagenesis screens.

## PARTIE II : LE POISSON ZEBRE EN CANCEROLOGIE

Mutagenesis	Mutated Genes	Cancer Phenotypes	References
ENU	<i>dnaaf1 (lrcc50)</i>	Seminoma (incidence: 90% after 2 years of age)	[40]
ENU	<i>bmpr1bb (alk6b)</i>	Testicular germ cell tumor	[41]
Retroviral insertion	Ribosomal proteins ( <i>rps8, rps15a, rpl7, rpl35, rpl36, rpl36a, rpl13, rpl23a, rps7, rps18, rps29</i> )	MPNST (incidence: 14% to 100% at 2 years of age)	[43]
Retroviral insertion	<i>nf2a</i>	MPNST	[43]
Retroviral insertion	<i>fgf8</i> (overexpression)	Neuroblastoma (incidence: 25–50% at 2 years of age)	[44]
ENU	<i>mybl2b (bmyb)</i> (Mutant crash&burn)	Altered cell proliferation and genome instability Increased MNNG-induced cancer susceptibility	[45]
ENU	<i>espl1 (separase)</i> (Mutant cease&desist)	High levels of polyploidy and aneuploidy. Increased MNNG-induced cancer susceptibility	[46]
ENU	<i>sufu</i> (Mutant dre) <i>hhip</i> (Mutant uki) <i>ptch1</i> (Mutant lep)	Activation of the Hedgehog signaling pathway	[47]
ENU	<i>tp53</i>	Deficiency in radiation-induced apoptosis Sarcomas (100% penetrance in adults)	[50]
ENU	<i>llgl2 (lgl2)</i> (Mutant pen)	Epidermal cell tumor	[51,52]
ENU	<i>myh11</i> (Mutant mlt)	Epithelial invasion and cystic expansion of the intestine	[53]

ENU, N-ethyl-N-nitrosourea; MNNG, N-methyl-N'-nitro-N-nitrosoguanidine; MPNST, malignant peripheral nerve sheath tumors.

Other large-scale forward genetic screens conducted at later developmental stages, on zebrafish larvae, also permitted the identification of genes involved in tumorigenesis. Indeed, zebrafish larvae carrying a mutation in the *llgl2* gene develop epidermal cell tumors at 5 days post-fertilization (dpf) [51,52], whereas mutations in the smooth muscle myosin *myh11* gene causes epithelial invasion and cystic expansion in the posterior intestinal lumen of 5 dpf zebrafish larvae [53].

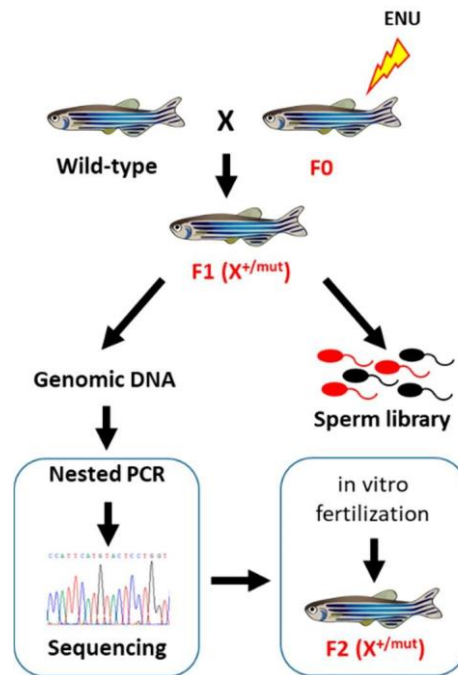
Most of the forward genetic screens aiming at discovering zebrafish cancer-related phenotypes were performed on mutagenized wild-type lines. However, in some cases mutagenesis was conducted on mutant or transgenic lines in order to take advantage of the characteristics of these lines. By outcrossing ENU treated wild-type zebrafish males with *golden* (*gol*<sup>-/-</sup>) mutant females, Keith Cheng and collaborators (Pennsylvania State College of Medicine, Hershey, PA, USA) screened for genomic instability and loss-of-heterozygosity of the golden pigmentation gene which is scored in large batches of embryos by the presence of unpigmented patches of the retinal pigment epithelium of the eye [54]. Twelve *genomic instability* (*gin1* to *gin12*) mutations were identified. Heterozygous adult from all *gin* lines develop various spontaneous tumors in skin, colon, kidney, liver, pancreas, ovary, testis, and neuronal tissues outlining the strong connection between genomic instability and cancer. Unfortunately, the genes carrying the *gin* mutations have not yet been cloned. To identify T-cell cancer phenotypes, the Nikolaus Trede Lab (University of Utah, Salt Lake City, UT, USA) ENU treated Tg(lck:EGFP) male transgenic zebrafish expressing the green fluorescent protein (GFP) in T-cells. Three mutant lines, *hulk* (*hlk*), *shrek* (*srk*) and *oscar the grouch* (*otg*) were identified by fluorescence microscopy for enlarged or extra-thymic GFP expression [55]. These three mutant lines develop T-cell acute lymphoblastic leukemia (T-ALL) malignancies that are phenotypically related to oncogene-induced leukemia. However, the responsible mutations have not been cloned yet, underlying the fact that the identification of cancer genes from ENU-mutagenesis zebrafish screen both requires important zebrafish housing structures and consequent efforts to identify the corresponding mutated genes by positional cloning in the zebrafish cancer lines.

In summary, large-scale genetic screens in zebrafish successfully resulted in the identification of lines displaying mutations in orthologues of mammalian tumor suppressor genes such as *NF2* or *TP53*, as well as in the discovery of a variety of genes that were not previously linked to cancer, including among others, *mybl2*, *dnnaf1* or multiple genes encoding ribosomal proteins (Table 2).

### Reverse Genetics and TILLING

In contrast to the unbiased nature of forward genetic screens that provide the opportunity to uncover novel genes involved in tumorigenesis, reverse genetic approaches allow the generation of zebrafish cancer models carrying mutations in known genes. Targeting Induced Local Lesions IN Genomes (TILLING) is a robust technology based on reverse genetics that may facilitate the identification of loss-of-function, hypomorphic or gain-of-function alleles in theoretically any selected gene. The method was first described in *Arabidopsis thaliana* in 2000 [56,57] but rapidly implemented in other organisms including mouse and zebrafish [58,59]. The approach consists in screening individual genomic DNA samples from a cohort of ENU-mutagenized F1 zebrafish to identify mutations that alter a chosen gene, while the sperm of the corresponding fish is cryopreserved for subsequent reconstitution of the mutant line by in vitro fertilization once desired mutations are identified (Figure 2). The screen for mutations could be done either using the CELI endonuclease cutting DNA heteroduplexes and thus identifying single base pair differences between mutant and wild-type alleles of target genes, and/or using DNA sequencing.





**Figure 2.** Overview of the target-selected mutagenesis in zebrafish. Male zebrafish treated with N-ethyl-N-nitrosourea (ENU) are mated to wild-type females. The resulting F1 male progeny are grown to breeding age and sacrificed to create a mutagenesis library consisting of genomic DNA, used to screen for F1 males harboring mutations in genes of interest (left) and cryopreserved sperm from each individual fish, for recovering the mutant lines using in vitro fertilization (right). This procedure permits the recovery of 50% heterozygous mutant zebrafish in the F2 generation.

The laboratory of Ronald Plasterk (Hubrecht Laboratory, Utrecht, The Netherlands) was the first to apply the TILLING technology to zebrafish [59]. As a proof of concept, Plasterk and coworkers searched for mutations within the *rag1* gene, a gene involved in V(D)J recombination in T-cells, by sequencing the corresponding gene from 2679 F1 fish from a random ENU-mutagenized stock. They identified 15 mutations, including one that is responsible for the production of a truncated protein. The homozygous zebrafish line carrying this *rag1* loss-of-function mutation is deficient in V(D)J recombination but viable and fertile. Subsequently, this target-selected mutagenesis approach was used to generate several mutant lines for known tumor suppressor genes (Table 3).

**Table 3.** Zebrafish mutants generated through the TILLING approach.

Gene	Cancer Types	References
<i>tp53</i>	MPNSTs (Incidence: 28% at 8.5 months of age)	[60]
<i>ptena, ptenb</i>	Ocular hemangiosarcomas (Incidence: 10.2% at 12 months of age)	[61,62]
<i>apc</i>	Intestinal, hepatic and pancreatic proliferative neoplasia (Incidence: 29% at 15 months of age)	[63]
<i>mlh1, msh2, msh6</i>	Primarily neurofibromas in the eye and abdomen. (Incidence: about 33% at 18 months of age)	[64]
<i>brca2</i>	Testicular Neoplasia (Incidence: 31% at 10–16 months of age)	[65]

MPNST, malignant peripheral nerve sheath tumor.

The TILLING methodology has been used to generate a zebrafish line carrying a missense mutation in the DNA binding domain of Tp53 [60]. This mutation *tp53*<sup>M214K</sup> is similar to those found in a number of human cancers. The

homozygous mutant embryos show a reduced gamma irradiation-induced apoptosis but develop normally. Interestingly, at the age of 8.5 months, 28% of *tp53*<sup>M214K</sup> mutant zebrafish develop malignant peripheral nerve sheath tumors (MPNST). Since *TP53* is the most frequently mutated tumor suppressor gene in human cancer, this mutant zebrafish line provides a powerful model to study the role of TP53 in carcinogenesis. The second most frequently mutated tumor suppressor in human cancers is *PTEN*. The study of *pten* in zebrafish is more complex since following a whole genome duplication that has occurred in the teleost genomes, zebrafish has retained two gene copies of *pten* in its genome, *ptena* and *ptenb*. However, mutations in both of these genes have been identified by TILLING [61]. The single homozygous mutants *ptena*<sup>-/-</sup> and *ptenb*<sup>-/-</sup> are viable with no developmental phenotype, whereas mutants lacking both *ptena* and *ptenb* (*ptena*<sup>-/-</sup>; *ptenb*<sup>-/-</sup>) die at day 5 dpf from pleiotropic defects. Interestingly, fish lacking three *pten* alleles (*ptena*<sup>+/-</sup>; *ptenb*<sup>-/-</sup> or *ptena*<sup>-/-</sup>; *ptenb*<sup>+/-</sup>) are viable and fertile but spontaneously develop ocular tumors diagnosed as hemangiosarcomas. This indicates that haploinsufficiency of the genes encoding Pten predisposes to cancer in zebrafish [62]. In humans, mutations in the tumor suppressor gene *APC* constitute the primary transforming event responsible for a large part of the sporadic and inherited colorectal cancers through the constitutive activation of the Wnt/ $\beta$ -catenin signaling pathway. A zebrafish *apc* loss-of-function mutation has been identified from an ENU mutagenesis library and homozygous *apc* mutant zebrafish die between 72 and 96 hpf with multiple defects including most prominently, cardiac malformations [66]. In contrast, the heterozygous *apc* mutant zebrafish do not present developmental defects but spontaneously develop intestinal adenomas, hepatomas and pancreatic adenomas [63]. The tumors accumulate  $\beta$ -catenin and express Wnt target genes showing that the signaling pathway is conserved and that *apc* mutant zebrafish line could serve as a model of familial adenomas polyposis. In humans, rare cases of young patients with brain tumors, skin neurofibromas, and café-au-lait spots that resemble the neurofibromatosis syndrome are associated with deficiencies in the mismatch repair (MMR) machinery genes *MSH2*, *MLH1*, *MSH6* or *PMS2* [67]. Using ENU-driven target-selected mutagenesis, zebrafish mutants with loss-of-function mutations in the *msh2*, *mlh1* and *msh6* genes have been described [64]. Homozygous mutant zebrafish in either one of these genes present genomic instability and develop tumors at an early age and low frequencies.

These tumors are predominantly neurofibromas that mimic an important part of the phenotype of human patients, where biallelic MMR inactivation causes a neurofibromatosis type I-like phenotype. Inherited mutations in *BRCA2* predispose to breast and ovarian cancer. TILLING identified a nonsense mutation in *brca2* (*brca2*<sup>Q658X</sup>) at a gene position similar to *BRCA2* mutations found in humans with hereditary breast and ovarian cancers. Homozygous *brca2* mutant zebrafish are viable but fail to develop ovaries and develop into infertile males with deficient spermatogenesis [65]. Furthermore, *brca2* mutant zebrafish are predisposed to testicular neoplasia, while tumorigenesis is enhanced in a *tp53* deficient genetic background [65,68]. Thus, the *brca2* mutant zebrafish line shed new light on the role of *brca2* in ovary development and tumorigenesis in reproductive organs, as well as in cancer development associated to heritable *BRCA2* and *TP53* mutations.

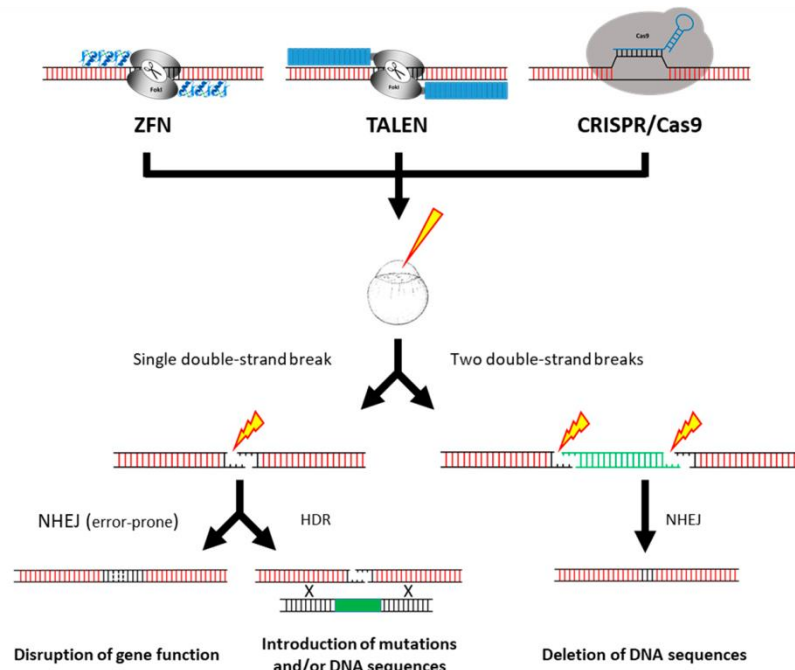
Altogether, TILLING is a powerful technology that raise zebrafish as a pertinent model in cancer research. However, the approach suffers from several drawbacks: (i) The procedure is costly, labor intensive and cannot be implemented in most individual labs. (ii) The approach relies on random chemical mutagenesis. It might statistically be more difficult to obtain mutants for small-size genes and some sequences could have a weak mutagenic potential. (iii) The use of random chemical mutagenesis makes that in addition to the mutation of interest, several other mutations are present in the fish genome. ENU-induced mutants should be outcrossed for several generations in order to exclude effects of unknown additional mutations.

### Gene Editing Technologies

The discovery of the programmable site-specific endonucleases, as the most versatile and efficient tools to modify any genomic sequence, completely revolutionized the field of reverse genetics applied to the zebrafish disease model [69,70]. Programmable site-specific endonucleases are engineered to induce double-strand DNA breaks specifically at chosen genomic target sites. These double-strand DNA breaks subsequently stimulate cellular DNA repair mechanisms such as the error-prone non-homologous end joining (NHEJ) and the homology-directed repair (HDR). The NHEJ mechanism, which often results in small random nucleotide insertions or deletion (*indel*) at the cut site, can be used to disrupt the function of a gene. In contrast, HDR allows the insertion of precise genomic



modifications in presence of a designed homologous DNA template. Another application of programmable site-specific endonucleases is the generation of large deletions, where two distant double-strand DNA breaks are induced on the same genomic DNA molecule (Figure 3). The programmable site-specific endonucleases used for genome editing are the zinc-finger nucleases (ZFNs), the transcription activator-like effector nucleases (TALENs) and the clustered regularly interspaced short palindromic repeats (CRISPR) RNA-guided Cas9 nucleases (CRISPR/Cas9) [71].



**Figure 3.** Gene targeting using programmable site-specific endonucleases. RNAs and/or proteins corresponding to the components of the ZFNs, TALENs and CRISPR/Cas9 programmable site-specific endonuclease systems are injected into zebrafish embryos at the 1-cell stage to induce double-strand DNA breaks at the target genomic sites. Cellular DNA repair mechanisms including non-homologous end joining (NHEJ) and homology-directed repair (HDR) are responsible for gene editing events such as the disruption of gene function, gene deletion or the introduction of precise modifications.

### Zinc Finger Nucleases

Zinc finger nucleases (ZFNs) are chimeric proteins comprising a DNA binding domain composed of C<sub>2</sub>H<sub>2</sub> type zinc fingers mainly derived from the transcriptional factor ZIF268/EGR1, and a cleavage domain from the bacterial endonuclease FokI [72]. The alpha helix of each zinc finger recognizes and binds to a 3 bp DNA sequence. Since the DNA binding domain of ZFNs is usually an array of three or four zinc fingers, it would recognize a nine to 18 bp target DNA sequence. Because FokI acts as a dimer, ZFNs function in pairs and cleave a spacer sequence to induce a double-strand DNA break located between the two ZFN binding sites. ZFNs were the first available direct gene targeting method in zebrafish, thus making any zebrafish gene knockout possible [73–75]. However, the engineering and selection of ZFNs with efficient and specific DNA binding activity remains highly challenging. In particular, the fact that the binding specificities of individual zinc fingers could overlap and depend on the context of the surrounding zinc fingers and DNA sequences renders their assembly difficult and has strongly limited the ZFN usages.

In spite of the difficulties associated to the design of efficient and specific ZFNs, the Thomas Look laboratory (Children’s Hospital, Harvard Medical School, Boston, MA, USA) successfully applied these programmable site-specific endonucleases to the inactivation of several cancer genes in zebrafish [76,77] (Table 4).

## PARTIE II : LE POISSON ZEBRE EN CANCEROLOGIE

**Table 4.** Cancer genes targeted by programmable site-specific endonucleases in zebrafish.

Nuclease	Mutated Genes	Cancer Types	References
ZFN	<i>nf1a, nf1b</i>	High grade gliomas and MPNSTs in the <i>tp53</i> <sup>-/-</sup> genetic background	[76]
ZFN	<i>tet2</i>	Generalized MDS at 24 months of age	[77]
TALEN	<i>rb1</i>	Primitive neuroectodermal tumors (Incidence: 11–33% at 3.5 months of age in F0 mosaics)	[78–80]
TALEN	<i>cdkn2a/b</i>	MPNSTs in F0 <i>tp53</i> <sup>-/-</sup> mutants (Incidence: 39% at 35 weeks of age)	[79]
TALEN	<i>tp53del/del</i>	Angiosarcoma, MPNSTs, germ cell tumors (Incidence: 37% at 12 months of age)	[81]
TALEN CRISPR	<i>irx1a</i> <i>irx1b</i>	Intestinal hyperplasia and testicular, ovarian, renal and bile duct tumors	[82]
CRISPR	<i>spred1, tp53, ptena/b, cdkn2a</i>	Melanocyte-specific tumor suppressor gene inactivation promotes KIT-, BRAF- and NRAS-driven melanoma formation	[83]
CRISPR	<i>ptch1</i>	T-cell-specific inactivation promotes NOTCH1 <sup>ICD</sup> -driven T-ALL development	[84]
CRISPR	<i>atrx</i>	Hemizygous loss of <i>atrx</i> promotes the development of epithelioid sarcoma, angiosarcoma and pleomorphic sarcoma in <i>tp53/nf1</i> -deficient mutants	[85]
CRISPR	<i>suz12a, suz12b</i>	Loss of <i>suz12</i> function accelerates the onset of MPNST and expands the spectrum of tumor types in <i>tp53/nf1</i> -deficient mutants	[86]
CRISPR	<i>arid1aa, arid1ab</i>	<i>arid1aa</i> or <i>arid1ab</i> deficiency increases the penetrance of MYCN-induced neuroblastoma in Tg(dbh:EGFP-MYCN) transgenics	[87]

## PARTIE II : LE POISSON ZEBRE EN CANCEROLOGIE

CRISPR	<i>lats2</i>	PNST in F0 mosaics	[88]
Tumor formation in the digestive system			
CRISPR	<i>dact2</i>	(active proliferation and hyperplasia in the pancreatic duct, the intrahepatic bile duct, the intestinal epithelium, pancreas and liver)	[89]
Loss of <i>twist3</i> prevents			
CRISPR	<i>twist3</i>	BRAF <sup>V600E</sup> -mediated tumorigenesis in thyroid	[90]

MDS, myelodysplastic syndrome; MPNST, malignant peripheral nerve sheath tumor; PNST, peripheral nerve sheath tumor; T-ALL, T-cell acute lymphoblastic leukemia.

As for *PTEN*, the zebrafish genome codes for two orthologues of *NF1*, *nf1a* and *nf1b*. ZFN-mediated inactivation of these genes revealed that homozygous loss-of-function of *nf1a* or *nf1b* alone leads to phenotypically normal zebrafish. In contrast, loss of both genes results in larval lethality with phenotypes that resemble aspects of the human neurofibromatosis type 1 [76]. Moreover, Nf1 cooperate with Tp53 to the development of high-grade gliomas and MPNSTs in adult zebrafish. Indeed, the combined loss of *tp53* and three of four *nf1* alleles (*nf1a*<sup>+/-</sup>; *nf1b*<sup>+/-</sup>; *tp53*<sup>+/-</sup>) significantly accelerates the onset and increases the penetrance of tumors as compared with the loss of *tp53* alone (*nf1a*<sup>+/+</sup>; *nf1b*<sup>+/+</sup>; *tp53*<sup>+/-</sup>) or with the concomitant loss of *tp53* and both alleles of *nf1b* with intact *nf1a* (*nf1a*<sup>+/+</sup>; *nf1b*<sup>-/-</sup>; *tp53*<sup>+/-</sup>). On the other side, ZFN-inactivation of *tet2*, a gene coding for one of the TET family members of DNA methylcytosine oxidases that converts 5-methylcytosine into 5-hydroxymethylcytosine during the DNA demethylation process, is responsible for generalized myelodysplastic syndromes (MDS) in zebrafish [77]. The *tet2*<sup>-/-</sup> zebrafish mutants are viable and constitute a good model of MDS since somatic loss-of-function mutations of *TET2* are frequently found in human MDS.

### ***Transcription Activator-Like Effector Nucleases***

Transcription activator-like effector nucleases (TALENs) are similar to ZFNs in their architectures. Both programmable site-specific endonucleases contain the endonuclease domain of FokI, but differ in their DNA-binding domain moiety. The DNA binding domain of the TALENs derives from transcription activator-like effectors (TALEs) which are transcription factors from phytopathogens of the *Xanthomonas* genus. The DNA binding domain of TALE proteins is composed of an array of several 34 amino acids repeat units, with each repeat recognizing and binding to a single nucleotide in the target DNA sequence [91,92]. Within these 34 amino acids conserved modules, two residues at position 12 and 13, called RDV (repeat variable di-residue) determine the binding specificity to its target nucleotide. For instance, in the natural TALEs, the most frequent RDVs, NI (Asn Ile), NG (Asn Gly), HD (His Asp) and NN (Asn Asn)/NK (Asn Lys) specifically recognize the nucleotides A, T, C and G, respectively. Thus, compared to ZFNs, the DNA recognition basis of TALENs is much simpler and predictable, making the TALENs an extremely efficient programmable site-specific endonuclease system. With 15–17 repeated modules, the TALEN technology offers a way to introduce double-strand DNA breaks with a high specificity and a strong efficacy in any gene and any organism including zebrafish [93,94].

TALENs have been shown to be efficient enough to induce somatic mosaic bi-allelic mutations in tumor suppressor genes in F0 fish allowing the study of their function in tumorigenesis, even when homozygous mutants are larval lethal [78,79] (Table 4). TALEN-mediated somatic bi-allelic inactivation of *rb1* induces medulloblastoma-like primitive neuroectodermal tumors in wild-type zebrafish [79,80], whereas *cdkn2a/b* inactivation induces MPNST

in the brain with high frequencies and early onset in F0 *tp53*-mutant zebrafish [79]. Using two pairs of TALENs designed to target both sides of the *tp53* gene (Figure 3), David Langenau and co-workers (Massachusetts General Hospital Research Institute, Boston, MA, USA) generate a *tp53<sup>del/del</sup>* zebrafish line [81]. Homozygous *tp53<sup>del/del</sup>* mutants start to develop tumors at 4 months of age, mainly leukemia with blast-like cells in the kidney marrow. At 7 months, angiosarcoma, MPNSTs or germ cell tumors are externally visible in a subset of *tp53<sup>del/del</sup>* zebrafish. This tumor spectrum is wider than what was observed for *tp53<sup>M214K</sup>* mutants identified by TILING and *tp53<sup>I166T</sup>* identified in an ENU forward genetic screen [50,60], indicating that tumor onset and phenotype might depend on the nature of the *tp53* mutation.

### ***Clustered Regularly Interspaced Short Palindromic Repeats RNA-Guided Cas9 Nucleases***

Cas9 is a dual RNA-guide DNA endonuclease associated with the clustered regularly interspaced short palindromic repeats (CRISPRs) playing a role of an adaptive immune system in archea and bacteria. In brief, a guide RNA (gRNA) recognizes a specific DNA sequence by hybridization and targets the Cas9 endonuclease that induces a double-strand DNA break at the RNA recognition site [95]. In contrast to ZFNs and TALENs that target DNA through a protein-DNA interaction, the recognition of the CRISPR/Cas9 target site is achieved via a RNA-DNA hybridization. In addition, while ZFNs and TALENs act as dimers, Cas9 induces double-strand DNA breaks as a monomer. The CRISPR/Cas9 can be programmed to cut different target sites by changing 20 nucleotides within the gRNA. The only limitation in designing a target site is the requirement for a protospacer adjacent motif (PAM) (NGG in the case of the *Streptococcus pyogenes* spCas9 enzyme) just adjacent to the 20-nt target sequence. Due to its remarkable simplicity, the CRISPR/Cas9 system rapidly became a genome-editing tool of choice in a wide variety of experimental model organism, including zebrafish [96]. The first use of CRISPR/Cas9-mediated mutagenesis in zebrafish was demonstrated by Woong Hwang and colleagues

(Massachusetts General Hospital, Charlestown, MA, USA) in 2013 [97]. The authors reported the targeting of 10 loci and showed mutagenesis rates between 24% and 59% in somatic tissues. At around the same time, Nannan Chang *et al.* (Peking University, Beijing, China) showed that the CRISPR/Cas9 is able to induce somatic bi-allelic conversion [98], while Li-En Jao *et al.* (Vanderbilt University School of Medicine, Nashville, TN, USA) demonstrated that CRISPR/Cas9-induced mutations are heritable, and the possibility to target five genomic loci simultaneously, resulting in multiple loss-of-function phenotypes in the same injected zebrafish [99]. Furthermore, the possibility to express the Cas9 protein under the control of tissue-specific promoters in transgenics renders tissue-specific gene disruption feasible in zebrafish [100,101].

While the CRISPR/Cas9 system appears to be an efficient and simple tool for genomic engineering, there are still several concerns over its specificity and possible unintended off-targets [102]. In zebrafish a study detected CRISPR/Cas9-mediated off-target mutagenesis in only 1/25 predicted off-target sites in germline, whereas in another study the off-target mutation rate was estimated between 1.1 to 2.5% [103,104]. These potential off-targets may easily be outcrossed away. Moreover, different Cas9 variants, including Cas9-HF1, eSpCas9, evoCas9 and HypaCas9, have been developed or under development to increase the specificity of the Cas9 enzyme and thus reduce the off-targets [105].

Despite a recent implementation, the CRISPR/Cas9 system has been already applied to the zebrafish cancer research field [82–90] (Table 4). Iroquois homeobox 1 (IRX1) is an homeobox-containing transcription factor playing a role in embryonic development and suspected to act as a tumor suppressor gene in head and neck squamous cell carcinoma, gastric cancers and glioma [106–109]. In zebrafish, *Irx1* is encoded by two ohnologues *irx1a* and *irx1b*. The group of Seung Woo Park (Yonsei University College of Medicine, Seoul, Korea) used the TALEN technology to inactivate *irx1a*, while *irx1b* mutants were generated using the CRISPR/Cas9 system [82]. Homozygous mutants for both *irx1a* and *irx1b* (*irx1a<sup>-/-</sup>*; *irx1b<sup>-/-</sup>*) are malformed and die within 6 months of age. In contrast, single homozygous mutants for *irx1a* (*irx1a<sup>-/-</sup>*) and *irx1b* (*irx1b<sup>-/-</sup>*) are viable and fertile but spontaneously develop hyperplasia and tumors in different organs where the genes are expressed, including intestine, testis, ovary, kidney and bile duct. In a search for genetic alterations in mucosal melanomas, *SPRED1*, a negative regulator of MAPK signaling, was found inactivated in 37% of the tumors [83]. Through the expression of Cas9 under the control of

the *mitfa* promoter in F0 mosaic zebrafish, Ablain *et al.* [83] showed that the melanocyte-specific targeting of *spred1* in addition to *ptena/b* and *tp53* results in late-onset melanomas,

which is not observed through *ptena/b* and *tp53* targeting alone. This CRISPR/Cas9-mediated mosaic tissue-specific gene knockout could be combined to the concomitant expression of oncogenes in the F0 zebrafish through the co-injection of vectors driving the expression of Cas9, gRNAs and oncogenes. For instance, it has been shown that *NRAS<sup>Q61R</sup>* expression combined with *tp53* loss in melanocytes is responsible for tumor formation within 3 week after plasmid injection in the embryos. In contrast, *BRAF<sup>V600E</sup>* expression together with *cdkn2a* loss-of-function leads to melanoma after several months [83]. Similarly, using the *rag2* promoter to drive the expression of Cas9 and the Notch1a<sup>CD</sup> oncogene into immature T-cell progenitors, the Alejandro Gutierrez group (Boston Children's Hospital, Boston, MA, USA) showed that inactivation of *ptch1* accelerates the Notch1a<sup>CD</sup>-induced T-ALL onset [84]. These examples illustrate the usefulness of the CRISPR/Cas9 system to assay for tumor suppressor gene activities. On the other side, the approach may also be used to characterize genes promoting tumorigenesis. Using an RNAseq data set from human papillary thyroid cancers, Viviana Anelli *et al.* (New York Presbyterian Hospital, New York City, NY, USA) [90], identify TWIST2 as a key effector upregulated by the oncogenic BRAF<sup>V600E</sup> mutation. The authors showed that the CRISPR/Cas9-inactivation of *twist3*, an orthologue of TWIST2, in a zebrafish *BRAF<sup>V600E</sup>* transgenic model, rescues defects in follicle structure and thyroid hormone production. These data indicate that *twist3* contributes to *BRAF<sup>V600E</sup>*-mediated transformation [90].

In sum, programmable site-specific endonucleases have been successfully used in zebrafish (i) to generate models harboring mutations in tumor suppressor genes (e.g., *nf1*, *rb1*; [76,78]); (ii) to engineer novel genetic alterations in tumor suppressor genes (e.g., *tp53<sup>del/del</sup>*; [81]); (iii) to demonstrate the tumor suppressor function of novel candidate genes (e.g., *irx1*, *spred1*, *arid1a* [82,83,87]); (iv) to characterize the role of genes involved in tumor development (e.g., *twist3*; [90]); (v) as well as to investigate the cooperation between different mutations in the tumorigenesis onset (e.g., *tp53* and *nf1* with *atrx* or *suz12* [85,86]). The use of programmable site-specific endonucleases can be applied to the engineering of zebrafish mutant lines [76,77], but also to the rapid generation of F0 mosaic somatic mutants [78,79,83,84,90]. In addition, expression of Cas9 under the control of tissue-specific promoters allows CRISPR/Cas9-mediated inactivation of cancer genes in specific tissues [83,84]. Finally, the programmable site-specific endonucleases have been mainly used to induce *indel* mutations and in some cases, used to generate genomic deletions via the targeting of two double-strand breaks [81,88]

(Figure 3). However, to our knowledge, precise point mutations and knock-ins generated through HDR have not yet been reported in zebrafish in the context of cancer research.

### Transgenesis

The ability to generate transgenic lines that stably express genes (e.g., genes involved in diseases or fluorescent marker genes) is central to numerous biomedical studies using zebrafish as a model organism. The Monte Westerfield lab (University of Oregon, Eugene, OR, USA) pioneered the stable introduction of exogenous DNA into the zebrafish genome in the late 1980s [110]. After injection of a linearized plasmid into the cytoplasm of fertilized eggs, these first transgenic zebrafish transmitted the bacterial DNA to their progeny with a germline transmission rate of about 15%. However, gene expression from the transgenic cassettes was often subjected to transcriptional silencing, presumably as the result of methylation and rearrangements of the foreign DNA integrated as concatemers [111–113]. This limitation was bypassed by flanking the transgenic cassettes with 18-bp recognition sites for the

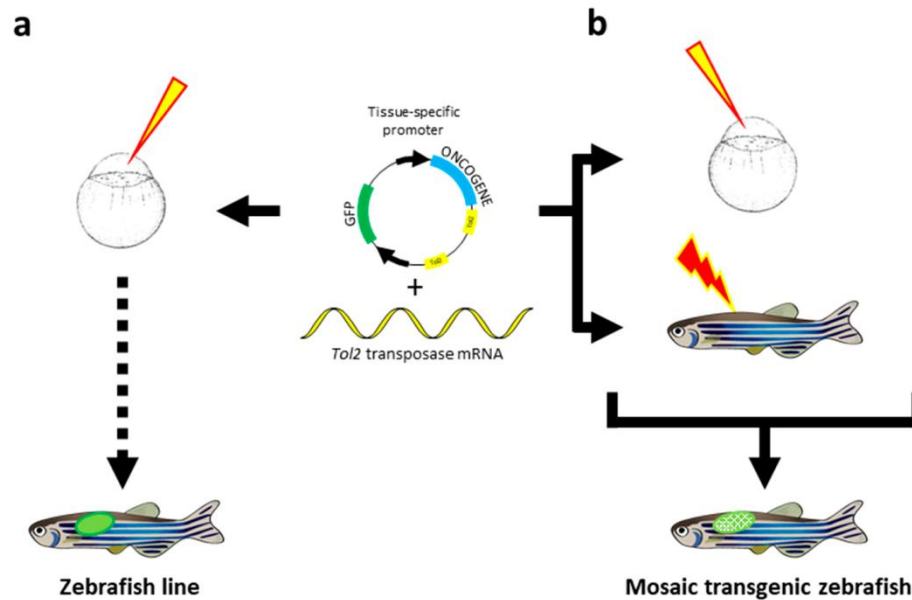
*Saccharomyces cerevisiae* I-SceI meganuclease enzyme. Co-injection into the embryo of the enzyme together with the transgenic vector carrying the restriction sites, linearizes the circular DNA before genome integration, reduces concatemerization, and favors transgene expression. Although this strategy was first implemented in the Japanese rice fish medaka (*Oryzias latipes*) by the group of Jean-Stéphane Joly (Institut de Neurobiologie A. Fessard, Gif-Sur-Yvette, France) [114], the approach became widely applied to the generation of transgenic zebrafish lines [115]. An alternative method based on the *Tol2* transposon identified in the medaka fish has been developed by Koichi Kawakami (The Graduate University of Advanced Studies, Mishima, Shizuoka, Japan) [116–118]. *Tol2* is an active transposon element that belongs to the hAT transposon family. The principle of its use in transgenesis is based on



the separation of the cis transposable elements called the short terminal repeats from the transposase enzyme. The transposase is in vitro transcribed and co-injected with a circular plasmid carrying the transgenic cassette flanked by the cis regulatory short terminal repeats (Figure 4a). This system allows the insertion of up to 11 kbp sequences of foreign DNA into the genome and produces germ-line transmission of the transgene in up to 50% of the injected zebrafish [119]. Due to transposon-mediated insertion, the prokaryotic sequences from the plasmid backbone are not integrated in the zebrafish host genome and transgenesis results in the integration of single copy per locus avoiding concatemer formation. The absence of concatemer integration together with the exclusion of the prokaryotic sequences prevents transgene silencing through heterochromatinization or chromosomal rearrangements. Finally, *Tol2*-mediated transgenesis usually leads to single copy insertions into the genome. All of these attributes made the *Tol2*-based approach a powerful and

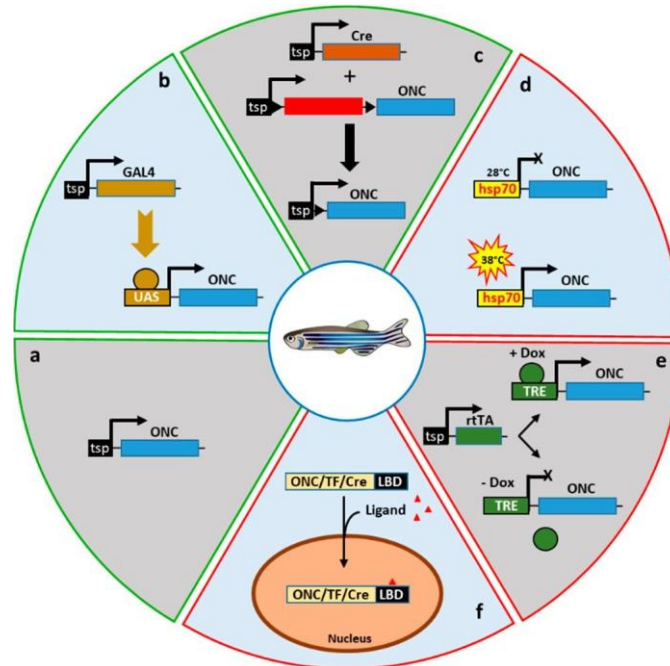
widely used methodology to generate transgenic zebrafish [120].

In 2003, in a landmark publication, David Langenau and coworkers (Harvard Medical School, Boston, MA, USA) were the first to demonstrate transgenic modeling of cancer in zebrafish [121]. The expression of murine *Myc* oncogene fused to GFP was targeted to the developing lymphocytes using the zebrafish *rag2* promoter. The transgenics developed T-ALL between 1 month and 5 months of age and the GFP-labeled leukemic cells were shown spreading from the thymus into the surrounding tissues and invading skeletal muscle and visceral organs. This study demonstrated that the expression of an oncogene is able to drive cancer formation in zebrafish. Furthermore, the co-expression of fluorescent markers allows in vivo monitoring of tumor progression in real-time, tumor cell isolation for further phenotypic or transcriptomic characterization as well as for serial transplantation into irradiated recipient zebrafish. These observations substantiate the power of the zebrafish model in cancer research and have opened avenues to the development of zebrafish transgenic lines applied to cancer modeling. Thus, over the years a large number of zebrafish lines expressing various oncogenes have been generated (Table A1).



**Figure 4.** Tol2-mediated transgenesis in zebrafish. (a) Stable transgenesis. Transposase mRNA synthesized by in vitro transcription and a plasmid DNA harboring the *Tol2* cis regulatory elements (yellow) are co-injected into zebrafish fertilized eggs. The transposase synthesized from mRNA catalyzes the integration of the excised transgenic sequences placed between the *Tol2* short terminal repeats into the genome. The injected embryos are raised and crossed with wild-type fish. The integrated transgene is transmitted to the F1 generation showing non-mosaic transgene expression. (b) Mosaic F0 transgenic zebrafish can be created either through the microinjection into fertilized eggs or through electroporation in adults (TEAZ) of the *Tol2* transposase and the construct containing the cis regulatory elements. A fluorescent marker gene (e.g., GFP) is often present in the construct to identify the transgenics.

However, transgenic zebrafish expressing oncogenes could sometime not survive sexual maturity. The transgenic Tg(rag2:GFP-Myc) zebrafish line for instance, shows a 100% leukemia incidence with a cancer onset mostly preceding reproductive age. Consequently, maintenance and propagation of the line requires labor-intensive in vitro fertilization [122]. Thus, the full exploitation of the transgenic zebrafish model in cancer research necessitates the development of strategies to allow conditional expression of the oncogenes in a spatial- and temporal-controlled fashion. Moreover, cancer phenotypes mostly result from somatic mutations rather than from germline genetic mutations. Then, controlled activation of oncogenes in zebrafish somatic tissues would also accurately mimic cancer onset. Fortunately, an exceptional toolbox of conditional transgenic approaches mainly adapted from *Drosophila* and mouse genetics, has been established in zebrafish (Figure 5).



**Figure 5.** Strategies to control oncogene expression in zebrafish cancer models. (a) Spatial regulation of oncogene (ONC) expression through the activity of tissue-specific promoter (tsp). (b) The Gal4/ UAS system to enable cell type specific oncogene expression. The yeast transcription factor Gal4 is expressed through the control of a tissue-specific promoter (tsp) and binds to the upstream activating sequence (UAS) to drive the expression of the oncogene (ONC). (c) Cre-mediated oncogene (ONC) expression. Cre-mediated recombination removes an intervening sequence and places the oncogene under the control of a promoter. (d) Temporal oncogene (ONC) activation through the induction of the activity of the heat shock promoter. (e) Temporal oncogene (ONC) activation using the Tet-ON system. The binding of doxycycline to rtTA allows the transcription factor to bind and activate transcription at tetO responsive elements (TRE)-containing promoters. (f) Ligand-dependent oncogene (ONC) activation. Fusing the ligand-binding domain (LBD) onto oncogenic transcription factors (MYC), transcription factors (LexA) or Cre recombinase confers ligand-dependent activities to these proteins. Strategies to spatially control oncogene expression are boxed in green, whereas temporal to control of oncogene expression is boxed in red.

### Tissue-Specific Promoters

In zebrafish, a number of tissue-specific promoters have been reported to faithfully reproduce endogenous gene expression patterns in transgenics. These tissue-specific promoters may be used to drive oncogene expression in define tissues (Figure 5a). In their princeps experiments, Langenau and colleagues [121] used the *rag2* promoter to drive expression of the *Myc* oncogene and induce T-ALL. However, since B-lymphoblasts also express *rag2* and because *MYC* is known to drive human B-ALL, the Tg(rag2:GFP-Myc) and Tg(rag2:MYC) transgenic zebrafish also develop B-ALL [123–126]. While the *rag2* gene shows a specific expression in immature lymphoid cells, the zebrafish *rag2* promoter used in

transgenics is active in both lymphoid and nonlymphoid cell populations including olfactory rosettes, sperm and mesenchymal progenitor cells [127–129]. Thus, the ectopic expression properties of the zebrafish *rag2* promoter in the mesenchymal cell compartments has been used to drive expression of *KRAS<sup>G12D</sup>* or a myristoylated constitutively active form of *Akt2* in order to generate transgenic models of rhabdomyosarcoma or liposarcoma, respectively [129,130]. Apart from *rag2*, a number of tissue-specific promoters have been used to drive oncogene expression into transgenic zebrafish and to induce cancer formation. These tissue-specific promoters include the myeloid-specific promoter *spi1* to induce AML [131–133], *mitfa* (melanocyte inducing transcription factor a) to induce melanoma [134–137], *cdh15* (M-cadherin) and *mylz2* (myosin light chain) to induce rhabdomyosarcoma [138], *ptfla* (pancreas associated transcription factor 1a) or *myod* to induce exocrine or endocrine pancreatic carcinomas respectively [139,140], *fabp10a* (fatty acid binding protein 10a) to induce hepatocellular carcinoma or intrahepatic cholangiocarcinoma [141–148], *dbh* (dopamine  $\beta$ -hydroxylase) and *sox10* to induce brain tumors [149–151], *pomc* (proopiomelanocortin) to induce pituitary adenoma [151], *tg* (thyroglobulin) to induce papillary thyroid carcinoma [90] or *flick* (Fugu lymphocyte-specific protein tyrosine kinase) to induce testicular germ cell tumors [152] (Table A1).

Elizabeth Patton, Leonard Zon and colleagues (Children’s Hospital, Boston, MA, USA) used the melanocyte-specific promoter *mitfa* to drive the expression of the *BRAF<sup>V600E</sup>* oncogene [134]. The transgenic line develops patches of ectopic melanocytes, termed fish (f)-nevi, but not melanoma. This is similar to human pigmented nevi, a benign lesion that also carries *BRAF<sup>V600E</sup>* mutations at a high frequency [153]. However, when present in a *tp53<sup>-/-</sup>* genetic background, *BRAF<sup>V600E</sup>* expression leads to highly invasive and transplantable melanoma, highlighting a synergistic interaction between the BRAF and TP53 pathways in the development of melanoma. *NRAS<sup>Q61K</sup>* and *HRAS<sup>G12V</sup>* transgenic zebrafish have also been generated to model human melanoma [135,136]. While *NRAS<sup>Q61K</sup>*-dependent progression to melanoma requires concomitant *tp53* loss-of-function mutation, *HRAS<sup>G12V</sup>* targeted expression alone causes ectopic melanocyte formation during early embryogenesis, melanocyte hyperplasia, dysplasia, invasion of loose connective tissues and rapid progression to deeply invasive melanoma. As such, Tg(*mitfa:HRAS<sup>G12V</sup>*) transgenics may serve as a model for familial atypical mole and melanoma (FAMM) syndrome. In addition to *BRAF<sup>V600E</sup>* and *NRAS<sup>Q61K</sup>*, a number of oncogenes have been shown to synergize with *tp53* mutations in transgenic zebrafish. For instance, wild-type transgenic zebrafish expressing the HBV X (HBx) antigen under the control of the liver-specific promoter *fabp10a* do not develop hepatocellular carcinoma (HCC) whereas *tp53<sup>-/-</sup>* transgenics do develop HCC [142,154]. In the same way, wild-type transgenic zebrafish expressing constitutively active *Akt2* in mesenchymal progenitors develop well-differentiated liposarcoma (WDLPS) at an incidence rate of 8%, while this rate raises to 29% in a *tp53<sup>-/-</sup>* genetic background [130]. Similarly, loss of *nf1* function accelerates disease onset and increases the penetrance of MYCN-induced neuroblastoma in Tg(*dbh:GFP-MYCN*) transgenic zebrafish [155]. The transgenic line Tg(*rag2:KRAS<sup>G12D</sup>*) expresses *KRAS<sup>G12D</sup>* in undifferentiated muscle satellite cells and develop tumors that resemble embryonal rhabdomyosarcoma, the most common childhood soft-tissue tumor [129]. Co-injection of multiple transgenes driving the expression of fluorescent proteins under the control of promoters reflecting different stages of the muscle cell differentiation is an efficient method to label cell subpopulations. Applied to the Tg(*rag2:KRAS<sup>G12D</sup>*) model, this elegant strategy shows that rhabdomyosarcoma are composed of distinct cell populations that are dynamically reorganized during tumor growth [129,156].

Thus, the transgenic approach based on the expression of oncogenes under the control of tissue-specific promoters, eventually in combination with other markers or mutations, is a straightforward method to model specific cancer types in zebrafish. However, in some cases, the use of ubiquitous promoters has



also proven efficacy in zebrafish cancer modeling. TEL-AML1 (also known as ETV6-RUNX1), generated by the t(12;21)(p13;q22) chromosomal translocation is the most common chimeric fusion gene in childhood cancer and is selectively associated with pre-B acute lymphoblastic leukemia (ALL). TEL-AML1 expression under the control of zebrafish *actin* or *Xenopus elongation factor 1 $\alpha$*  ubiquitous promoters leads to lymphoid progenitor expansion that evolved in oligoclonal B-lineage ALL in about 3% of the transgenic zebrafish at the 8 to 12 months of age [157]. Interestingly, lymphoid-specific expression of TEL-AML1 from the *rag2* promoter failed to cause lymphoid hyperplasia, suggesting that the fusion protein acts prior to the committed lymphoid progenitor stage. Of note, in transgenic mouse models, expression of TEL-AML1 does not result in any hematological disorder, unless these mice are subjected to ENU or extremely low-frequency (ELF) magnetic fields exposure [158–160], highlighting the value of the TEL-AML1 transgenic zebrafish model.

### *The Gal4/UAS System*

The Gal4/UAS system is a binary system derived from the yeast *Saccharomyces cerevisiae* and used to drive conditional transgene expression in zebrafish [161]. It consists of the transcriptional activator Gal4 that controls gene expression through the binding to its UAS (upstream activation sequence) recognition DNA motif. To achieve tissue-specific gene expression, Gal4 or a Gal4 derivative, is placed under the control of a tissue-specific promoter whereas a minimal promoter containing UAS sequences drives the expression of the gene of interest. The gene of interest whose expression strictly depends on Gal4 binding to the UAS will be exclusively expressed in the tissues where Gal4 is present (Figure 5b). The Gal4/UAS expression system has been implemented in various animal models including *Drosophila* [162], *Xenopus* [163] or mice [164], while Nico Scheer and José Campos-Ortega (Universität zu Köln, Cologne, Germany) were the first to apply this conditional expression technique to transgenic zebrafish in 1999 [165]. A key property of this system is the separate integration of the two components in a driver line expressing Gal4 and in an effector UAS line allowing the silent propagation of toxic or lethal genes and their conditional activation only in double-transgenic offspring from the crosses between driver and effector lines.

The versatility of the Gal4/UAS system has largely been applied to the generation of zebrafish cancer models. A transgenic Tg(UAS:GFP-HRAS<sup>G12V</sup>) line develops melanoma when crossed with lines expressing Gal4 under the *kita* enhancer [166], leukemia when crossed with lines expressing Gal4 under the *fli1* promoter [167], glioma when crossed with lines expressing Gal4 under the *zic4* enhancer [168] or chordoma when crossed with lines expressing Gal4 under the notochord-specific promoter *shhb* (also known as *twhh*, *tiggywinkle* hedgehog) promoter [169]. On the opposite, a same driver line can be used to assay for the tumorigenic activity of different oncogenes or diverse mutations within the same oncogene. For instance, while the *ptf1a* (pancreas associated transcription factor 1a) promoter is active in cerebellum, hindbrain and pancreas, Tg(*ptf1a*:Gal4; UAS:KRAS<sup>G12V</sup>) double-transgenics develop pancreatic tumors [170], whereas Tg(*ptf1a*:Gal4; UAS:AKT1<sup>Myr</sup>) present glioma [171]. In another study, the Steven Leach lab (Johns Hopkins Medical Institutions, Baltimore, MD, USA) used a Tg(*ptf1a*:Gal4) driver line to compare the ability of 12 different KRAS mutations to drive pancreatic tumorigenesis in vivo [172].

### *Cre-Mediated Recombination*

The Cre-mediated recombination is another way to control transgene expression. Cre is a site-specific recombinase recognizing and inducing recombination at specific 34 bp target sites called loxP sites. Depending on the orientation and location of loxP sites on the DNA molecules, Cre-mediated recombination is able to induce different DNA rearrangement events. These events are the deletion of a DNA sequence located between two loxP sites in direct orientation, the inversion of a DNA sequence located between two loxP sites in an inverted orientation, the insertion of a circular DNA molecule containing a loxP site into another loxP site-containing DNA molecule or a chromosomal translocation between two loxP site-containing linear DNA molecules. In transgenic mouse cancer models, the Cre-loxP conditional system has been used to achieve tissue-specific expression of transgenes, often recapitulating aspects of the human disease [173,174]. In a classical setting, a target transgenic mouse line carrying a promoter and an oncogene separated by an intervening sequence that contains elements preventing expression of the oncogene and flanked by loxP sites in direct orientation is mated with an effector

transgenic mouse line expressing the Cre recombinase. In the double-transgenic offspring, Cre-mediated recombination occurs at the loxP sites, the intervening STOP sequence is deleted, and the oncogene is then placed under the control of the promoter leading to tumor formation [173] (Figure 5c). The tight regulation of transgene expression by this binary system is essential since many oncogenes play crucial roles during development and continuous oncogene expression may cause embryonic lethality, abnormal development or infertility in the transgenics preventing the maintenance of the lines.

The first group to apply the Cre-loxP system to zebrafish cancer modeling was Langenau *et al.* [122] in order to bypass the early lethality of the Tg(rag2:GFP-Myc) T-ALL transgenic model. A Tg(rag2:loxP-DsRed-loxP-GFP-Myc) target transgenic line that expresses the red fluorescent protein DsRed under the control of the *rag2* promoter was generated. When Cre recombinase was introduced into the transgenic embryos via mRNA microinjection, the DsRed DNA part was excised allowing GFP-Myc expression and leukemia development. However, only partial Cre-mediated recombination occurred and only about 6% of the Cre-injected zebrafish developed

T-ALL, whereas the incidence of T-ALL is 100% in the Tg(rag2:GFP-Myc) animals [121,122]. Using the same approach, the group of Jae-Hak Park (Seoul National University, Seoul, Korea) generated a Tg(nes:loxP-mCherry-loxP-GFP-KRAS<sup>G12V</sup>) transgenic line. The injection of Cre mRNAs places KRAS<sup>G12V</sup> under the control of the *nestin* promoter leading to extensive apoptosis of neural progenitor cells followed by severe edema of the brain and early death of the injected transgenic zebrafish. However, overexpression of KRAS<sup>G12V</sup> was not able to induce brain tumor development [175]. More recently, effector lines stably expressing Cre were generated to achieve tissue-specific expression of oncogenes as a rational alternative to Cre mRNA injection. Indeed, when an effector line expressing Cre under the control of the *ela3l* (elastase 3l) promoter was mated to a Tg(ubb:loxP-mCherry-loxP-GFP-KRAS<sup>G12V</sup>) transgenic line, mCherry was excised from the target transgene and GFP-KRAS<sup>G12V</sup> expressed in the endocrine pancreas. As a result, 40% of the double-transgenic zebrafish develop pancreatic endocrine tumors by the age of 12 months [176].

### Heat Shock Inducible Promoters

Genetic tools allowing temporal control of oncogene expression are essential in cancer modelling. In this regard, promoters of heat shock proteins can successfully be time-controlled in transgenic zebrafish [177,178]. In zebrafish, the *hsp70* (heat shock protein 70) promoter is widely used for raising ectopic and controlled expression of transgenes. This promoter consists in a 1.5 kbp genomic fragment containing contiguous 5-bp DNA consensus sequences known as heat shock elements (HSEs). These elements are binding sites for the heat shock transcription factors Hsf, which are repressed through the association with the Hsp90 protein complex under unstressed conditions. As protein unfolding increases with temperature, the chaperone protein Hsp90 is recruited from the Hsf-Hsp90 protein complexes, Hsf is released and free to bind to its DNA recognition sites to activate transcription [179]. The heat shock treatment of live zebrafish is usually carried out by soaking the embryos for 30 min in water at 38 °C.

The *hsp70* promoter has been used to generate zebrafish transgenic lines expressing various oncogenes in a controlled fashion (Figure 5d). For instance, the Marina Mione lab (Centre for Integrative Biology, University of Trento, Trento, Italy) generated a Tg(*hsp70*:GFP-HRAS<sup>G12V</sup>) transgenic line expressing the *HRAS<sup>G12V</sup>* oncogene under the control of the *hsp70* heat-inducible promoter [180]. Twenty-four hours old transgenic embryos subjected to heat shock developed traits of a Costello syndrome-like phenotype associated with an increase in cellular hyperproliferation and senescence. However, no obvious cancer formation was described in this model. In contrast, expression of the oncogenic fusion proteins AML1-ETO (RUNX1-RUNX1T1) or BCR-ABL1 under the *hsp70* promoter is responsible for the development of myeloid leukemia-like lesions after the heat shock of the transgenic embryos [181,182], whereas expression of the human EWS-FLI1 (EWSR1-FLI1) fusion oncoprotein under the *hsp70* promoter induces tumors with histology strongly resembling that of human Ewing's sarcoma in heat shock-treated zebrafish [183]. Li-Jing Shen, Fang-Yuan Chen and colleagues (Shanghai

Jiao Tong University School of Medicine, Shanghai, China) used a bidirectional promoter containing 8 HSEs to drive the expression of the GFP marker on one side and of oncogenes, either MYCN or the fusion product RUNX1-

EVI1 (RUNX1-MECOM), on the other DNA strand [184,185]. In both cases, heat shock treatments of the transgenic embryos led to an altered hematopoiesis and an increase in AML blasts, serving as models of myeloid malignancies. One early application of heat shock promoters to zebrafish cancer models was their use to control Cre expression. The cross of a Tg(hsp70:Cre) transgenic line with the Tg(rag2:loxP-DsRedloxP-GFP-Myc) T-ALL model leads to double-transgenics that develop leukemia in 81% of heat shocked larvae [186]. This indicates that expression of Cre from a stable controlled transgene is far more efficient than from Cre mRNA injection in embryos. However, due to leakiness of the *hsp70* promoter about 13% of the non-heat shocked double transgenics also develop leukemia. Nevertheless, Tg(hsp70:Cre) transgenic lines have been used to activate oncogene expression such as KRAS<sup>G12D</sup> or NUP98-HOXA9, in double-transgenic zebrafish [187,188].

Thus, despite the leakiness of the *hsp70* promoter and the limitation of the heat shock inducible approach to embryos, this transgenic strategy based on heat shock promoters successfully achieve temporal control of oncogene expression applied to cancer modelling in zebrafish.

### *Tetracycline Regulated Expression*

The tetracycline (Tet) system comes from *Escherichia coli* where it controls the expression of genes involved in tetracycline resistance. This system relies on the tetracycline repressor (tetR) that reacts to tetracycline or to its more stable derivative doxycycline (Dox). Through the fusion of tetR to the transactivation domain of the transcription factor VP16, a Dox-controlled transactivator (tTA) has been generated [189]. In this system named Tet-OFF, tTA binds to tetracycline operator (tetO) sequences and activates transcription from a minimal promoter in absence of Dox, whereas Dox binding to tTA inhibits tTA-mediated transcriptional activation [190]. Activation of the Tet-OFF transcription then relies on the removal of Dox from organisms making this approach somehow inconvenient. From tTA, Manfred Gossen, Herman Bujard and colleagues (Universität Heidelberg, Heidelberg, Germany) developed a Dox-dependent transcriptional activator named rtTA [191]. In this Tet-ON binary system, the rtTA transactivator is inactive in absence of Dox, while upon Dox addition, rtTA binds to TetO-containing promoters and activate transcription (Figure 5e). The Tet-ON system has been applied to temporally regulate gene expression in a number of organisms, including *Drosophila* [192], mice [193] or zebrafish [194]. Conditional expression based on the Tet-ON system in transgenic zebrafish appears sometime leaky and whereas the induction by Dox is rapid and strong, inactivation of the transgene upon Dox removal is very slow. However, the rtTA system has been successfully used in transgenic zebrafish to model brain [195] and liver cancers [196]. In particular, Tet-ON-based strategies have been largely applied to hepatocellular carcinoma modelling by the group of Zhiyuan Gong (National University of Singapore, Singapore) [196–201]. In these studies, the oncogenes *xmrk*, *Myc* or *kras*<sup>G12V</sup> are expressed under the control of a tetO-containing minimal promoter, while rtTA expression is driven by the liver-specific promoter *fabp10a* [196,197,199]. In this manner, the spatial expression of the oncogene is achieved by the tissue-specific expression of rtTA and the temporal control relies on Dox supply. For instance, Tet-ON conditional expression of *xmrk*, a *Xiphophorus* hyperactive version of epidermal growth factor receptor, induces already after three weeks of Dox treatment, hepatocellular carcinoma leading to diminished growth and increased lethality in both juvenile and adult transgenic zebrafish. Moreover, induced liver tumors regressed rapidly upon inducer withdrawal, leading to complete rescue in about four weeks [196]. In another study, in order to investigate the relationship between Kras and RhoA, the Zhiyuan Gong lab used the Tet-ON strategy to express concomitantly in the same transgenics, the *kras*<sup>G12V</sup> oncogene on one side, and constitutively active (RhoA<sup>G14V</sup>) or dominant negative (RhoA<sup>T19N</sup>) versions of *rhoA* on the other side [199]. This approach not only shows that activation of RhoA inhibits the oncogenic effect of Kras<sup>G12V</sup> but also demonstrates that the Tet-ON system could be applied to decipher signaling pathways involved in tumorigenesis.

### *Ligand-Binding Domain Fusion Proteins*

Steroid hormone receptors are modular transcription factors organized into structurally and functionally defined domains [202,203]. Fusing the ligand-binding domain (LBD) onto other transcription factors or tyrosine kinases confers hormone-dependent activities on these proteins [204–207]. To activate the LBDs from estrogen receptor (ER), progesterone receptor (PR), glucocorticoid receptor (GR) and androgen receptor (AR), the most common hormone analogues used are 4-hydroxy-tamoxifen (4-OHT), mifepristone (RU486), dexamethasone and

mibolone, respectively (Figure 5f). The Thomas Look lab (Children's Hospital, Harvard Medical School, Boston, MA, USA) used this conditional approach to control leukemia development in transgenic zebrafish [208]. A stable transgenic line Tg(rag2:MYC-ER<sup>T2</sup>), in which the zebrafish *rag2* promoter drives expression of human MYC oncogene fused to a modified LBD from the estrogen receptor (ER<sup>T2</sup>) that is posttranslationally induced by addition of 4-OHT but not by endogenous estrogens [209], has been generated. In this model, 4-OHT treatment induces MYC activation and disease development within 5 weeks of treatment, and withdrawal of 4-OHT results in T-ALL apoptosis and tumor regression. In contrast, untreated siblings do not develop blood abnormalities suggesting that the activity of the MYC-ER<sup>T2</sup> fusion protein is tightly regulated without observable leakiness. Hence, using a tissue-specific promoter and a LBD fusion protein, two levels of regulation, spatial and temporal are combined in a single transgene to achieve oncogene activity control. In another work, Joji Nakayama, Zhiyuan Gong and colleagues described a transgenic zebrafish line expressing a Twist1a-ER<sup>T2</sup> fusion driven by the liver-specific promoter *fapb10* together with the *xmrk* oncogene under the control of the Tet-ON system as a model for metastatic dissemination of liver cancer cells induced by both 4-OHT and Dox [210]. This possibility to activate simultaneously or sequentially different cancer genes confers a huge flexibility to the transgenic zebrafish cancer model.

LexPR results from the fusion of the DNA-binding domain of bacterial repressor protein LexA, a truncated LBD of the progesterone receptor and the activation domain from RELA (NF- $\alpha$ B/p65). The activation domain and the LexA DNA-binding motif make the LexPR a transcriptional activator recognizing LexO sequences and activating transcription from LexO-containing minimal promoters (LexOp). Furthermore, the use of progesterone receptor LBD containing a 42 amino acid C-terminal deletion renders the LBD unable to bind progesterone or other endogenous hormones but still able to bind mifepristone (RU486), a well characterized anti-progesterone compound [211]. When fused to other proteins, this truncated progesterone-binding domain confers RU486 responsiveness [212,213]. Then, the LexPR conditional system is a binary transcriptional system activating transcription from LexOp promoters in a mifepristone dependent fashion. Alexander Emelyanov and Serguei Parinov (National University of Singapore, Singapore) were the first to apply this system in transgenic zebrafish and to demonstrate its potential for oncogene activation [214]. One weakness of the system, also observed with other binary gene expression systems, is the variability in expression strength between larvae of the same clutch and between cells within the same larva. However, the LexPR system has been used in transgenic zebrafish to generate hepatocellular carcinoma models through the expression of LexPR under the control of the liver-specific promoter *fapb10a* and LexOp driving the expression of *kras*<sup>G12V</sup> or *tgfb1a* [215–217]. Similarly, when the LexOp Nras<sup>Q61K</sup>/Kras<sup>G12V</sup> target oncogene is activated by LexPR under the control of the melanocyte-specific *mitfa* promoter or under the control of the intestine-specific *ifabp* (*fapb2*) promoter, double-transgenic zebrafish lines develop, in presence of mifepristone, melanoma or intestinal tumors, respectively [218,219]. In addition, the reversibility of the LexPR system allows the investigation of oncogene addiction of tumors in the transgenics. When Tg(*fapb10*:LexPR; LexOp:*kras*<sup>G12V</sup>) transgenics with apparent liver tumors, are transferred to mifepristone-free water, a regression of mifepristone-induced tumors is observed indicating that tumor maintenance requires continual *kras*<sup>G12V</sup> expression [215]. Finally, the LexPR can be combined with other expression systems to develop more sophisticated zebrafish cancer models. For instance, the Tg(*fapb10a*:LexPR; LexOp:Cre; *fapb10a*:loxP-mCherry-loxP-GFP-*kras*<sup>G12V</sup>) transgenic line associate the LexPR system to Cre-mediated recombination [220]. These triple-transgenic zebrafish contain three different constructs obtained via crossbreeding of the driver/Cre-effector line with a *fapb10a*:loxP-mCherry-loxP-GFP-*kras*<sup>G12V</sup> line. By applying RU486 to triple-transgenics, LexPR activator produced from the driver exclusively activates Cre expression in the liver, which subsequently removes DNA sequences coding for mCherry.

After Cre-mediated recombination, the liver-specific expression of *kras*<sup>G12V</sup> is constitutively activated and the triple-transgenic zebrafish develop hepatocellular carcinoma.

Owing to the success of inducible Cre/loxP applications in mouse models [221,222], the chimeric Cre-LBD system has been also transferred into zebrafish to achieve spatial and temporal control of transgene expression [223] (Figure 5f). Additionally, the large number of CreER<sup>T2</sup> expressing lines [224] offers a strong potential in terms of generation of zebrafish cancer models. However, only few applications of the conditional CreER<sup>T2</sup> technology have been reported so far in the field of cancer modelling [225,226]. Kalasekar *et al.* [225] generated a model of hepatocellular carcinoma based on the CreER<sup>T2</sup>-mediated conditional expression of an activated version of  $\beta$ -



catenin (Xla.Cttnb1<sup>S33A, S37A, T41A, S45A</sup>, and hereafter-named Cttnb1<sup>ACT</sup>). In this system, a Tg(fabp10a:CreER<sup>T2</sup>) line expressing Cre recombinase fused to the modified LBD of the estrogen receptor downstream of the hepatocyte-specific *fabp10a* promoter is used as a driver line, whereas a Tg(fabp10a:loxP-BFP-loxP-Cttnb1<sup>ACT</sup>) transgenic line serves as a switch target line. Treatment of double-transgenic larvae with 4-OHT from 3 dpf to 6 dpf results in expression of activated  $\beta$ -catenin in most hepatocyte leading to hepatocellular carcinoma development in about 30% of the adult zebrafish at 3 months of age. However, untreated double-transgenics also develop hepatocellular carcinoma at a similar frequency. This observation is explained by a leakiness of the CreER<sup>T2</sup> system. Indeed, using a fluorescent color switch Tg(ubi:loxP-GFP-loxP-mCherry) reporter line crossed with the CreER<sup>T2</sup> driver line, the authors showed that at 6 dpf, 48% of the hepatocytes expressed mCherry upon 4-OHT treatment, but 6% of the hepatocytes are also mCherry<sup>+</sup> in absence of treatment. Thus, a subset of larval hepatocytes expressing the activated  $\beta$ -catenin is sufficient to drive liver tumorigenesis. Furthermore, hepatocellular carcinoma penetrance was significantly lower in the CreER<sup>T2</sup>-based model

(about 30%) than in the constitutive Tg(fabp10a:Cttnb1<sup>ACT</sup>) model in which activated  $\beta$ -catenin is directly under the control of the *fabp10a* promoter (about 85%) [147]. One possible explanation for the difference in penetrance is that activated  $\beta$ -catenin is expressed slightly earlier in the constitutive Tg(fabp10a:Cttnb1<sup>ACT</sup>) model than in the CreER<sup>T2</sup>-based model due to the time required to transcribe and translate CreER<sup>T2</sup> and to excise the BFP/STOP cassette preventing the expression of the activated  $\beta$ -catenin [225]. Then, the potential leakiness and the possible reduced efficacy of the CreER<sup>T2</sup> system imply the requirement for further improvements before its wider application to zebrafish cancer modelling. Finally, it worth noting that in contrast to the LexPR strategy, 4-OHT-mediated activation of the CreER<sup>T2</sup> systems induces irreversible oncogene activation. This might not be a drawback in terms of cancer modelling since mutations are stably acquired during tumorigenesis [227,228], and in certain application such as drug screening, the maintenance of the mutation status in absence of the inducer could be an advantage.

### Optogenetics

Optogenetics broadly refers to biological techniques involving light-gated proteins to control cellular behavior. It provides a precise tool to modulate spatially and temporally cellular activities [229–231]. Although optogenetics relies generally on light-sensitive protein domains, an optical control could also be applied to uncage the inducer of a protein activity. Zhiping Feng, Shimon Weiss, David Bensimon and co-workers (University of California at Los Angeles, Los Angeles, CA, USA; Ecole Normale Supérieure, Paris, France) reported an original approach to induce oncogene expression and tumor formation in transgenic zebrafish using optical control [232]. The strategy utilizes caged-cyclofen, a synthetically modified estrogen receptor inducer, which allows stringent light-dependent activation of proteins fused to the modified estrogen receptor LDB, ER<sup>T2</sup>. In presence of caged-cyclofen, these ER<sup>T2</sup> fusion proteins are sequestered in the cytoplasm through their interaction with chaperones. Optical-induced uncaging of cyclofen releases the fusion proteins from their chaperone complexes, allowing them to translocate into the nucleus and activate transcription (e.g., Gal4) or induce recombination (e.g., Cre). Based on this principle, two light-inducible strategies for transient or constitutive *KRAS*<sup>G12V</sup> expression in transgenic zebrafish were implemented. The transient system combines the effector transgenic construct ubi:Gal4-ER<sup>T2</sup> where the Gal4-ER<sup>T2</sup> fusion is expressed under the control of the *ubiquitin* promoter, and the target transgene UAS:*KRAS*<sup>G12V</sup>-T2A-CFP. Photo-activation of caged-cyclofen allows Gal4-ER<sup>T2</sup> to induce *KRAS*<sup>G12V</sup> expression. Once uncaged cyclofen diffuses, the expression of *KRAS*<sup>G12V</sup> is turned off and its mRNA and protein products are slowly degraded, resulting in a transient expression of the oncogene. Furthermore, the co-expression of the CFP fluorescent protein reports on the activity of the target transgene. The authors showed that transient activation of *KRAS*<sup>G12V</sup> through photo-activation of caged-cyclofen at 1 dpf does not result in tumor formation. In contrast, periodic *KRAS*<sup>G12V</sup> expression achieved by 1 day exposure to cyclofen every 5 days during 2 months, leads to tumor development in about 3% of the transgenic zebrafish within 12 months [232]. Constitutive *KRAS*<sup>G12V</sup> expression is obtained through the injection of the ubi:loxP-EosFP-loxP-*KRAS*<sup>G12V</sup>-T2A-mTFP target construct into transgenic Tg(ubi:CreER<sup>T2</sup>) zebrafish, subsequently exposed to caged-cyclofen prior photo-activation at 1 dpf. About 10% of treated embryos will then develop tumors before 1 year of age. Feng *et al.* [232] were also able to induce *KRAS*<sup>G12V</sup> expression at the single-cell level using two-photon uncaging, thus highlighting the potential of the approach although none of the fish developed a tumor in this condition.

### *Mosaic Somatic Transgenesis*

The generation of stable transgenic zebrafish models of cancer is tedious and time-consuming, and even more difficult when the studies require the generation of multiple transgenic lines to activate oncogene expression or to investigate the cooperation of various genetic alterations. Mosaic transient (or somatic) transgenesis appears then as a powerful alternative strategy for rapid functional studies (Figure 4b) [233]. One or several transgenes are injected into one-cell stage wild-type or transgenic embryos, and randomly integrates into the genome of the F0 mosaic zebrafish [234]. Co-injection of different transgenes in the embryo leads to their co-integration into given cells allowing the co-expression of the transgenes in these cells [235]. Moreover, given that cancer generally results from somatic sequential mutations rather than multiple germline genetic alterations, the mosaic transient transgenesis approach has also a rational basis. The Leonard Zon Lab (Children's Hospital, Boston, MA, USA) has largely demonstrated that the simplicity of the technique allows rapid and robust melanoma modelling by combining mosaic somatic expression and/or inactivation of different oncogenes and tumor-suppressor genes in zebrafish [83,236,237]. This approach has also been applied to demonstrate the oncogenic role of long non-coding (lnc) RNAs such as *THOR* [238].

Oncogene expression through transgenesis via the injection of transgenes into one-cell stage embryos implies that these oncogenes will be expressed from early embryogenesis, with potentially deleterious effects during development. To bypass this possible limitation, Richard White and colleagues (Memorial Sloan Kettering Cancer Center, New York City, NY, USA) implemented a method called TEAZ (Transgene Electroporation in Adult Zebrafish) to deliver transgenes driving oncogene as well as CRISPR/Cas9 components expression, directly into somatic tissues of adult zebrafish using electroporation [239]. One interesting outcome of this approach is that it is possible to track over the time the spreading of the tumor from its origin, at the electroporation point.

Then, the continuous progress in the improvement of conditional transgenesis approaches in zebrafish has generated an immense toolbox for controlling oncogene and tumor-suppressor gene activities in specific cells and a given time. This set of tools offers a variety of modelling possibilities in the field of cancer research.

### Conclusions

Over the past two decades, the progresses accomplished in the implementation of sophisticated genetic and reverse genetic tools allowed the emergence of a remarkable diversity of genetically engineered zebrafish cancer models. The precision of the reverse genetic toolbox enables the generation of zebrafish lines carrying genetic modifications that are identical to the alterations found in patients. Transcriptomic and oncogenomic comparisons between human and zebrafish tumors reveal common molecular signatures and potentially lead to the discovery of novel genes involved in tumorigenesis [29,81,90,129,135,240,241]. The possibility to associate several genetic alterations in zebrafish allows combinatorial genetic modelling to recapitulate the genetic alterations found in humans and to decipher the cross-talk between different pathways involved in tumor initiation or progression [83,149,199]. Thus, the zebrafish contributes novel insights in tumor biology and provides suited cancer models that can ultimately being used to identify anti-cancer drugs [242,243].

**Author Contributions:** P.-O.A. wrote the manuscript. L.R., P.V., P.-O.A. collected data from the literature; L.R., P.V. and X.L.B. revised the manuscript. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work was supported by the CNRS, Inserm, the University of Lille, the Ligue Contre le Cancer – Comité du Pas-de-Calais and the GIP Cancéropôle Nord-Ouest.

**Acknowledgments:** We are grateful to Lizzy Griffiths for the drawings of zebrafish, freely available online at <http://zebrafishart.blogspot.com/> (assessed on 13 April 2020).

**Conflicts of Interest:** The authors declare no conflict of interest.

### Appendix A

**Table A1.** Transgenic zebrafish lines developing cancers.

## PARTIE II : LE POISSON ZEBRE EN CANCEROLOGIE

Transgenic Lines	Cancer Types	References
Tg(rag2: GFP-Myc)		[121]
Tg(rag2:loxP-DsRed-loxP-GFP-Myc); mRNA-Cre		[122]
Tg(rag2:loxP-DsRed-loxP-GFP-Myc; hsp70:Cre)		[186]
Tg(rag2:MYC-ER <sup>T2</sup> )	T-ALL	[208]
Tg(rag2:NOTCH1 <sup>ICN</sup> -GFP)		[244]
Tg(rag2:GFP-Myc; rag2:GIMAP5; rag2:GIMAP7)		[245]
Tg(rag2:GFP-bcl2) Tg(rag2:GFP-Myc; rag2:GFP-bcl2)	T-LBL	[246]
Tg(act/xEF1:TEL-AML1)/[ETV6-RUNX1]		[157]
Tg(rag2: GFP-Myc)	B-ALL	[123]
Tg(rag2:MYC)		[124,125]
Tg(spi1:MYST3-NCOA2-GFP) [MOZ-TIF2]		[131]
Tg(hsp70:AML1-ETO)		[181]
Tg(MYCN:HSE:GFP)		[184]
Tg(RUNX1-EVI1:HSE:GFP)		[185]
Tg(spi1:tel-jak2a)/Tg(CMV:tel-jak2a)		[247,248]
Tg(spi1:loxP-GFP-loxP-NUP98-HOXA9; hsp70:Cre)	AML	[188]
Tg(fli1:Gal4FF; UAS:GFP-HRAS <sup>G12V</sup> )		[167]
Tg(hsp70:p210BCR/ABL1)		[182]
Tg(spi1:FLT3 <sup>ITD</sup> -2A-GFP)		[132]
Tg(spi1:SOX4-GFP)		[133]
Tg(mitfa:BRAF <sup>V600E</sup> ); tp53 <sup>-/-</sup>		[134]
Tg(mitfa:BRAFV600E); mitf <sup>avc7/vc7</sup>		[249]
Tg(mitfa:GFP-NRAS <sup>Q61K</sup> ); tp53 <sup>-/-</sup>		[135]
Tg(mitfa:mitfa; mitfa:NRAS <sup>Q61K</sup> ); Casper		[237]
Tg(mitfa:HRAS <sup>G12V</sup> ; mitfa:GFP)	Melanoma	[136]
Tg(UAS:GFP-HRAS <sup>G12V</sup> ; mitfa/kita:Gal4-VP16)		[166,250]
Tg(mitfa:GNAQ <sup>Q209P</sup> ); tp53 <sup>-/-</sup>		[137]
Tg(kita/mitfa:LexPR-Cerulean; Kras <sup>G12V</sup> /Hras <sup>G12V</sup> /Nras <sup>Q61K</sup> ) lexOp:RFP-		[218]
Tg(mitfa:BRAF <sup>V600E</sup> ); tp53 <sup>-/-</sup> ; prdm1a <sup>+/-</sup>		[251]

## PARTIE II : LE POISSON ZEBRE EN CANCEROLOGIE

Tg(rag2:KRAS <sup>G12D</sup> )		Rhabdomyosarcoma	[129]
Tg(cdh15:KRAS <sup>G12D</sup> ), Tg(mylz2:KRAS <sup>G12D</sup> )			[138]
<b>Table A1. Cont.</b>			
Transgenic Lines	Cancer Types	References	
Tg(ptf1a:GFP-KRAS <sup>G12V</sup> )			[139]
Tg(ptf1a:GAL4-VP16; Tg(myod:MYCN)	UAS:KRAS <sup>mut</sup>	Pancreas tumors	[170,172] [140]
Tg(ptf1a:CreER <sup>T2</sup> ; ubb:loxP-STOP-loxP-Gal4; UAS:KRAS <sup>G12V</sup> )			[226]
Tg(ela3I:Cre; ubb:loxP-STOP-loxP-KRAS <sup>G12D</sup> )			[176]
Tg(act:loxP-EGFP-loxP-KRAS <sup>G12D</sup> ; hsp70:Cre)	Various		[187]
Tg(act: EWS-FLI1)/Tg(hsp70: EWS-FLI1) (tp53 <sup>-/-</sup> )	Ewing sarcoma		[183]
Tg(rag2:myr-Akt2) Tg(krt4:myr-AKT1)	Liposarcoma		[130] [252]
Tg(fabp10a:GFP-kras <sup>G12V</sup> )			[141]
Tg(lexA:GFP-kras <sup>G12V</sup> ; fabp10a:LexPR)			[215]
Tg(TetO:Myc; fabp10a:rtTA)			[197,198]
Tg(TetO:xmrk; fabp10a:rtTA)			[196,198]
Tg(TetO:xmrk; fabp10a:rtTA; fabp10a:mCherry-T2A-twist1a-ER <sup>T2</sup> )			[210]
Tg(fabp10a:HBx-RFP); tp53 <sup>-/-</sup> /Tg(fabp10a:src); tp53 <sup>-/-</sup>			[142,201]
Tg(fabp10a:UHRF1-GFP)			[143]
Tg(fabp10a:rtTA2s-M2; TRE2:EGFP-kras <sup>G12V</sup> )			[199–201]
Tg(fabp10a:rtTA2s-M2; TRE2:EGFP-kras <sup>G12V</sup> ; TRE2:EGFP-rhoA <sup>T19N</sup> )	Liver tumors		[199]
Tg(fabp10a:nras <sup>Q61K</sup> )			[144]
Tg(fabp10a:RPIA)			[145]
Tg(fabp10a:LexPR; LexOp:Cre; fabp10a:loxP-mCherry-loxP-EGFP-kras <sup>G12V</sup> )			[220]
Tg(fabp10a:LexPR-T2a-mcherry; LexOp:tgfb1a)			[216,217]
Tg(fabp10a: AURKA <sup>V352I</sup> ; myl7:EGFP)			[146]
Tg(fabp10a: Xla.Cttnb1 <sup>S33A, S37A, T41A, S45A</sup> )			[147]
Tg(fabp10a:CreER <sup>T2</sup> ; fabp10a:loxP-BFP-loxP-Xla.Cttnb1 <sup>S33A, S37A, T41A, S45A</sup> )			[225]
Tg(fabp10a:edn1)			[148]
Tg(ifabp:LexPR; LexOp:EGFP-kras <sup>G12V</sup> )	Intestinal tumors		[219]



## PARTIE II : LE POISSON ZEBRE EN CANCEROLOGIE

Tg(UAS:GFP-HRAS <sup>G12V</sup> ; twhh/mü4465:Gal4)	Chordoma	[169]
Tg(UAS:myr-AKT1; ptf1a:Gal4-VP16)		[171]
Tg(UAS:GFP-RAC1 <sup>G12V</sup> ; ptf1a:Gal4-VP16)		[171]
Tg(UAS:smoa1-GFP; krt5:Gal4-VP16)		[171]
Tg(UAS:mCherry-KRAS <sup>G12V</sup> ; krt5/gfap:Gal4-VP16)		[253]
Tg(TRE:mCherry-KRAS <sup>G12V</sup> ; krt5/gfap:rtTa)		[195]
Tg(UAS:GFP-HRAS <sup>G12V</sup> ; zic4:Gal4-VP16)		[195]
Tg(dbh:GFP-MYCN), Tg(dbh:ALK <sup>F1174L</sup> )		[195]
		[168]
	Brain tumors	[149]
Tg(dbh:EGFP-MYCN; dbh:LMO1)		[254]
Tg(dbh:EGFP-MYCN); nf1a <sup>-/-</sup>		[155]
Tg(sox10:mCherry-NRAS/NRAS <sup>Q61R</sup> /NRAS <sup>S17N</sup> ); p53 <sup>-/-</sup>		[150]
Tg(dbh:MYC)/Tg(dbh:MYCN)		[151]
Tg(dbh:GFP-MYCN; dbh:LIN28B)		[255]
Tg(actb2:KIT <sup>D816V</sup> -GFP)	Mastocytosis	[256]
Tg(pomc:pttg; POMC:GFP)	Pituitary adenoma	[257]
Tg(tg:BRAF <sup>V600E</sup> ; tg:TdTomato)	Papillary thyroid carcinoma	[90]
Tg(flck:Tag)/Tg(flck:scl)/Tg(flck:scl; flck:lmo1)	Testicular germ cell tumors	[152]
Tg(ubi:loxP-EosFP-loxP-KRAS <sup>G12V</sup> -T2A-mTFP; ubi:CreER <sup>T2</sup> ) Tg(UAS:KRAS <sup>G12V</sup> -T2A-CFP; ubi:Gal4-ER <sup>T2</sup> )	Abdominal tumors	[232]

### References

1. Plehn, M. Ueber geschwülste bei kaltblütern. *Z. Krebs.* **1906**, *4*, 525–564. [[CrossRef](#)]
2. Kosswig, C. Über bastarde der teleostier platypoecilus und xiphophorus. *Z. Indukt. Abstamm. -Vererb.* **1927**, *44*, 253.
3. Kosswig, C. Über kreuzungen zwischen den teleostiern xiphophorus helleri und platypoecilus maculatus. *Z. Indukt. Abstamm. -Vererb.* **1928**, *47*, 150–158.
4. Haussler, G. Über melanobildung bei bastarden von xiphophorus helleri und platypoecilus maculatus var. Rubra. *Klin. Wochenschr.* **1928**, *7*, 1561–1562. [[CrossRef](#)]
5. Gordon, M. Hereditary basis of melanosis in hybrid fishes. *Am. J. Cancer* **1931**, *15*, 1495–1523.
6. Walter, R.B.; Kazianis, S. Xiphophorus interspecies hybrids as genetic models of induced neoplasia. *ILAR J.* **2001**, *42*, 299–321. [[CrossRef](#)]
7. Patton, E.E.; Mitchell, D.L.; Nairn, R.S. Genetic and environmental melanoma models in fish. *Pigment Cell Melanoma Res.* **2010**, *23*, 314–337. [[CrossRef](#)]
8. Harshbarger, J.C.; Slatick, M.S. Lesser known aquarium fish tumor models. *Mar. Biotechnol.* **2001**, *3*, S115–S129. [[CrossRef](#)] [[PubMed](#)]
9. Engeszer, R.E.; Patterson, L.B.; Rao, A.A.; Parichy, D.M. Zebrafish in the wild: A review of natural history and new notes from the field. *Zebrafish* **2007**, *4*, 21–40. [[CrossRef](#)] [[PubMed](#)]

10. Sundin, J.; Morgan, R.; Finnøen, M.H.; Dey, A.; Sarkar, K.; Jutfelt, F. On the observation of wild zebrafish (*Danio rerio*) in India. *Zebrafish* **2019**, *16*, 546–553. [[CrossRef](#)] [[PubMed](#)]
11. Laale, H.W. The biology and use of zebrafish, *Brachydanio rerio* in fisheries research. A literature review. *J. Fish. Biol.* **1977**, *10*, 121–173. [[CrossRef](#)]
12. Grunwald, D.J.; Streisinger, G. Induction of recessive lethal and specific locus mutations in the zebrafish with ethyl nitrosourea. *Genet. Res.* **1992**, *59*, 103–116. [[CrossRef](#)] [[PubMed](#)]
13. Howe, K.; Clark, M.D.; Torroja, C.F.; Torrance, J.; Berthelot, C.; Muffato, M.; Collins, J.E.; Humphray, S.; McLaren, K.; Matthews, L.; *et al.* The zebrafish reference genome sequence and its relationship to the human genome. *Nature* **2013**, *496*, 498–503. [[CrossRef](#)] [[PubMed](#)]
14. Rennekamp, A.J.; Peterson, R.T. 15 years of zebrafish chemical screening. *Curr. Opin. Chem. Biol.* **2015**, *24*, 58–70. [[CrossRef](#)]
15. Dang, M.; Fogley, R.; Zon, L.I. Identifying novel cancer therapies using chemical genetics and zebrafish. *Adv. Exp. Med. Biol.* **2016**, *916*, 103–124. [[CrossRef](#)]
16. White, R.M.; Sessa, A.; Burke, C.; Bowman, T.; LeBlanc, J.; Ceol, C.; Bourque, C.; Dovey, M.; Goessling, W.; Burns, C.E.; *et al.* Transparent adult zebrafish as a tool for in vivo transplantation analysis. *Cell Stem Cell* **2008**, *2*, 183–189. [[CrossRef](#)]
17. Antinucci, P.; Hindges, R. A crystal-clear zebrafish for in vivo imaging. *Sci. Rep.* **2016**, *6*, 29490. [[CrossRef](#)]
18. Lawson, N.D.; Weinstein, B.M. In vivo imaging of embryonic vascular development using transgenic zebrafish. *Dev. Biol.* **2002**, *248*, 307–318. [[CrossRef](#)]
19. Choi, J.; Dong, L.; Ahn, J.; Dao, D.; Hammerschmidt, M.; Chen, J.N. FoxH1 negatively modulates flk1 gene expression and vascular formation in zebrafish. *Dev. Biol.* **2007**, *304*, 735–744. [[CrossRef](#)]
20. Traver, D.; Paw, B.H.; Poss, K.D.; Penberthy, W.T.; Lin, S.; Zon, L.I. Transplantation and in vivo imaging of multilineage engraftment in zebrafish bloodless mutants. *Nat. Immunol.* **2003**, *4*, 1238–1246. [[CrossRef](#)]
21. Amatruda, J.F.; Shepard, J.L.; Stern, H.M.; Zon, L.I. Zebrafish as a cancer model system. *Cancer Cell* **2002**, *1*, 229–231. [[CrossRef](#)]
22. Kent, M.L.; Spitsbergen, J.M.; Matthews, J.M.; Fournie, J.W.; Westerfield, M. Diseases of Zebrafish in Research Facilities. Zebrafish International Resource Center. 2016. Available online: [https://zebrafish.org/wiki/health/disease\\_manual/start](https://zebrafish.org/wiki/health/disease_manual/start) (accessed on 10 April 2020).
23. Matthews, J.L. Common diseases of laboratory zebrafish. *Methods Cell Biol.* **2004**, *77*, 617–643. [[CrossRef](#)]
24. Smolowitz, R.; Hanley, J.; Richmond, H. A three-year retrospective study of abdominal tumors in zebrafish maintained in an aquatic laboratory animal facility. *Biol. Bull.* **2002**, *203*, 265–266. [[CrossRef](#)] [[PubMed](#)]
25. Stanton, M. Diethylnitrosamine-induced hepatic degeneration and neoplasia in the aquarium fish, *Brachydanio rerio*. *J. Natl. Cancer Inst.* **1965**, *34*, 117–130. [[CrossRef](#)] [[PubMed](#)]
26. Mizgireuv, I.V.; Majorova, I.G.; Gorodinskaya, V.M.; Khudoley, V.V.; Revskoy, S.Y. Carcinogenic effect of N-nitrosodimethylamine on diploid and triploid zebrafish (*Danio rerio*). *Toxicol. Pathol.* **2004**, *32*, 514–518. [[CrossRef](#)] [[PubMed](#)]
27. Mizgireuv, I.V.; Revskoy, S.Y. Transplantable tumor lines generated in clonal zebrafish. *Cancer Res.* **2006**, *66*, 3120–3125. [[CrossRef](#)] [[PubMed](#)]
28. Spitsbergen, J.M.; Tsai, H.-W.; Reddy, A.; Miller, T.; Arbogast, D.; Hendricks, J.D.; Bailey, G.S. Neoplasia in zebrafish (*Danio rerio*) treated with 7,12-dimethylbenz[a]anthracene by two exposure routes at different developmental stages. *Toxicol. Pathol.* **2000**, *28*, 716–725. [[CrossRef](#)]
29. Lam, S.H.; Wu, Y.L.; Vega, V.B.; Miller, L.D.; Spitsbergen, J.; Tong, Y.; Zhan, H.; Govindarajan, K.R.; Lee, S.; Mathavan, S.; *et al.* Conservation of gene expression signatures between zebrafish and human liver tumors and tumor progression. *Nat. Biotechnol.* **2006**, *24*, 73–75. [[CrossRef](#)]
30. Spitsbergen, J.M.; Tsai, H.W.; Reddy, A.; Miller, T.; Arbogast, D.; Hendricks, J.D.; Bailey, G.S. Neoplasia in zebrafish (*Danio rerio*) treated with N-methyl-N'-nitro-N-nitrosoguanidine by three exposure routes at different developmental stages. *Toxicol. Pathol.* **2000**, *28*, 705–715. [[CrossRef](#)]
31. Beckwith, L.G.; Moore, J.L.; Tsao-Wu, G.S.; Harshbarger, J.C.; Cheng, K.C. Ethylnitrosourea induces neoplasia in zebrafish (*Danio rerio*). *Lab. Invest.* **2000**, *80*, 379–385. [[CrossRef](#)]
32. Lam, S.H.; Gong, Z. Modeling liver cancer using zebrafish: A comparative oncogenomics approach. *Cell Cycle* **2006**, *5*, 573–577. [[CrossRef](#)] [[PubMed](#)]

33. Zimmering, S.; Thompson, E. Comparison of rates of sex-linked recessive lethals induced by ethylnitrosourea (ENU) in postmeiotic cells of the male and oogonia of the female *Drosophila*. *Environ. Mutagen.* **1984**, *6*, 617–619. [[CrossRef](#)] [[PubMed](#)]
34. Yoshikawa, I.; Ayaki, T.; Ohshima, K. Comparative studies of dose-response curves for recessive lethal mutations induced by ethylnitrosourea in spermatogonia and in spermatozoa of *Drosophila melanogaster*. *Environ. Mutagen.* **1984**, *6*, 489–496. [[CrossRef](#)] [[PubMed](#)]
35. Russell, W.L.; Kelly, E.M.; Hunsicker, P.R.; Bangham, J.W.; Maddux, S.C.; Phipps, E.L. Specific-locus test shows ethylnitrosourea to be the most potent mutagen in the mouse. *Proc. Natl. Acad. Sci. USA* **1979**, *76*, 5818–5819. [[CrossRef](#)] [[PubMed](#)]
36. Mullins, M.C.; Hammerschmidt, M.; Haffter, P.; Nüsslein-Volhard, C. Large-scale mutagenesis in the zebrafish: In search of genes controlling development in a vertebrate. *Curr. Biol.* **1994**, *4*, 189–202. [[CrossRef](#)]
37. Driever, W.; Solnica-Krezel, L.; Schier, A.F.; Neuhauss, S.C.; Malicki, J.; Stemple, D.L.; Stainier, D.Y.; Zwartkruis, F.; Abdelilah, S.; Rangini, Z.; *et al.* A genetic screen for mutations affecting embryogenesis in zebrafish. *Development* **1996**, *123*, 37–46.
38. Haffter, P.; Granato, M.; Brand, M.; Mullins, M.C.; Hammerschmidt, M.; Kane, D.A.; Odenthal, J.; van Eeden, F.J.; Jiang, Y.J.; Heisenberg, C.P.; *et al.* The identification of genes with unique and essential functions in the development of the zebrafish, *Danio rerio*. *Development* **1996**, *123*, 1–36.
39. Van Rooijen, E.; Giles, R.H.; Voest, E.E.; van Rooijen, C.; Schulte-Merker, S.; van Eeden, F.J. LRRC50, a conserved ciliary protein implicated in polycystic kidney disease. *J. Am. Soc. Nephrol.* **2008**, *19*, 1128–1138. [[CrossRef](#)]
40. Basten, S.G.; Davis, E.E.; Gillis, A.J.; van Rooijen, E.; Stoop, H.; Babala, N.; Logister, I.; Heath, Z.G.; Jonges, T.N.; Katsanis, N.; *et al.* Mutations in LRRC50 predispose zebrafish and humans to seminomas. *PLoS Genet.* **2013**, *9*, e1003384. [[CrossRef](#)]
41. Neumann, J.C.; Chandler, G.L.; Damoulis, V.A.; Fustino, N.J.; Lillard, K.; Looijenga, L.; Margraf, L.; Rakheja, D.; Amatruda, J.F. Mutation in the type IB bone morphogenetic protein receptor Alk6b impairs germ-cell differentiation and causes germ-cell tumors in zebrafish. *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 13153–13158. [[CrossRef](#)]
42. Amsterdam, A.; Burgess, S.; Golling, G.; Chen, W.; Sun, Z.; Townsend, K.; Farrington, S.; Haldi, M.; Hopkins, N. A large-scale insertional mutagenesis screen in zebrafish. *Genes Dev.* **1999**, *13*, 2713–2724. [[CrossRef](#)] [[PubMed](#)]
43. Amsterdam, A.; Sadler, K.C.; Lai, K.; Farrington, S.; Bronson, R.T.; Lees, J.A.; Hopkins, N. Many ribosomal protein genes are cancer genes in zebrafish. *PLoS Biol.* **2004**, *2*, E139. [[CrossRef](#)] [[PubMed](#)]
44. Amsterdam, A.; Lai, K.; Komisarczuk, A.Z.; Becker, T.S.; Bronson, R.T.; Hopkins, N.; Lees, J.A. Zebrafish hagoromo mutants up-regulate fgf8 postembryonically and develop neuroblastoma. *Mol. Cancer Res.* **2009**, *7*, 841–850. [[CrossRef](#)] [[PubMed](#)]
45. Shepard, J.L.; Amatruda, J.F.; Stern, H.M.; Subramanian, A.; Finkelstein, D.; Ziai, J.; Finley, K.R.; Pfaff, K.L.; Hersey, C.; Zhou, Y.; *et al.* A zebrafish bmyb mutation causes genome instability and increased cancer susceptibility. *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 13194–13199. [[CrossRef](#)]
46. Shepard, J.L.; Amatruda, J.F.; Finkelstein, D.; Ziai, J.; Finley, K.R.; Stern, H.M.; Chiang, K.; Hersey, C.; Barut, B.; Freeman, J.L.; *et al.* A mutation in separase causes genome instability and increased susceptibility to epithelial cancer. *Genes Dev.* **2007**, *21*, 55–59. [[CrossRef](#)]
47. Koudijs, M.J.; den Broeder, M.J.; Keijser, A.; Wienholds, E.; Houwing, S.; van Rooijen, E.M.; Geisler, R.; van Eeden, F.J. The zebrafish mutants dre, uki, and lep encode negative regulators of the hedgehog signaling pathway. *PLoS Genet.* **2005**, *1*, e19. [[CrossRef](#)]
48. Niyaz, M.; Khan, M.S.; Mudassar, S. Hedgehog signaling: An achilles' heel in cancer. *Transl. Oncol.* **2019**, *12*, 1334–1344. [[CrossRef](#)]
49. Jeng, K.S.; Chang, C.F.; Lin, S.S. Sonic hedgehog signaling in organogenesis, tumors, and tumor microenvironments. *Int. J. Mol. Sci.* **2020**, *21*, 758. [[CrossRef](#)]
50. Parant, J.M.; George, S.A.; Holden, J.A.; Yost, H.J. Genetic modeling of Li-Fraumeni syndrome in zebrafish. *Dis. Model. Mech.* **2010**, *3*, 45–56. [[CrossRef](#)]
51. Sonawane, M.; Carpio, Y.; Geisler, R.; Schwarz, H.; Maischein, H.M.; Nüsslein-Volhard, C. Zebrafish penner/lethal giant larvae 2 functions in hemidesmosome formation, maintenance of cellular morphology and growth regulation in the developing basal epidermis. *Development* **2005**, *132*, 3255–3265. [[CrossRef](#)]

52. Reischauer, S.; Levesque, M.P.; Nüsslein-Volhard, C.; Sonawane, M. Lgl2 executes its function as a tumor suppressor by regulating ErbB signaling in the zebrafish epidermis. *PLoS Genet.* **2009**, *5*, e1000720. [[CrossRef](#)] [[PubMed](#)]
53. Wallace, K.N.; Dolan, A.C.; Seiler, C.; Smith, E.M.; Yusuff, S.; Chaille-Arnold, L.; Judson, B.; Sierk, R.; Yengo, C.; Sweeney, H.L.; *et al.* Mutation of smooth muscle myosin causes epithelial invasion and cystic expansion of the zebrafish intestine. *Dev. Cell* **2005**, *8*, 717–726. [[CrossRef](#)] [[PubMed](#)]
54. Moore, J.L.; Rush, L.M.; Breneman, C.; Mohideen, M.A.; Cheng, K.C. Zebrafish genomic instability mutants and cancer susceptibility. *Genetics* **2006**, *174*, 585–600. [[CrossRef](#)] [[PubMed](#)]
55. Frazer, J.K.; Meeker, N.D.; Rudner, L.; Bradley, D.F.; Smith, A.C.; Demarest, B.; Joshi, D.; Locke, E.E.; Hutchinson, S.A.; Tripp, S.; *et al.* Heritable T-cell malignancy models established in a zebrafish phenotypic screen. *Leukemia* **2009**, *23*, 1825–1835. [[CrossRef](#)]
56. McCallum, C.M.; Comai, L.; Greene, E.A.; Henikoff, S. Targeted screening for induced mutations. *Nat. Biotechnol.* **2000**, *18*, 455–457. [[CrossRef](#)]
57. McCallum, C.M.; Comai, L.; Greene, E.A.; Henikoff, S. Targeting induced local lesions IN genomes (TILLING) for plant functional genomics. *Plant. Physiol.* **2000**, *123*, 439–442. [[CrossRef](#)]
58. Coghill, E.L.; Hugill, A.; Parkinson, N.; Davison, C.; Glenister, P.; Clements, S.; Hunter, J.; Cox, R.D.; Brown, S.D. A gene-driven approach to the identification of ENU mutants in the mouse. *Nat. Genet.* **2002**, *30*, 255–256. [[CrossRef](#)]
59. Wienholds, E.; Schulte-Merker, S.; Walderich, B.; Plasterk, R.H. Target-selected inactivation of the zebrafish *rag1* gene. *Science* **2002**, *297*, 99–102. [[CrossRef](#)]
60. Berghmans, S.; Murphey, R.D.; Wienholds, E.; Neuberg, D.; Kutok, J.L.; Fletcher, C.D.; Morris, J.P.; Liu, T.X.; Schulte-Merker, S.; Kanki, J.P.; *et al.* tp53 mutant zebrafish develop malignant peripheral nerve sheath tumors. *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 407–412. [[CrossRef](#)]
61. Faucherre, A.; Taylor, G.S.; Overvoorde, J.; Dixon, J.E.; den Hertog, J. Zebrafish pten genes have overlapping and non-redundant functions in tumorigenesis and embryonic development. *Oncogene* **2008**, *27*, 1079–1086. [[CrossRef](#)]
62. Choorapoikayil, S.; Kuiper, R.V.; de Bruin, A.; den Hertog, J. Haploinsufficiency of the genes encoding the tumor suppressor Pten predisposes zebrafish to hemangiosarcoma. *Dis. Model. Mech.* **2012**, *5*, 241–247. [[CrossRef](#)] [[PubMed](#)]
63. Haramis, A.P.; Hurlstone, A.; van der Velden, Y.; Begthel, H.; van den Born, M.; Offerhaus, G.J.; Clevers, H.C. Adenomatous polyposis coli-deficient zebrafish are susceptible to digestive tract neoplasia. *EMBO Rep.* **2006**, *7*, 444–449. [[CrossRef](#)] [[PubMed](#)]
64. Feitsma, H.; Kuiper, R.V.; Korving, J.; Nijman, I.J.; Cuppen, E. Zebrafish with mutations in mismatch repair genes develop neurofibromas and other tumors. *Cancer Res.* **2008**, *68*, 5059–5066. [[CrossRef](#)] [[PubMed](#)]
65. Shive, H.R.; West, R.R.; Embree, L.J.; Azuma, M.; Sood, R.; Liu, P.; Hickstein, D.D. *brca2* in zebrafish ovarian development, spermatogenesis, and tumorigenesis. *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 19350–19355. [[CrossRef](#)] [[PubMed](#)]
66. Hurlstone, A.F.; Haramis, A.P.; Wienholds, E.; Begthel, H.; Korving, J.; Van Eeden, F.; Cuppen, E.; Zivkovic, D.; Plasterk, R.H.; Clevers, H. The Wnt/beta-catenin pathway regulates cardiac valve formation. *Nature* **2003**, *425*, 633–637. [[CrossRef](#)]
67. Wimmer, K.; Rosenbaum, T.; Messiaen, L. Connections between constitutional mismatch repair deficiency syndrome and neurofibromatosis type 1. *Clin. Genet.* **2017**, *91*, 507–519. [[CrossRef](#)]
68. Shive, H.R.; West, R.R.; Embree, L.J.; Golden, C.D.; Hickstein, D.D. BRCA2 and TP53 collaborate in tumorigenesis in zebrafish. *PLoS ONE* **2014**, *9*, e87177. [[CrossRef](#)]
69. Parant, J.M.; Yeh, J.R. Approaches to inactivate genes in zebrafish. *Adv. Exp. Med. Biol.* **2016**, *916*, 61–86. [[CrossRef](#)]
70. Gut, P.; Reischauer, S.; Stainier, D.Y.R.; Arnaout, R. Little fish, big data: Zebrafish as model for cardiovascular and metabolic disease. *Physiol. Rev.* **2017**, *97*, 889–938. [[CrossRef](#)]
71. Gaj, T.; Gersbach, C.A.; Barbas, C.F., 3rd. ZFN, TALEN, and CRISPR/Cas-based methods for genome engineering. *Trends Biotechnol.* **2013**, *31*, 397–405. [[CrossRef](#)]
72. Urnov, F.D.; Rebar, E.J.; Holmes, M.C.; Zhang, H.S.; Gregory, P.D. Genome editing with engineered zinc finger nucleases. *Nat. Rev. Genet.* **2010**, *11*, 636–646. [[CrossRef](#)] [[PubMed](#)]
73. Doyon, Y.; McCammon, J.M.; Miller, J.C.; Faraji, F.; Ngo, C.; Katibah, G.E.; Amora, R.; Hocking, T.D.; Zhang, L.; Rebar, E.J.; *et al.* Heritable targeted gene disruption in zebrafish using designed zinc-finger nucleases. *Nat. Biotechnol.* **2008**, *26*, 702–708. [[CrossRef](#)] [[PubMed](#)]

74. Meng, X.; Noyes, B.; Zhu, L.J.; Lawson, N.D.; Wolfe, S.A. Targeted gene inactivation in zebrafish using engineered zinc-finger nucleases. *Nat. Biotechnol.* **2008**, *26*, 695–701. [[CrossRef](#)] [[PubMed](#)]
75. Ekker, S.C. Zinc finger-based knockout punches for zebrafish genes. *Zebrafish* **2008**, *5*, 121–123. [[CrossRef](#)]
76. Shin, J.; Padmanabhan, A.; de Groh, E.D.; Lee, J.S.; Haidar, S.; Dahlberg, S.; Guo, F.; He, S.; Wolman, M.A.; Granato, M.; *et al.* Zebrafish neurofibromatosis type 1 genes have redundant functions in tumorigenesis and embryonic development. *Dis. Model. Mech.* **2012**, *5*, 881–894. [[CrossRef](#)]
77. Gjini, E.; Mansour, M.R.; Sander, J.D.; Moritz, N.; Nguyen, A.T.; Kesarsing, M.; Gans, E.; He, S.; Chen, S.; Ko, M.; *et al.* A zebrafish model of myelodysplastic syndrome produced through tet2 genomic editing. *Mol. Cell. Biol.* **2015**, *35*, 789–804. [[CrossRef](#)]
78. Solin, S.L.; Shive, H.R.; Woolard, K.D.; Essner, J.J.; McGrail, M. Rapid tumor induction in zebrafish by TALEN-mediated somatic inactivation of the retinoblastoma1 tumor suppressor rb1. *Sci. Rep.* **2015**, *5*, 13745. [[CrossRef](#)]
79. Shim, J.; Choi, J.H.; Park, M.H.; Kim, H.; Kim, J.H.; Kim, S.Y.; Hong, D.; Kim, S.; Lee, J.E.; Kim, C.H.; *et al.* Development of zebrafish medulloblastoma-like PNET model by TALEN-mediated somatic gene inactivation. *Oncotarget* **2017**, *8*, 55280–55297. [[CrossRef](#)]
80. Schultz, L.E.; Haltom, J.A.; Almeida, M.P.; Wierson, W.A.; Solin, S.L.; Weiss, T.J.; Helmer, J.A.; Sandquist, E.J.; Shive, H.R.; McGrail, M. Epigenetic regulators Rbbp4 and Hdac1 are overexpressed in a zebrafish model of RB1 embryonal brain tumor, and are required for neural progenitor survival and proliferation. *Dis. Model. Mech.* **2018**, *11*, dmm034124. [[CrossRef](#)]
81. Ignatius, M.S.; Hayes, M.N.; Moore, F.E.; Tang, Q.; Garcia, S.P.; Blackburn, P.R.; Baxi, K.; Wang, L.; Jin, A.; Ramakrishnan, A.; *et al.* tp53 deficiency causes a wide tumor spectrum and increases embryonal rhabdomyosarcoma metastasis in zebrafish. *Elife* **2018**, *7*, e37202. [[CrossRef](#)]
82. Jung, I.H.; Jung, D.E.; Chung, Y.Y.; Kim, K.S.; Park, S.W. Iroquois homeobox 1 acts as a true tumor suppressor in multiple organs by regulating cell cycle progression. *Neoplasia* **2019**, *21*, 1003–1014. [[CrossRef](#)] [[PubMed](#)]
83. Ablain, J.; Xu, M.; Rothschild, H.; Jordan, R.C.; Mito, J.K.; Daniels, B.H.; Bell, C.F.; Joseph, N.M.; Wu, H.; Bastian, B.C.; *et al.* Human tumor genomics and zebrafish modeling identify SPRED1 loss as a driver of mucosal melanoma. *Science* **2018**, *362*, 1055–1060. [[CrossRef](#)] [[PubMed](#)]
84. Burns, M.A.; Liao, Z.W.; Yamagata, N.; Pouliot, G.P.; Stevenson, K.E.; Neuberg, D.S.; Thorner, A.R.; Ducar, M.; Silverman, E.A.; Hunger, S.P.; *et al.* Hedgehog pathway mutations drive oncogenic transformation in high-risk T-cell acute lymphoblastic leukemia. *Leukemia* **2018**, *32*, 2126–2137. [[CrossRef](#)] [[PubMed](#)]
85. Oppel, F.; Tao, T.; Shi, H.; Ross, K.N.; Zimmerman, M.W.; He, S.; Tong, G.; Aster, J.C.; Look, A.T. Loss of atrx cooperates with p53-deficiency to promote the development of sarcomas and other malignancies. *PLoS Genet.* **2019**, *15*, e1008039. [[CrossRef](#)]
86. Oppel, F.; Ki, D.H.; Zimmermann, M.W.; Ross, K.N.; Tao, T.; Shi, H.; He, S.; Aster, J.C.; Look, A.T. suz12 inactivation in p53 and nf1 deficient zebrafish accelerates the onset of MPNSTs and expands the spectrum of tumor types to include adenocarcinoma, leukemia, and soft tissue sarcoma. *Dis. Model. Mech.* **2020**, Epub ahead of print, PMID: 32651197. [[CrossRef](#)]
87. Shi, H.; Tao, T.; Abraham, B.J.; Durbin, A.D.; Zimmerman, M.W.; Kadoch, C.; Look, A.T. ARID1A loss in neuroblastoma promotes the adrenergic-to-mesenchymal transition by regulating enhancer-mediated gene expression. *Sci. Adv.* **2020**, *6*, eaaz3440. [[CrossRef](#)]
88. Brandt, Z.J.; North, P.N.; Link, B.A. Somatic mutations of lats2 cause peripheral nerve sheath tumors in zebrafish. *Cells* **2019**, *8*, E972. [[CrossRef](#)]
89. Kim, D.H.; Kim, E.J.; Kim, D.H.; Park, S.W. Dact2 is involved in the regulation of epithelial-mesenchymal transition. *Biochem. Biophys. Res. Commun.* **2020**, *524*, 190–197. [[CrossRef](#)]
90. Anelli, V.; Villefranc, J.A.; Chhangawala, S.; Martinez-McFaline, R.; Riva, E.; Nguyen, A.; Verma, A.; Bareja, R.; Chen, Z.; Scognamiglio, T.; *et al.* Oncogenic BRAF disrupts thyroid morphogenesis and function via twist expression. *Elife* **2017**, *6*, e20728. [[CrossRef](#)]
91. Boch, J.; Scholze, H.; Schornack, S.; Landgraf, A.; Hahn, S.; Kay, S.; Lahaye, T.; Nickstadt, A.; Bonas, U. Breaking the code of DNA binding specificity of TAL-type III effectors. *Science* **2009**, *326*, 1509–1512. [[CrossRef](#)]
92. Moscou, M.J.; Bogdanove, A.J. A simple cipher governs DNA recognition by TAL effectors. *Science* **2009**, *326*, 1501. [[CrossRef](#)] [[PubMed](#)]



93. Huang, P.; Xiao, A.; Zhou, M.; Zhu, Z.; Lin, S.; Zhang, B. Heritable gene targeting in zebrafish using customized TALENs. *Nat. Biotechnol.* **2011**, *29*, 699–700. [[CrossRef](#)] [[PubMed](#)]
94. Sander, J.D.; Cade, L.; Khayter, C.; Reyon, D.; Peterson, R.T.; Joung, J.K.; Yeh, J.R. Targeted gene disruption in somatic zebrafish cells using engineered TALENs. *Nat. Biotechnol.* **2011**, *29*, 697–698. [[CrossRef](#)] [[PubMed](#)]
95. Doudna, J.A.; Charpentier, E. Genome editing. The new frontier of genome engineering with CRISPR-Cas9. *Science* **2014**, *346*, 1258096. [[CrossRef](#)] [[PubMed](#)]
96. Liu, K.; Petree, C.; Requena, T.; Varshney, P.; Varshney, G.K. Expanding the CRISPR toolbox in zebrafish for studying development and disease. *Front. Cell Dev. Biol.* **2019**, *7*, 13. [[CrossRef](#)] [[PubMed](#)]
97. Hwang, W.Y.; Fu, Y.; Reyon, D.; Maeder, M.L.; Tsai, S.Q.; Sander, J.D.; Peterson, R.T.; Yeh, J.R.; Joung, J.K. Efficient genome editing in zebrafish using a CRISPR-Cas system. *Nat. Biotechnol.* **2013**, *31*, 227–229. [[CrossRef](#)]
98. Chang, N.; Sun, C.; Gao, L.; Zhu, D.; Xu, X.; Zhu, X.; Xiong, J.W.; Xi, J.J. Genome editing with RNA-guided Cas9 nuclease in zebrafish embryos. *Cell Res.* **2013**, *23*, 465–472. [[CrossRef](#)]
99. Jao, L.E.; Wentz, S.R.; Chen, W. Efficient multiplex biallelic zebrafish genome editing using a CRISPR nuclease system. *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 13904–13909. [[CrossRef](#)]
100. Ablain, J.; Durand, E.M.; Yang, S.; Zhou, Y.; Zon, L.I. A CRISPR/Cas9 vector system for tissue-specific gene disruption in zebrafish. *Dev. Cell* **2015**, *32*, 756–764. [[CrossRef](#)]
101. Ablain, J.; Zon, L.I. Tissue-specific gene targeting using CRISPR/Cas9. *Methods Cell Biol.* **2016**, *135*, 189–202. [[CrossRef](#)]
102. Fu, Y.; Foden, J.A.; Khayter, C.; Maeder, M.L.; Reyon, D.; Joung, J.K.; Sander, J.D. High-frequency off-target mutagenesis induced by CRISPR-Cas nucleases in human cells. *Nat. Biotechnol.* **2013**, *31*, 822–826. [[CrossRef](#)] [[PubMed](#)]
103. Varshney, G.K.; Pei, W.; LaFave, M.C.; Idol, J.; Xu, L.; Gallardo, V.; Carrington, B.; Bishop, K.; Jones, M.; Li, M.; *et al.* High-throughput gene targeting and phenotyping in zebrafish using CRISPR/Cas9. *Genome Res.* **2015**, *25*, 1030–1042. [[CrossRef](#)] [[PubMed](#)]
104. Hruscha, A.; Krawitz, P.; Rechenberg, A.; Heinrich, V.; Hecht, J.; Haass, C.; Schmid, B. Efficient CRISPR/Cas9 genome editing with low off-target effects in zebrafish. *Development* **2013**, *140*, 4982–4987. [[CrossRef](#)] [[PubMed](#)]
105. Jamal, M.; Ullah, A.; Ahsan, M.; Tyagi, R.; Habib, Z.; Rehman, K. Improving CRISPR-Cas9 on-target specificity. *Curr. Issues Mol. Biol.* **2018**, *26*, 65–80. [[CrossRef](#)] [[PubMed](#)]
106. Cavodeassi, F.; Modolell, J.; Gómez-Skarmeta, J.L. The Iroquois family of genes: From body building to neural patterning. *Development* **2001**, *128*, 2847–2855.
107. Liu, X.; Zhang, J.; Liu, L.; Jiang, Y.; Ji, J.; Yan, R.; Zhu, Z.; Yu, Y. Protein arginine methyltransferase 5-mediated epigenetic silencing of IRX1 contributes to tumorigenicity and metastasis of gastric cancer. *Biochim. Biophys. Acta Mol. Basis Dis.* **2018**, *1864*, 2835–2844. [[CrossRef](#)]
108. Zhang, P.; Liu, N.; Xu, X.; Wang, Z.; Cheng, Y.; Jin, W.; Wang, X.; Yang, H.; Liu, H.; Zhang, Y.; *et al.* Clinical significance of Iroquois Homeobox Gene—IRX1 in human glioma. *Mol. Med. Rep.* **2018**, *17*, 4651–4656. [[CrossRef](#)]
109. Bennett, K.L.; Karpenko, M.; Lin, M.T.; Claus, R.; Arab, K.; Dyckhoff, G.; Plinkert, P.; Herpel, E.; Smiraglia, D.; Plass, C. Frequently methylated tumor suppressor genes in head and neck squamous cell carcinoma. *Cancer Res.* **2008**, *68*, 4494–4499. [[CrossRef](#)]
110. Stuart, G.W.; McMurray, J.V.; Westerfield, M. Replication, integration and stable germ-line transmission of foreign sequences injected into early zebrafish embryos. *Development* **1988**, *103*, 403–412.
111. Culp, P.; Nüsslein-Volhard, C.; Hopkins, N. High-frequency germ-line transmission of plasmid DNA sequences injected into fertilized zebrafish eggs. *Proc. Natl. Acad. Sci. USA* **1991**, *88*, 7953–7957. [[CrossRef](#)]
112. Gibbs, P.D.; Peek, A.; Thorgaard, G. An in vivo screen for the luciferase transgene in zebrafish. *Mol. Mar. Biol. Biotechnol.* **1994**, *3*, 307–316. [[PubMed](#)]
113. Udvardi, A.J.; Linney, E. Windows into development: Historic, current, and future perspectives on transgenic zebrafish. *Dev. Biol.* **2003**, *256*, 1–17. [[CrossRef](#)]
114. Thermes, V.; Grabher, C.; Ristoratore, F.; Bourrat, F.; Choulika, A.; Wittbrodt, J.; Joly, J.S. I-SceI meganuclease mediates highly efficient transgenesis in fish. *Mech. Dev.* **2002**, *118*, 91–98. [[CrossRef](#)]
115. Soroldoni, D.; Hogan, B.M.; Oates, A.C. Simple and efficient transgenesis with meganuclease constructs in zebrafish. *Methods Mol. Biol.* **2009**, *546*, 117–130. [[CrossRef](#)] [[PubMed](#)]
116. Kawakami, K.; Koga, A.; Hori, H.; Shima, A. Excision of the tol2 transposable element of the medaka fish, *Oryzias latipes*, in zebrafish, *Danio rerio*. *Gene* **1998**, *225*, 17–22. [[CrossRef](#)]

117. Kawakami, K.; Shima, A. Identification of the Tol2 transposase of the medaka fish *Oryzias latipes* that catalyzes excision of a nonautonomous Tol2 element in zebrafish *Danio rerio*. *Gene* **1999**, *240*, 239–244. [[CrossRef](#)]
118. Kawakami, K.; Shima, A.; Kawakami, N. Identification of a functional transposase of the Tol2 element, an Ac-like element from the Japanese medaka fish, and its transposition in the zebrafish germ lineage. *Proc. Natl. Acad. Sci. USA* **2000**, *97*, 11403–11408. [[CrossRef](#)]
119. Kawakami, K. Transgenesis and gene trap methods in zebrafish by using the Tol2 transposable element. *Methods Cell Biol.* **2004**, *77*, 201–222. [[CrossRef](#)]
120. Clark, K.J.; Urban, M.D.; Skuster, K.J.; Ekker, S.C. Transgenic zebrafish using transposable elements. *Methods Cell Biol.* **2011**, *104*, 137–149. [[CrossRef](#)]
121. Langenau, D.M.; Traver, D.; Ferrando, A.A.; Kutok, J.L.; Aster, J.C.; Kanki, J.P.; Lin, S.; Prochownik, E.; Trede, N.S.; Zon, L.I.; *et al.* Myc-induced T cell leukemia in transgenic zebrafish. *Science* **2003**, *299*, 887–890. [[CrossRef](#)]
122. Langenau, D.M.; Feng, H.; Berghmans, S.; Kanki, J.P.; Kutok, J.L.; Look, A.T. Cre/lox-regulated transgenic zebrafish model with conditional myc-induced T cell acute lymphoblastic leukemia. *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 6068–6073. [[CrossRef](#)] [[PubMed](#)]
123. Garcia, E.G.; Iyer, S.; Garcia, S.P.; Loontjens, S.; Sadreyev, R.I.; Speleman, F.; Langenau, D.M. Cell of origin dictates aggression and stem cell number in acute lymphoblastic leukemia. *Leukemia* **2018**, *32*, 1860–1865. [[CrossRef](#)] [[PubMed](#)]
124. Borga, C.; Park, G.; Foster, C.; Burroughs-Garcia, J.; Marchesin, M.; Shah, R.; Hasan, A.; Ahmed, S.T.; Bresolin, S.; Batchelor, L.; *et al.* Simultaneous B and T cell acute lymphoblastic leukemias in zebrafish driven by transgenic MYC: Implications for oncogenesis and lymphopoiesis. *Leukemia* **2019**, *33*, 333–347. [[CrossRef](#)] [[PubMed](#)]
125. Park, G.; Burroughs-Garcia, J.; Foster, C.A.; Hasan, A.; Borga, C.; Frazer, J.K. Zebrafish B cell acute lymphoblastic leukemia: New findings in an old model. *Oncotarget* **2020**, *11*, 1292–1305. [[CrossRef](#)]
126. Sinha, A.A.; Park, G.; Frazer, J.K. Tackling acute lymphoblastic leukemia-one fish at a time. *Int. J. Mol. Sci.* **2019**, *20*, 5313. [[CrossRef](#)]
127. Jessen, J.R.; Jessen, T.N.; Vogel, S.S.; Lin, S. Concurrent expression of recombination activating genes 1 and 2 in zebrafish olfactory sensory neurons. *Genesis* **2001**, *29*, 156–162. [[CrossRef](#)]
128. Langenau, D.M.; Ferrando, A.A.; Traver, D.; Kutok, J.L.; Hezel, J.P.; Kanki, J.P.; Zon, L.I.; Look, A.T.; Trede, N.S. In vivo tracking of T cell development, ablation, and engraftment in transgenic zebrafish. *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 7369–7374. [[CrossRef](#)]
129. Langenau, D.M.; Keefe, M.D.; Storer, N.Y.; Guyon, J.R.; Kutok, J.L.; Le, X.; Goessling, W.; Neubergh, D.S.; Kunkel, L.M.; Zon, L.I. Effects of RAS on the genesis of embryonal rhabdomyosarcoma. *Genes Dev.* **2007**, *21*, 1382–1395. [[CrossRef](#)]
130. Gutierrez, A.; Snyder, E.L.; Marino-Enriquez, A.; Zhang, Y.X.; Sioletic, S.; Kozakewich, E.; Grebliunaite, R.; Ou, W.B.; Sicinska, E.; Raut, C.P.; *et al.* Aberrant AKT activation drives well-differentiated liposarcoma. *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 16386–16391. [[CrossRef](#)]
131. Zhuravleva, J.; Paggetti, J.; Martin, L.; Hammann, A.; Solary, E.; Bastie, J.N.; Delva, L. MOZ/TIF2-induced acute myeloid leukaemia in transgenic fish. *Br. J. Haematol.* **2008**, *143*, 378–382. [[CrossRef](#)]
132. Lu, J.W.; Hou, H.A.; Hsieh, M.S.; Tien, H.F.; Lin, L.I. Overexpression of FLT3-ITD driven by spi-1 results in expanded myelopoiesis with leukemic phenotype in zebrafish. *Leukemia* **2016**, *30*, 2098–2101. [[CrossRef](#)] [[PubMed](#)]
133. Lu, J.W.; Hsieh, M.S.; Hou, H.A.; Chen, C.Y.; Tien, H.F.; Lin, L.I. Overexpression of SOX4 correlates with poor prognosis of acute myeloid leukemia and is leukemogenic in zebrafish. *Blood Cancer J.* **2017**, *7*, e593. [[CrossRef](#)] [[PubMed](#)]
134. Patton, E.E.; Widlund, H.R.; Kutok, J.L.; Kopani, K.R.; Amatruda, J.F.; Murphey, R.D.; Berghmans, S.; Mayhall, E.A.; Traver, D.; Fletcher, C.D.; *et al.* BRAF mutations are sufficient to promote nevi formation and cooperate with p53 in the genesis of melanoma. *Curr. Biol.* **2005**, *15*, 249–254. [[CrossRef](#)] [[PubMed](#)]
135. Dovey, M.; White, R.M.; Zon, L.I. Oncogenic NRAS cooperates with p53 loss to generate melanoma in zebrafish. *Zebrafish* **2009**, *6*, 397–404. [[CrossRef](#)] [[PubMed](#)]
136. Michailidou, C.; Jones, M.; Walker, P.; Kamarashev, J.; Kelly, A.; Hurlstone, A.F. Dissecting the roles of Raf- and PI3K-signalling pathways in melanoma formation and progression in a zebrafish model. *Dis. Model. Mech.* **2009**, *2*, 399–411. [[CrossRef](#)]
137. Mouti, M.A.; Dee, C.; Coupland, S.E.; Hurlstone, A.F. Minimal contribution of ERK1/2-MAPK signalling towards the maintenance of oncogenic GNAQQ209P-driven uveal melanomas in zebrafish. *Oncotarget* **2016**, *7*, 39654–39670. [[CrossRef](#)]

138. Storer, N.Y.; White, R.M.; Uong, A.; Price, E.; Nielsen, G.P.; Langenau, D.M.; Zon, L.I. Zebrafish rhabdomyosarcoma reflects the developmental stage of oncogene expression during myogenesis. *Development* **2013**, *140*, 3040–3050. [[CrossRef](#)]
139. Park, S.W.; Davison, J.M.; Rhee, J.; Hruban, R.H.; Maitra, A.; Leach, S.D. Oncogenic KRAS induces progenitor cell expansion and malignant transformation in zebrafish exocrine pancreas. *Gastroenterology* **2008**, *134*, 2080–2090. [[CrossRef](#)]
140. Yang, H.W.; Kutok, J.L.; Lee, N.H.; Piao, H.Y.; Fletcher, C.D.; Kanki, J.P.; Look, A.T. Targeted expression of human MYCN selectively causes pancreatic neuroendocrine tumors in transgenic zebrafish. *Cancer Res.* **2004**, *64*, 7256–7262. [[CrossRef](#)]
141. Nguyen, A.T.; Emelyanov, A.; Koh, C.H.; Spitsbergen, J.M.; Lam, S.H.; Mathavan, S.; Parinov, S.; Gong, Z. A high level of liver-specific expression of oncogenic Kras(V12) drives robust liver tumorigenesis in transgenic zebrafish. *Dis. Model. Mech.* **2011**, *4*, 801–813. [[CrossRef](#)]
142. Lu, J.W.; Yang, W.Y.; Tsai, S.M.; Lin, Y.M.; Chang, P.H.; Chen, J.R.; Wang, H.D.; Wu, J.L.; Jin, S.L.; Yuh, C.H. Liver-specific expressions of HBx and src in the p53 mutant trigger hepatocarcinogenesis in zebrafish. *PLoS ONE* **2013**, *8*, e76951. [[CrossRef](#)] [[PubMed](#)]
143. Mudbhary, R.; Hoshida, Y.; Chernyavskaya, Y.; Jacob, V.; Villanueva, A.; Fiel, M.I.; Chen, X.; Kojima, K.; Thung, S.; Bronson, R.T.; et al. UHRF1 overexpression drives DNA hypomethylation and hepatocellular carcinoma. *Cancer Cell* **2014**, *25*, 196–209. [[CrossRef](#)]
144. Wang, J.; Leng, X.; Wang, G.; Wan, X.; Cao, H. The construction of intrahepatic cholangiocarcinoma model in zebrafish. *Sci. Rep.* **2017**, *7*, 13419. [[CrossRef](#)]
145. Chou, Y.T.; Chen, L.Y.; Tsai, S.L.; Tu, H.C.; Lu, J.W.; Ciou, S.C.; Wang, H.D.; Yuh, C.H. Ribose-5-phosphate isomerase A overexpression promotes liver cancer development in transgenic zebrafish via activation of ERK and  $\beta$ -catenin pathways. *Carcinogenesis* **2019**, *40*, 461–473. [[CrossRef](#)] [[PubMed](#)]
146. Su, Z.L.; Su, C.W.; Huang, Y.L.; Yang, W.Y.; Sampurna, B.P.; Ouchi, T.; Lee, K.L.; Wu, C.S.; Wang, H.D.; Yuh, C.H. A novel AURKA mutant-induced early-onset severe hepatocarcinogenesis greater than wild-type via activating different pathways in zebrafish. *Cancers* **2019**, *11*, 927. [[CrossRef](#)] [[PubMed](#)]
147. Evason, K.J.; Francisco, M.T.; Juric, V.; Balakrishnan, S.; del Pilar Lopez Pazmino, M.; Gordan, J.D.; Kakar, S.; Spitsbergen, J.; Goga, A.; Stainier, D.Y. Identification of chemical inhibitors of  $\beta$ -catenin-driven liver tumorigenesis in zebrafish. *PLoS Genet.* **2015**, *11*, e1005305. [[CrossRef](#)]
148. Lu, J.W.; Liao, C.Y.; Yang, W.Y.; Lin, Y.M.; Jin, S.L.; Wang, H.D.; Yuh, C.H. Overexpression of endothelin 1 triggers hepatocarcinogenesis in zebrafish and promotes cell proliferation and migration through the AKT pathway. *PLoS ONE* **2014**, *9*, e85318. [[CrossRef](#)]
149. Zhu, S.; Lee, J.S.; Guo, F.; Shin, J.; Perez-Atayde, A.R.; Kutok, J.L.; Rodig, S.J.; Neuberger, D.S.; Helman, D.; Feng, H.; et al. Activated ALK collaborates with MYCN in neuroblastoma pathogenesis. *Cancer Cell* **2012**, *21*, 362–373. [[CrossRef](#)]
150. Modzelewska, K.; Boer, E.F.; Mosbrugger, T.L.; Picard, D.; Anderson, D.; Miles, R.R.; Kroll, M.; Oslund, W.; Pysker, T.J.; Schiffman, J.D.; et al. MEK inhibitors reverse growth of embryonal brain tumors derived from oligoneural precursor cells. *Cell Rep.* **2016**, *17*, 1255–1264. [[CrossRef](#)]
151. Zimmerman, M.W.; Liu, Y.; He, S.; Durbin, A.D.; Abraham, B.J.; Easton, J.; Shao, Y.; Xu, B.; Zhu, S.; Zhang, X.; et al. MYC drives a subset of high-risk pediatric neuroblastomas and is activated through mechanisms including enhancer hijacking and focal enhancer amplification. *Cancer Discov.* **2018**, *8*, 320–335. [[CrossRef](#)]
152. Gill, J.A.; Lowe, L.; Nguyen, J.; Liu, P.P.; Blake, T.; Venkatesh, B.; Aplan, P.D. Enforced expression of Simian virus 40 large T-antigen leads to testicular germ cell tumors in zebrafish. *Zebrafish* **2010**, *7*, 333–341. [[CrossRef](#)] [[PubMed](#)]
153. Damsky, W.E.; Bosenberg, M. Melanocytic nevi and melanoma: Unraveling a complex relationship. *Oncogene* **2017**, *36*, 5771–5792. [[CrossRef](#)] [[PubMed](#)]
154. Shieh, Y.S.; Chang, Y.S.; Hong, J.R.; Chen, L.J.; Jou, L.K.; Hsu, C.C.; Her, G.M. Increase of hepatic fat accumulation by liver specific expression of Hepatitis B virus X protein in zebrafish. *Biochim. Biophys. Acta* **2010**, *1801*, 721–730. [[CrossRef](#)] [[PubMed](#)]
155. He, S.; Mansour, M.R.; Zimmerman, M.W.; Ki, D.H.; Layden, H.M.; Akahane, K.; Gjini, E.; de Groh, E.D.; Perez-Atayde, A.R.; Zhu, S.; et al. Synergy between loss of NF1 and overexpression of MYCN in neuroblastoma is mediated by the GAP-related domain. *Elife* **2016**, *5*, e14713. [[CrossRef](#)] [[PubMed](#)]



156. Ignatius, M.S.; Chen, E.; Elpek, N.M.; Fuller, A.Z.; Tenente, I.M.; Clagg, R.; Liu, S.; Blackburn, J.S.; Linardic, C.M.; Rosenberg, A.E.; *et al.* In vivo imaging of tumor-propagating cells, regional tumor heterogeneity, and dynamic cell movements in embryonal rhabdomyosarcoma. *Cancer Cell* **2012**, *21*, 680–693. [[CrossRef](#)]
157. Sabaawy, H.E.; Azuma, M.; Embree, L.J.; Tsai, H.J.; Starost, M.F.; Hickstein, D.D. TEL-AML1 transgenic zebrafish model of precursor B cell acute lymphoblastic leukemia. *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 15166–15171. [[CrossRef](#)] [[PubMed](#)]
158. Andreasson, P.; Schwaller, J.; Anastasiadou, E.; Aster, J.; Gilliland, D.G. The expression of ETV6/CBFA2 (TEL/AML1) is not sufficient for the transformation of hematopoietic cell lines in vitro or the induction of hematologic disease in vivo. *Cancer Genet. Cytogenet.* **2001**, *130*, 93–104. [[CrossRef](#)]
159. Schindler, J.W.; Van Buren, D.; Foudi, A.; Krejci, O.; Qin, J.; Orkin, S.H.; Hock, H. TEL-AML1 corrupts hematopoietic stem cells to persist in the bone marrow and initiate leukemia. *Cell Stem Cell* **2009**, *5*, 43–53. [[CrossRef](#)]
160. Campos-Sanchez, E.; Vicente-Dueñas, C.; Rodríguez-Hernández, G.; Capstick, M.; Kuster, N.; Dasenbrock, C.; Sánchez-García, I.; Cobaleda, C. Novel ETV6-RUNX1 mouse model to study the role of ELF-MF in childhood b-acute lymphoblastic leukemia: A pilot study. *Bioelectromagnetics* **2019**, *40*, 343–353. [[CrossRef](#)]
161. Halpern, M.E.; Rhee, J.; Goll, M.G.; Akitake, C.M.; Parsons, M.; Leach, S.D. Gal4/UAS transgenic tools and their application to zebrafish. *Zebrafish* **2008**, *5*, 97–110. [[CrossRef](#)]
162. Fischer, J.A.; Giniger, E.; Maniatis, T.; Ptashne, M. GAL4 activates transcription in Drosophila. *Nature* **1988**, *332*, 853–856. [[CrossRef](#)] [[PubMed](#)]
163. Hartley, K.O.; Nutt, S.L.; Amaya, E. Targeted gene expression in transgenic Xenopus using the binary Gal4-UAS system. *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 1377–1382. [[CrossRef](#)] [[PubMed](#)]
164. Ornitz, D.M.; Moreadith, R.W.; Leder, P. Binary system for regulating transgene expression in mice: Targeting int-2 gene expression with yeast GAL4/UAS control elements. *Proc. Natl. Acad. Sci. USA* **1991**, *88*, 698–702. [[CrossRef](#)] [[PubMed](#)]
165. Scheer, N.; Campos-Ortega, J.A. Use of the Gal4-UAS technique for targeted gene expression in the zebrafish. *Mech. Dev.* **1999**, *80*, 153–158. [[CrossRef](#)]
166. Anelli, V.; Santoriello, C.; Distel, M.; Köster, R.W.; Ciccarelli, F.D.; Mione, M. Global repression of cancer gene expression in a zebrafish model of melanoma is linked to epigenetic regulation. *Zebrafish* **2010**, *6*, 417–424. [[CrossRef](#)]
167. Alghisi, E.; Distel, M.; Malagola, M.; Anelli, V.; Santoriello, C.; Herwig, L.; Krudewig, A.; Henkel, C.V.; Russo, D.; Mione, M.C. Targeting oncogene expression to endothelial cells induces proliferation of the myelo-erythroid lineage by repressing the Notch pathway. *Leukemia* **2013**, *27*, 2229–2241. [[CrossRef](#)]
168. Mayrhofer, M.; Gourain, V.; Reischl, M.; Affaticati, P.; Jenett, A.; Joly, J.S.; Benelli, M.; Demichelis, F.; Poliani, P.L.; Sieger, D.; *et al.* A novel brain tumour model in zebrafish reveals the role of YAP activation in MAPK- and PI3K-induced malignant growth. *Dis. Model. Mech.* **2017**, *10*, 15–28. [[CrossRef](#)]
169. Burger, A.; Vasilyev, A.; Tomar, R.; Selig, M.K.; Nielsen, G.P.; Peterson, R.T.; Drummond, I.A.; Haber, D.A. A zebrafish model of chordoma initiated by notochord-driven expression of HRASV12. *Dis. Model. Mech.* **2014**, *7*, 907–913. [[CrossRef](#)]
170. Liu, S.; Leach, S.D. Screening pancreatic oncogenes in zebrafish using the Gal4/UAS system. *Methods Cell Biol.* **2011**, *105*, 367–381. [[CrossRef](#)]
171. Jung, I.H.; Leem, G.L.; Jung, D.E.; Kim, M.H.; Kim, E.Y.; Kim, S.H.; Park, H.C.; Park, S.W. Glioma is formed by active Akt1 alone and promoted by active Rac1 in transgenic zebrafish. *Neuro Oncol.* **2013**, *15*, 290–304. [[CrossRef](#)]
172. Park, J.T.; Johnson, N.; Liu, S.; Levesque, M.; Wang, Y.J.; Ho, H.; Huso, D.; Maitra, A.; Parsons, M.J.; Prescott, J.D.; *et al.* Differential in vivo tumorigenicity of diverse KRAS mutations in vertebrate pancreas: A comprehensive survey. *Oncogene* **2015**, *34*, 2801–2806. [[CrossRef](#)]
173. Lakso, M.; Sauer, B.; Mosinger, B., Jr.; Lee, E.J.; Manning, R.W.; Yu, S.H.; Mulder, K.L.; Westphal, H. Targeted oncogene activation by site-specific recombination in transgenic mice. *Proc. Natl. Acad. Sci. USA* **1992**, *89*, 6232–6236. [[CrossRef](#)] [[PubMed](#)]
174. Lewandoski, M. Conditional control of gene expression in the mouse. *Nat. Rev. Genet.* **2001**, *2*, 743–755. [[CrossRef](#)] [[PubMed](#)]
175. Seok, S.H.; Na, Y.R.; Han, J.H.; Kim, T.H.; Jung, H.; Lee, B.H.; Emelyanov, A.; Parinov, S.; Park, J.H. Cre/loxP-regulated transgenic zebrafish model for neural progenitor-specific oncogenic Kras expression. *Cancer Sci.* **2010**, *101*, 149–154. [[CrossRef](#)] [[PubMed](#)]
176. Oh, S.; Park, J.T. Zebrafish model of KRAS-initiated pancreatic endocrine tumor. *Anim. Cells Syst.* **2019**, *23*, 209–218. [[CrossRef](#)] [[PubMed](#)]

177. Halloran, M.C.; Sato-Maeda, M.; Warren, J.T.; Su, F.; Lele, Z.; Krone, P.H.; Kuwada, J.Y.; Shoji, W. Laser-induced gene expression in specific cells of transgenic zebrafish. *Development* **2000**, *127*, 1953–1960.
178. Shoji, W.; Sato-Maeda, M. Application of heat shock promoter in transgenic zebrafish. *Dev. Growth Differ.* **2008**, *50*, 401–406. [[CrossRef](#)]
179. Zou, J.; Guo, Y.; Guettouche, T.; Smith, D.F.; Voellmy, R. Repression of heat shock transcription factor HSF1 activation by HSP90 (HSP90 complex) that forms a stress-sensitive complex with HSF1. *Cell* **1998**, *94*, 471–480. [[CrossRef](#)]
180. Santoriello, C.; Deflorian, G.; Pezzimenti, F.; Kawakami, K.; Lanfranccone, L.; d'Adda di Fagagna, F.; Mione, M. Expression of H-RASV12 in a zebrafish model of Costello syndrome causes cellular senescence in adult proliferating cells. *Dis. Model. Mech.* **2009**, *2*, 56–67. [[CrossRef](#)]
181. Yeh, J.R.; Munson, K.M.; Chao, Y.L.; Peterson, Q.P.; Macrae, C.A.; Peterson, R.T. AML1-ETO reprograms hematopoietic cell fate by downregulating scl expression. *Development* **2008**, *135*, 401–410. [[CrossRef](#)]
182. Xu, M.; Ye, Y.; Ye, Z.; Xu, S.; Liu, W.; Xu, J.; Zhang, Y.; Liu, Q.; Huang, Z.; Zhang, W. Human BCR/ABL1 induces chronic myeloid leukemia-like disease in zebrafish. *Haematologica* **2020**, *105*, 674–686. [[CrossRef](#)] [[PubMed](#)]
183. Leacock, S.W.; Basse, A.N.; Chandler, G.L.; Kirk, A.M.; Rakheja, D.; Amatruda, J.F. A zebrafish transgenic model of Ewing's sarcoma reveals conserved mediators of EWS-FLI1 tumorigenesis. *Dis. Model. Mech.* **2012**, *5*, 95–106. [[CrossRef](#)] [[PubMed](#)]
184. Shen, L.J.; Chen, F.Y.; Zhang, Y.; Cao, L.F.; Kuang, Y.; Zhong, M.; Wang, T.; Zhong, H. MYCN transgenic zebrafish model with the characterization of acute myeloid leukemia and altered hematopoiesis. *PLoS ONE* **2013**, *8*, e59070. [[CrossRef](#)] [[PubMed](#)]
185. Shen, L.; Zhu, J.; Chen, F.; Lin, W.; Cai, J.; Zhong, J.; Zhong, H. RUNX1-Evi-1 fusion gene inhibited differentiation and apoptosis in myelopoiesis: An in vivo study. *BMC Cancer* **2015**, *15*, 970. [[CrossRef](#)] [[PubMed](#)]
186. Feng, H.; Langenau, D.M.; Madge, J.A.; Quinkertz, A.; Gutierrez, A.; Neuberger, D.S.; Kanki, J.P.; Look, A.T. Heat-shock induction of T-cell lymphoma/leukaemia in conditional Cre/lox-regulated transgenic zebrafish. *Br. J. Haematol.* **2007**, *138*, 169–175. [[CrossRef](#)]
187. Le, X.; Langenau, D.M.; Keefe, M.D.; Kutok, J.L.; Neuberger, D.S.; Zon, L.I. Heat shock-inducible Cre/Lox approaches to induce diverse types of tumors and hyperplasia in transgenic zebrafish. *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 9410–9415. [[CrossRef](#)]
188. Forrester, A.M.; Grabher, C.; McBride, E.R.; Boyd, E.R.; Vigerstad, M.H.; Edgar, A.; Kai, F.B.; Da'as, S.I.; Payne, E.; Look, A.T.; et al. NUP98-HOXA9-transgenic zebrafish develop a myeloproliferative neoplasm and provide new insight into mechanisms of myeloid leukaemogenesis. *Br. J. Haematol.* **2011**, *155*, 167–181. [[CrossRef](#)]
189. Berens, C.; Hillen, W. Gene regulation by tetracyclines. Constraints of resistance regulation in bacteria shape TetR for application in eukaryotes. *Eur. J. Biochem.* **2003**, *270*, 3109–3121. [[CrossRef](#)]
190. Gossen, M.; Bujard, H. Tight control of gene expression in mammalian cells by tetracycline-responsive promoters. *Proc. Natl. Acad. Sci. USA* **1992**, *89*, 5547–5551. [[CrossRef](#)]
191. Gossen, M.; Freundlieb, S.; Bender, G.; Müller, G.; Hillen, W.; Bujard, H. Transcriptional activation by tetracyclines in mammalian cells. *Science* **1995**, *268*, 1766–1769. [[CrossRef](#)]
192. Stebbins, M.J.; Urlinger, S.; Byrne, G.; Bello, B.; Hillen, W.; Yin, J.C. Tetracycline-inducible systems for Drosophila. *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 10775–10780. [[CrossRef](#)] [[PubMed](#)]
193. Furth, P.A.; St Onge, L.; Böger, H.; Gruss, P.; Gossen, M.; Kistner, A.; Bujard, H.; Hennighausen, L. Temporal control of gene expression in transgenic mice by a tetracycline-responsive promoter. *Proc. Natl. Acad. Sci. USA* **1994**, *91*, 9302–9306. [[CrossRef](#)] [[PubMed](#)]
194. Huang, C.J.; Jou, T.S.; Ho, Y.L.; Lee, W.H.; Jeng, Y.T.; Hsieh, F.J.; Tsai, H.J. Conditional expression of a myocardium-specific transgene in zebrafish transgenic lines. *Dev. Dyn.* **2005**, *233*, 1294–1303. [[CrossRef](#)] [[PubMed](#)]
195. Ju, B.; Chen, W.; Orr, B.A.; Spitsbergen, J.M.; Jia, S.; Eden, C.J.; Henson, H.E.; Taylor, M.R. Oncogenic KRAS promotes malignant brain tumors in zebrafish. *Mol. Cancer* **2015**, *14*, 18. [[CrossRef](#)]
196. Li, Z.; Huang, X.; Zhan, H.; Zeng, Z.; Li, C.; Spitsbergen, J.M.; Meierjohann, S.; Schartl, M.; Gong, Z. Inducible and repressible oncogene-addicted hepatocellular carcinoma in Tet-on xmrk transgenic zebrafish. *J. Hepatol.* **2012**, *56*, 419–425. [[CrossRef](#)]

197. Li, Z.; Zheng, W.; Wang, Z.; Zeng, Z.; Zhan, H.; Li, C.; Zhou, L.; Yan, C.; Spitsbergen, J.M.; Gong, Z. A transgenic zebrafish liver tumor model with inducible Myc expression reveals conserved Myc signatures with mammalian liver tumors. *Dis. Model. Mech.* **2013**, *6*, 414–423. [[CrossRef](#)]
198. Zheng, W.; Li, Z.; Nguyen, A.T.; Li, C.; Emelyanov, A.; Gong, Z. Xmrk, kras and myc transgenic zebrafish liver cancer models share molecular signatures with subsets of human hepatocellular carcinoma. *PLoS ONE* **2014**, *9*, e91179. [[CrossRef](#)]
199. Chew, T.W.; Liu, X.J.; Liu, L.; Spitsbergen, J.M.; Gong, Z.; Low, B.C. Crosstalk of Ras and Rho: Activation of RhoA abates Kras-induced liver tumorigenesis in transgenic zebrafish models. *Oncogene* **2014**, *33*, 2717–2727. [[CrossRef](#)]
200. Li, Y.; Li, H.; Spitsbergen, J.M.; Gong, Z. Males develop faster and more severe hepatocellular carcinoma than females in kras(V12) transgenic zebrafish. *Sci. Rep.* **2017**, *7*, 41280. [[CrossRef](#)]
201. Yang, Q.; Yan, C.; Wang, X.; Gong, Z. Leptin induces muscle wasting in a zebrafish kras-driven hepatocellular carcinoma (HCC) model. *Dis. Model. Mech.* **2019**, *12*, dmm038240. [[CrossRef](#)]
202. Evans, R.M. The steroid and thyroid hormone receptor superfamily. *Science* **1988**, *240*, 889–895. [[CrossRef](#)] [[PubMed](#)]
203. Webster, N.J.; Green, S.; Jin, J.R.; Chambon, P. The hormone-binding domains of the estrogen and glucocorticoid receptors contain an inducible transcription activation function. *Cell* **1988**, *54*, 199–207. [[CrossRef](#)]
204. Picard, D.; Salser, S.J.; Yamamoto, K.R. A movable and regulable inactivation function within the steroid binding domain of the glucocorticoid receptor. *Cell* **1988**, *54*, 1073–1080. [[CrossRef](#)]
205. Eilers, M.; Picard, D.; Yamamoto, K.R.; Bishop, J.M. Chimaeras of myc oncoprotein and steroid receptors cause hormone-dependent transformation of cells. *Nature* **1989**, *340*, 66–68. [[CrossRef](#)]
206. Jackson, P.; Baltimore, D.; Picard, D. Hormone-conditional transformation by fusion proteins of c-Abl and its transforming variants. *Embo J.* **1993**, *12*, 2809–2819. [[CrossRef](#)]
207. Picard, D. Regulation of protein function through expression of chimaeric proteins. *Curr. Opin. Biotechnol.* **1994**, *5*, 511–515. [[CrossRef](#)]
208. Gutierrez, A.; Grebliunaite, R.; Feng, H.; Kozakewich, E.; Zhu, S.; Guo, F.; Payne, E.; Mansour, M.; Dahlberg, S.E.; Neuberg, D.S.; *et al.* Pten mediates Myc oncogene dependence in a conditional zebrafish model of T cell acute lymphoblastic leukemia. *J. Exp. Med.* **2011**, *208*, 1595–1603. [[CrossRef](#)]
209. Littlewood, T.D.; Hancock, D.C.; Danielian, P.S.; Parker, M.G.; Evan, G.I. A modified oestrogen receptor ligand-binding domain as an improved switch for the regulation of heterologous proteins. *Nucleic Acids Res.* **1995**, *23*, 1686–1690. [[CrossRef](#)]
210. Nakayama, J.; Lu, J.W.; Makinoshima, H.; Gong, Z. A novel zebrafish model of metastasis identifies the hsd11β1 inhibitor adrenosterone as a suppressor of epithelial-mesenchymal transition and metastatic dissemination. *Mol. Cancer Res.* **2020**, *18*, 477–487. [[CrossRef](#)]
211. Baulieu, E.E. Contraception and other clinical applications of RU 486, an antiprogestone at the receptor. *Science* **1989**, *245*, 1351–1357. [[CrossRef](#)]
212. Wang, Y.; O'Malley, B.W., Jr.; Tsai, S.Y.; O'Malley, B.W. A regulatory system for use in gene transfer. *Proc. Natl. Acad. Sci. USA* **1994**, *91*, 8180–8184. [[CrossRef](#)] [[PubMed](#)]
213. Kellendonk, C.; Tronche, F.; Monaghan, A.P.; Angrand, P.O.; Stewart, F.; Schütz, G. Regulation of Cre recombinase activity by the synthetic steroid RU 486. *Nucleic Acids Res.* **1996**, *24*, 1404–1411. [[CrossRef](#)] [[PubMed](#)]
214. Emelyanov, A.; Parinov, S. Mifepristone-inducible LexPR system to drive and control gene expression in transgenic zebrafish. *Dev. Biol.* **2008**, *320*, 113–121. [[CrossRef](#)]
215. Nguyen, A.T.; Emelyanov, A.; Koh, C.H.; Spitsbergen, J.M.; Parinov, S.; Gong, Z. An inducible kras(V12) transgenic zebrafish model for liver tumorigenesis and chemical drug screening. *Dis. Model. Mech.* **2012**, *5*, 63–72. [[CrossRef](#)]
216. Yan, C.; Yang, Q.; Shen, H.M.; Spitsbergen, J.M.; Gong, Z. Chronically high level of tgfb1a induction causes both hepatocellular carcinoma and cholangiocarcinoma via a dominant Erk pathway in zebrafish. *Oncotarget* **2017**, *8*, 77096–77109. [[CrossRef](#)]
217. Yan, C.; Yang, Q.; Gong, Z. Transgenic expression of tgfb1a induces hepatic inflammation, fibrosis and metastasis in zebrafish. *Biochem. Biophys. Res. Commun.* **2019**, *509*, 175–181. [[CrossRef](#)]
218. Kenyon, A.; Gavriouchkina, D.; Zorman, J.; Chong-Morrison, V.; Napolitani, G.; Cerundolo, V.; Sauka-Spengler, T. Generation of a double binary transgenic zebrafish model to study myeloid gene regulation in response to oncogene activation in melanocytes. *Dis. Model. Mech.* **2018**, *11*, dmm030056. [[CrossRef](#)]

219. Lu, J.W.; Raghuram, D.; Fong, P.A.; Gong, Z. Inducible intestine-specific expression of *krasv12* triggers intestinal tumorigenesis in transgenic zebrafish. *Neoplasia* **2018**, *20*, 1187–1197. [[CrossRef](#)]
220. Nguyen, A.T.; Koh, V.; Spitsbergen, J.M.; Gong, Z. Development of a conditional liver tumor model by mifepristone-inducible Cre recombination to control oncogenic *kras* V12 expression in transgenic zebrafish. *Sci. Rep.* **2016**, *6*, 19559. [[CrossRef](#)]
221. Schwenk, F.; Kuhn, R.; Angrand, P.O.; Rajewsky, K.; Stewart, A.F. Temporally and spatially regulated somatic mutagenesis in mice. *Nucleic Acids Res.* **1998**, *26*, 1427–1432. [[CrossRef](#)]
222. Indra, A.K.; Warot, X.; Brocard, J.; Bornert, J.M.; Xiao, J.H.; Chambon, P.; Metzger, D. Temporally-controlled site-specific mutagenesis in the basal layer of the epidermis: Comparison of the recombinase activity of the tamoxifen-inducible Cre-ER(T) and Cre-ER(T2) recombinases. *Nucleic Acids Res.* **1999**, *27*, 4324–4327. [[CrossRef](#)] [[PubMed](#)]
223. Jopling, C.; Sleep, E.; Raya, M.; Martí, M.; Raya, A.; Izpisua Belmonte, J.C. Zebrafish heart regeneration occurs by cardiomyocyte dedifferentiation and proliferation. *Nature* **2010**, *464*, 606–609. [[CrossRef](#)] [[PubMed](#)]
224. Jungke, P.; Hans, S.; Brand, M. The zebrafish CreZoo: An easy-to-handle database for novel CreER(T2)-driver lines. *Zebrafish* **2013**, *10*, 259–263. [[CrossRef](#)] [[PubMed](#)]
225. Kalasekar, S.M.; Kotiyal, S.; Conley, C.; Phan, C.; Young, A.; Evason, K.J. Heterogeneous beta-catenin activation is sufficient to cause hepatocellular carcinoma in zebrafish. *Biol. Open* **2019**, *8*, bio047829. [[CrossRef](#)]
226. Park, J.T.; Leach, S.D. Zebrafish model of KRAS-initiated pancreatic cancer. *Anim. Cells Syst.* **2018**, *22*, 353–359. [[CrossRef](#)]
227. Tomasetti, C.; Marchionni, L.; Nowak, M.A.; Parmigiani, G.; Vogelstein, B. Only three driver gene mutations are required for the development of lung and colorectal cancers. *Proc. Natl. Acad. Sci. USA* **2015**, *112*, 118–123. [[CrossRef](#)]
228. Hornsby, C.; Page, K.M.; Tomlinson, I.P. What can we learn from the population incidence of cancer? Armitage and Doll revisited. *Lancet Oncol.* **2007**, *8*, 1030–1038. [[CrossRef](#)]
229. Häusser, M. Optogenetics: The age of light. *Nat. Methods* **2014**, *11*, 1012–1014. [[CrossRef](#)]
230. Zhang, K.; Cui, B. Optogenetic control of intracellular signaling pathways. *Trends Biotechnol.* **2015**, *33*, 92–100. [[CrossRef](#)]
231. Rogers, K.W.; Müller, P. Optogenetic approaches to investigate spatiotemporal signaling during development. *Curr. Top. Dev. Biol.* **2020**, *137*, 37–77. [[CrossRef](#)]
232. Feng, Z.; Nam, S.; Hamouri, F.; Aujard, I.; Ducos, B.; Vriz, S.; Volovitch, M.; Jullien, L.; Lin, S.; Weiss, S.; *et al.* Optical control of tumor induction in the zebrafish. *Sci. Rep.* **2017**, *7*, 9195. [[CrossRef](#)] [[PubMed](#)]
233. Ung, C.Y.; Guo, F.; Zhang, X.; Zhu, Z.; Zhu, S. Mosaic zebrafish transgenesis for functional genomic analysis of candidate cooperative genes in tumor pathogenesis. *J. Vis. Exp.* **2015**, 52567. [[CrossRef](#)] [[PubMed](#)]
234. Köster, R.W.; Fraser, S.E. Tracing transgene expression in living zebrafish embryos. *Dev. Biol.* **2001**, *233*, 329–346. [[CrossRef](#)] [[PubMed](#)]
235. Langenau, D.M.; Keefe, M.D.; Storer, N.Y.; Jette, C.A.; Smith, A.C.; Ceol, C.J.; Bourque, C.; Look, A.T.; Zon, L.I. Co-injection strategies to modify radiation sensitivity and tumor initiation in transgenic Zebrafish. *Oncogene* **2008**, *27*, 4242–4248. [[CrossRef](#)] [[PubMed](#)]
236. Ceol, C.J.; Houvras, Y.; Jane-Valbuena, J.; Bilodeau, S.; Orlando, D.A.; Battisti, V.; Fritsch, L.; Lin, W.M.; Hollmann, T.J.; Ferré, F.; *et al.* The histone methyltransferase SETDB1 is recurrently amplified in melanoma and accelerates its onset. *Nature* **2011**, *471*, 513–517. [[CrossRef](#)]
237. McConnell, A.M.; Mito, J.K.; Ablain, J.; Dang, M.; Formichella, L.; Fisher, D.E.; Zon, L.I. Neural crest state activation in NRAS driven melanoma, but not in NRAS-driven melanocyte expansion. *Dev. Biol.* **2019**, *449*, 107–114. [[CrossRef](#)]
238. Hosono, Y.; Niknafs, Y.S.; Prensner, J.R.; Iyer, M.K.; Dhanasekaran, S.M.; Mehra, R.; Pitchiaya, S.; Tien, J.; Escara-Wilke, J.; Poliakov, A.; *et al.* Oncogenic role of THOR, a conserved cancer/testis long non-coding RNA. *Cell* **2017**, *171*, 1559–1572. [[CrossRef](#)]
239. Callahan, S.J.; Tepan, S.; Zhang, Y.M.; Lindsay, H.; Burger, A.; Campbell, N.R.; Kim, I.S.; Hollmann, T.J.; Studer, L.; Mosimann, C.; *et al.* Cancer modeling by transgene electroporation in adult zebrafish (TEAZ). *Dis. Model. Mech.* **2018**, *11*, dmm034561. [[CrossRef](#)]
240. Ung, C.Y.; Lam, S.H.; Gong, Z. Comparative transcriptome analyses revealed conserved biological and transcription factor target modules between the zebrafish and human tumors. *Zebrafish* **2009**, *6*, 425–431. [[CrossRef](#)]
241. Zhang, G.; Hoersch, S.; Amsterdam, A.; Whittaker, C.A.; Beert, E.; Catchen, J.M.; Farrington, S.; Postlethwait, J.H.; Legius, E.; Hopkins, N.; *et al.* Comparative oncogenomic analysis of copy number alterations in human and zebrafish tumors enables cancer driver discovery. *PLoS Genet.* **2013**, *9*, e1003734. [[CrossRef](#)]

242. Nakayama, J.; Makinoshima, H. Zebrafish-based screening models for the identification of anti-metastatic drugs. *Molecules* **2020**, *25*, 2407. [[CrossRef](#)] [[PubMed](#)]
243. Fazio, M.; Ablain, J.; Chuan, Y.; Langenau, D.M.; Zon, L.I. Zebrafish patient avatars in cancer biology and precision cancer therapy. *Nat. Rev. Cancer* **2020**, *20*, 263–273. [[CrossRef](#)] [[PubMed](#)]
244. Chen, J.; Jette, C.; Kanki, J.P.; Aster, J.C.; Look, A.T.; Griffin, J.D. NOTCH1-induced T-cell leukemia in transgenic zebrafish. *Leukemia* **2007**, *21*, 462–471. [[CrossRef](#)] [[PubMed](#)]
245. Liao, W.S.; Tan, S.H.; Ngoc, P.C.T.; Wang, C.Q.; Tergaonkar, V.; Feng, H.; Gong, Z.; Osato, M.; Look, A.T.; Sanda, T. Aberrant activation of the GIMAP enhancer by oncogenic transcription factors in T-cell acute lymphoblastic leukemia. *Leukemia* **2017**, *31*, 1798–1807. [[CrossRef](#)]
246. Langenau, D.M.; Jette, C.; Berghmans, S.; Palomero, T.; Kanki, J.P.; Kutok, J.L.; Look, A.T. Suppression of apoptosis by bcl-2 overexpression in lymphoid cells of transgenic zebrafish. *Blood* **2005**, *105*, 3278–3285. [[CrossRef](#)]
247. Onnebo, S.M.; Condrón, M.M.; McPhee, D.O.; Lieschke, G.J.; Ward, A.C. Hematopoietic perturbation in zebrafish expressing a tel-jak2a fusion. *Exp. Hematol.* **2005**, *33*, 182–188. [[CrossRef](#)]
248. Onnebo, S.M.; Rasighaemi, P.; Kumar, J.; Liongue, C.; Ward, A.C. Alternative TEL-JAK2 fusions associated with T-cell acute lymphoblastic leukemia and atypical chronic myelogenous leukemia dissected in zebrafish. *Haematologica* **2012**, *97*, 1895–1903. [[CrossRef](#)]
249. Lister, J.A.; Capper, A.; Zeng, Z.; Mathers, M.E.; Richardson, J.; Paranthaman, K.; Jackson, I.J.; Patton, E.E. A conditional zebrafish MITF mutation reveals MITF levels are critical for melanoma promotion vs. regression in vivo. *J. Invest. Derm.* **2014**, *134*, 133–140. [[CrossRef](#)]
250. Santoriello, C.; Gennaro, E.; Anelli, V.; Distel, M.; Kelly, A.; Köster, R.W.; Hurlstone, A.; Mione, M. Kita driven expression of oncogenic HRAS leads to early onset and highly penetrant melanoma in zebrafish. *PLoS ONE* **2010**, *5*, e15170. [[CrossRef](#)]
251. Iwanaga, R.; Truong, B.T.; Lambert, K.A.; Vyas, R.; Orlicky, D.; Shellman, Y.G.; Tan, A.C.; Artinger, K.B. Loss of prdm1a accelerates melanoma onset and progression. *bioRxiv* **2019**, Epub ahead of print, PMID: 32562448. [[CrossRef](#)]
252. Chu, C.Y.; Chen, C.F.; Rajendran, R.S.; Shen, C.N.; Chen, T.H.; Yen, C.C.; Chuang, C.K.; Lin, D.S.; Hsiao, C.D. Overexpression of Akt1 enhances adipogenesis and leads to lipoma formation in zebrafish. *PLoS ONE* **2012**, *7*, e36474. [[CrossRef](#)] [[PubMed](#)]
253. Ju, B.; Chen, W.; Spitsbergen, J.M.; Lu, J.; Vogel, P.; Peters, J.L.; Wang, Y.D.; Orr, B.A.; Wu, J.; Henson, H.E.; *et al.* Activation of sonic hedgehog signaling in neural progenitor cells promotes glioma development in the zebrafish optic pathway. *Oncogenesis* **2014**, *3*, e96. [[CrossRef](#)] [[PubMed](#)]
254. Zhu, S.; Zhang, X.; Weichert-Leahey, N.; Dong, Z.; Zhang, C.; Lopez, G.; Tao, T.; He, S.; Wood, A.C.; Oldridge, D.; *et al.* LMO1 synergizes with MYCN to promote neuroblastoma initiation and metastasis. *Cancer Cell* **2017**, *32*, C310–C323. [[CrossRef](#)] [[PubMed](#)]
255. Tao, T.; Shi, H.; Mariani, L.; Abraham, B.J.; Durbin, A.D.; Zimmerman, M.W.; Powers, J.T.; Missios, P.; Ross, K.N.; Perez-Atayde, A.R.; *et al.* LIN28B regulates transcription and potentiates MYCN-induced neuroblastoma through binding to ZNF143 at target gene promoters. *Proc. Natl. Acad. Sci. USA* **2020**, *28*, 16516–16526. [[CrossRef](#)] [[PubMed](#)]
256. Balci, T.B.; Prykhozhiy, S.V.; The, E.M.; Da’as, S.I.; McBride, E.; Liwski, R.; Chute, I.C.; Leger, D.; Lewis, S.M.; Berman, J.N. A transgenic zebrafish model expressing KIT-D816V recapitulates features of aggressive systemic mastocytosis. *Br. J. Haematol.* **2014**, *167*, 48–61. [[CrossRef](#)] [[PubMed](#)]
257. Liu, N.A.; Jiang, H.; Ben-Shlomo, A.; Wawrowsky, K.; Fan, X.M.; Lin, S.; Melmed, S. Targeting zebrafish and murine pituitary corticotroph tumors with a cyclin-dependent kinase (CDK) inhibitor. *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 8414–8419. [[CrossRef](#)]

### 3.3 Conclusion

L'amélioration et la multiplication des technologies a permis des progrès considérables concernant notre compréhension du cancer. Des mutations récurrentes pour certains types de cancers ont été identifiées par séquençage de prélèvements issus de patients. L'étude *in vitro* des lignées cellulaires cancéreuses permet d'identifier des mécanismes, des gènes et des voies de signalisation associées à ces mutations, mais ne permet pas l'étude des cellules cancéreuses dans un système biologique complexe. Les interactions entre les cellules cancéreuses et le microenvironnement par exemple conditionnent le développement tumoral (Polyak *et al.*, 2009). La génération de modèles portant les mutations et développant des tumeurs est donc indispensable pour simuler la situation clinique. Dans ce domaine, le poisson zèbre s'est imposé comme un incontournable grâce à ses nombreuses qualités, ne serait-ce que son temps de génération court. Les tumeurs développées par les lignées mutantes et transgéniques ont été analysées et sont très similaires, tant au niveau moléculaire que histologique, aux tumeurs humaines. Ces modèles contribuent au décryptage de certains mécanismes oncogéniques ainsi qu'à la découverte de nouvelles molécules thérapeutiques (Li *et al.*, 2012).

# DISCUSSION

---

## 1. Le poisson zèbre comme modèle d'étude de la répression Polycomb

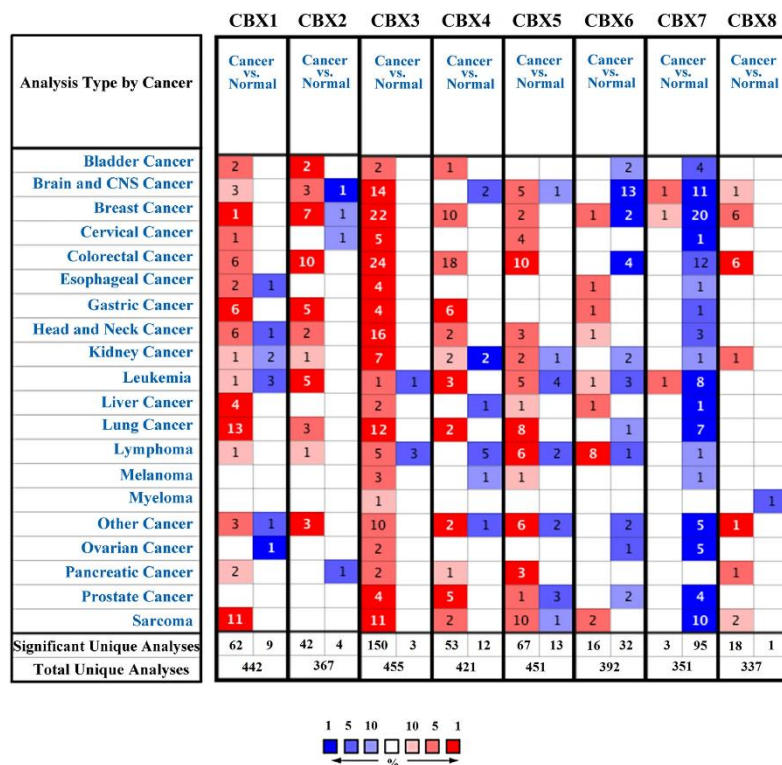
### 1.1 La conservation des gènes *cbx*

Le gène *Pc* de la drosophile a été conservé au cours de l'évolution. La recherche de séquences homologues dans les génomes séquencés chez les mammifères a identifié plusieurs orthologues : CBX2, CBX4, CBX6, CBX7 et CBX8. Ces orthologues sont tous conservés chez les téléostéens. La famille de protéines CBX est connue pour son rôle dans le recrutement spécifique du complexe PRC1 notamment au cours du développement. Les études fonctionnelles sur le modèle murin montrent que ces protéines ne sont pas indispensables au développement précoce des embryons. Les phénotypes observés sont réduits en comparaison des phénotypes observés en cas de perte de fonction de RNF2 (Table 2). Même si l'activité des complexes PRC1 non canoniques peut en partie expliquer ces résultats, cela suggérerait une redondance fonctionnelle des protéines CBX. Néanmoins, la conservation de chaque orthologue évoque une autre réalité. Les sous-unités du complexe PRC1 ne présentent pas toutes le même degré de conservation. La protéine RING1 par exemple est absente chez le poisson zèbre. Les protéines PCGF ne sont également pas toutes conservées, aucun orthologue de PCGF3 n'a été identifié chez le poisson zèbre. La protéine PCGF3 participe avec PCGF5 à la formation d'un complexe PRC1 non canonique. Leurs fonctions sont très similaires. L'absence de PCGF3 chez le poisson zèbre confirme la redondance fonctionnelle de ces deux protéines. La conservation de chacune des protéines *Cbx* chez le poisson zèbre indique, qu'au moins au sein de cette espèce, la spécificité d'action de chaque protéine est nécessaire.

Bien que les gènes *cbx* ait été identifiés chez le poisson zèbres, leur profil d'expression restait inconnu, tant au niveau spatial que temporel. Les gènes *cbx2*, *cbx4*, *cbx6a/b*, *cbx7a/b* et *cbx8a/b* sont tous déposés maternellement dans les œufs, exprimés au cours du développement ainsi qu'au stade adulte. L'hybridation *in situ* et la RT-PCR ne sont pas des méthodes quantitatives, mais laissent apercevoir des différences concernant l'abondance des transcrits au cours du développement. Une approche par RT-qPCR pourra permettre d'affiner ces résultats et de comparer l'expression des transcrits pendant le développement avec les quantités présentes dans les organes des poissons adultes. Les profils d'expression des gènes sont également très similaires, même si quelques différences sont à noter. Les bourgeons des nageoires pectorales expriment par exemple tous les gènes *cbx* à l'exception de *cbx6b*. Les protéines *Cbx* pourraient donc avoir un rôle spécifique dans la détermination et la régionalisation des cellules progénitrices, et notamment au niveau des membres. Tous les gènes *cbx* sont exprimés



dans les principaux organes des poissons adultes. Ces travaux initient la compréhension des fonctions des gènes *cbx* chez le poisson zèbre. La caractérisation fonctionnelle chez les mammifères a été amorcée mais les fonctions spécifiques dans un organe particulier ou un type cellulaire ne sont pas toujours claires. L'expression gènes *cbx* semblent pouvoir être utilisée comme marqueurs prédictifs le cancer du sein (Liang *et al.*, 2017). Les gènes sont surexprimés dans des types de cancers variés (Figure 22), mais les corrélations avec la survie des patients ne sont pas évidentes.



**Figure 22 : Expression relative des ARNm des 8 protéines CBX dans plusieurs types de cancers dans Oncomine.** Les surexpressions et gains de copies sont indiquées en rouges, les sous-expressions ou perte de copie en bleues. Les nombres correspondent au nombre d'études qui ont passé les critères de thresholds. Figure de Li *et al.*, 2020.

La sous-expression du gène *CBX2* est associée avec une moyenne de survie plus longue (Li *et al.*, 2020). Chez le poisson, le profil d'expression de *cbx2* est semblable à celui de *Ezh2*. Le transcrite est fortement exprimé lors du développement précoce puis diminue au cours du temps. Ces résultats tendent à montrer que *cbx2* est préférentiellement exprimé dans des cellules non-différenciée. De ce fait, la sous-expression de *CBX2* dans les cancers pourrait indiquer un certain état différencié des cellules cancéreuses et ainsi un pronostic plus favorable. Le poisson zèbre expriment chaque gène *Cbx* au stade adulte. Ainsi, l'évolution de l'expression de ces gènes après apparition de néoplasie pourra être évaluée afin de vérifier la valeur prédictive de ces gènes.

Enfin, nous avons identifié la conservation de plusieurs ohnologues dans la famille des *cbx* chez les téléostéens. Chez le poisson zèbre en particulier, *cbx6*, *cbx7* et *cbx8* sont conservés en pair d'ohnologues. Après les évènements de duplication du génome dans l'histoire évolutive des vertébrés



et des téléostéens, la majorité des copies ont été perdues au cours des générations suivantes. Les copies conservées présentent généralement un avantage sélectif. De ce fait, les gènes *cbx6a/b*, *cbx7a/b*, *cbx8a/b* pourrait avoir gagné de nouvelles fonctions chez le poisson zèbre. La présence de ces orthologues apparaît ainsi comme une limite dans l'utilisation du modèle poisson zèbre dans la compréhension des protéines Cbx et la transposabilité des connaissances chez les mammifères.

## 1.2 Le complexe PRC2 chez le poisson zèbre

### 1.2.1 Ezh1 et Ezh2

L'étude des séquences protéiques du domaine catalytique de Ezh2 et Ezh1 chez l'homme et chez plusieurs espèces de poissons permettent de retracer une partie de l'histoire évolutive de ces deux protéines. La sous-unité catalytique de PRC2 n'était représentée que par une seule protéine, E(z). L'homologie de séquence est plus importante entre E(z) et EZH2 (Laible *et al.*, 1997). L'apparition d'un deuxième orthologue, *ezh1*, chez les vertébrés à mâchoires suggère que *ezh1* soit le résultat d'une duplication de *ezh2*.

D'un point de vue phylogénie, la classe des chondrichthyens est divisée en deux sous-classes monophylétiques, les holocéphales et les élastombranchés (Inoue *et al.*, 2010). Cette théorie est un sujet de débat depuis plusieurs années. Le gène *ezh1* n'a pas été identifié dans la sous-classe des holocéphales dans nos analyses, mais est présent chez les élastombranchés et laisse penser que la classe des chondrichthyens soit un groupe polyphylétique. Une étude des hormones GnRH (gonadotropin releasing hormone) et CRF (corticotropin releasing factor) d'un point de vue phylogénique constate également des incohérences parmi les deux sous-classes concernant l'expression de ces hormones. L'une des hypothèses avancées est l'évolution particulière du téléencéphales des holocéphales et la pression de sélection qui a conduit à la perte de certains gènes (Lovejoy *et al.*, 2018). Des études complémentaires sont nécessaires pour déterminer avec précision l'histoire évolutive des chondrichthyens.

L'un de nos résultats les plus surprenants concerne la présence d'un transcrite *ezh1* d'origine maternelle dans les œufs des poissons zèbres. Ce résultat est validé par deux méthodes, par RT-PCR et par hybridation *in situ*, ce qui entre en conflit avec plusieurs données précédemment publiés sur le poisson zèbre. Dans ces études, aucun transcrite *ezh1* n'a été observé au stade zygotique. L'équipe de Chrispijn travaille sur la souche de poisson zèbre AB, tandis que la souche TU est utilisé au laboratoire. Des variations d'expression existent entre les différentes souches de poissons zèbres et peuvent être vraiment importantes selon les gènes (Holden & Brown, 2017). L'expression de *ezh1* au stade zygotique est très faible comparée aux autres stades de développement et à l'expression de *ezh2* par

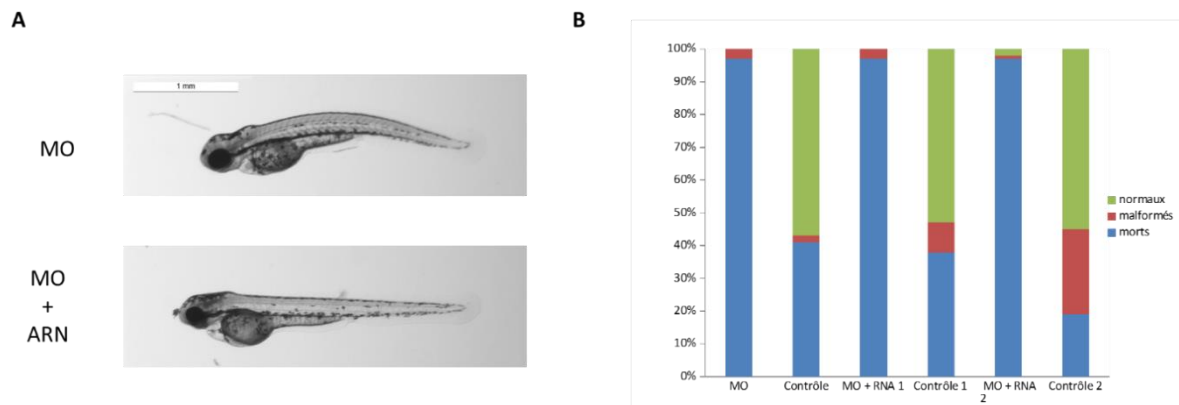
exemple. De ce fait, un transcrit *ezh1* pourrait être faiblement exprimé dans la souche TU et non exprimé dans la souche AB de par les variations d'expression entre les souches. Il existe des variations génétiques entre les différentes, mais également au sein d'une même souche. Le site [SNPfisher.nichd.nih.gov](http://SNPfisher.nichd.nih.gov) compile les variations reportées, à l'échelle du nucléotide, sur l'ensemble du génome (Butler *et al.*, 2015). Une recherche avec les coordonnées du gène *ezh1* (chr3 : 36,859,701-36,892,651) révèle plus de 200 mutations. Ce polymorphisme peut non seulement impacter l'expression du gène, mais également sa fonction. Toutefois, dans la souche TU ainsi que la souche AB, l'expression augmente au cours du développement, et est retrouvée principalement au sein de tissu différencié. Ces données confirment les observations sur le modèle murin (Margueron *et al.*, 2008). Le complexe PRC2 contenant la sous-unité Ezh1 semble intervenir dans les cellules différenciées afin de maintenir leur statut de différenciation. Les mutants *ezh1*<sup>-/-</sup> sont viables et confirme que le rôle de Ezh1 dans le développement n'est pas indispensable. L'expression de *ezh2* est augmentée chez les embryons mutants. Cette augmentation pourrait permettre de compenser la perte de *ezh1*, mais pourrait également expliquer un autre mécanisme. Le gène *ezh2*, à l'inverse de *ezh1*, est fortement exprimé au cours du développement précoce, puis voit son expression décroître dans les stades plus tardifs. Le rôle de Ezh2 est particulièrement important dans le développement précoce. Le KO de *Ezh2* chez la souris est létal au cours de la gastrulation (O'Carroll *et al.*, 2001) et autour de 12 jpf chez le poisson zèbre (Dupret *et al.*, 2017), ce qui suggère un rôle particulier dans les cellules indifférenciées, pour maintenir le statut souche ou engager les cellules dans une voie de différenciation. La maturation des cellules neurales au sein de l'hippocampe murin par exemple est dépendante de l'expression du gène *Psd-95*. Le KO de Ezh2 augmente l'expression de ce gène. A l'inverse, le KO de Ezh1 diminue l'expression de *Psd-95* (Henriquez *et al.*, 2013). Les complexes PRC2 contenant Ezh1 et Ezh2 semblent donc avoir une activité temporellement définie en fonction de la différenciation des cellules. Dans ce contexte, Ezh1 pourrait participer à la mise sous silence de *ezh2* afin d'entériner le statut différencié des cellules. Des boucles d'autorégulation ont par ailleurs été mises en évidence sur les protéines Cbx (Morey *et al.*, 2012) et JARID2 (Al-Raawi & Kanhere, 2020). Il pourrait être intéressant de vérifier la présence de la marque H3K27me3 et de Ezh1 sur le promoteur du gène *ezh2* dans des cellules matures. Ces données promeuvent une diversité de fonction de Ezh1 et Ezh2.

### 1.2.2 Eed

Les lignées mutantes KO pour les gènes *ezh1* et *ezh2* ont permis d'approfondir nos connaissances sur leur expression et leur rôle dans le développement précoce. La protéine Ezh1 compense très faiblement la perte de Ezh2, avec une réduction de la survie moyenne des embryons réduite de 2 jours. La protéine Ezh2 peut agir indépendamment du complexe PRC2. La protéine peut par exemple se lier

à des facteurs de transcription et favoriser l'expression de gènes, méthyler des protéines autres que les histones ou encore fixer des ARNnc (Huang *et al.*, 2021). L'utilisation des TALENs pour cibler le gène *ezh2* chez le poisson zèbre a conduit à l'insertion de 22 nucléotides dans le second exon de ce gène, ce qui forme un codon stop prématuré après le 60<sup>ème</sup> acide aminé. La protéine prédite est donc dépourvue de tous ses domaines protéiques (Dupret *et al.*, 2017). Ainsi, il est difficile de distinguer les fonctions de Ezh2 dépendante et indépendante du PRC2. Afin d'obtenir des informations sur les actions spécifiques du complexe PRC2, nous avons généré une nouvelle lignée mutante *eed*<sup>-/-</sup>. Les hétérozygotes sont viables et fertiles tandis que les mutants meurent autour de 11 jpf. Ce résultat est cohérent avec d'autres KO de protéines Polycomb chez le poisson zèbre (San *et al.*, 2016 ; Dupret *et al.*, 2017 ; Oppel *et al.*, 2020) mais en contraste avec ce qu'il se passe chez les mammifères (Faust *et al.*, 1995). Chez les souris, la perte de fonction de Eed, Ezh2 ou Suz12 est associée avec un arrêt du développement lors de la gastrulation. Nous avons observé qu'un transcrit *eed* d'origine maternelle est déposé dans les œufs. L'activation du génome zygotique varie beaucoup selon les espèces. Chez les souris, le génome zygotique est activé dès le stade deux cellules, 24 heures après la fécondation. Chez le poisson zèbre, la transcription du génome embryonnaire débute après 3 hpf au cours de la blastula (Jukam *et al.*, 2017). A ce stade, le rôle et les futures positions de certaines cellules embryonnaires ont d'ores et déjà été déterminées. Il semble donc que les transcrits d'origine maternelle chez le poisson zèbre possèdent une fonction plus importante que chez les souris. Il est donc possible que la déposition maternelle du transcrit *eed* masque des rôles de PRC2 dans le développement précoce. L'équipe de Rougeot a généré un mutant *ezh2* maternel zygotique et montre une absence de la marque H3K27me3 dès 48 hpf associés à des phénotypes variés tels qu'une réduction de la taille des yeux, l'absence de nageoire, ou une réduction de la taille de la tête des embryons (Rougeot *et al.*, 2019). Une approche avec des morpholinos a été essayée au laboratoire. Les morpholinos sont des petites séquences d'oligonucléotide antisens d'un ARNm et bloquent leur traduction (Nasevicius & Ekker, 2000). Deux morpholinos, dirigés contre le site d'initiation de la traduction et un site d'épissage de *eed* ont été designés, et ont été injectés dans les œufs. La co-injection avec un ARN *EED* humain doit permettre de sauver les phénotypes et d'assurer la spécificité des morpholinos. En plus d'une ponte de qualité désastreuse, les embryons co-injectés présentent des oedèmes cardiaques et des courbures de la colonne vertébrale à 72 hpf (Figure 23). L'injection des ARNs *EED* ne permet pas de sauver le phénotype et de vérifier la pertinence de nos résultats. Les morpholinos sont des courtes séquences d'une vingtaine de nucléotides. Les effets hors-cibles peuvent être nombreux et difficile à appréhender. Une nouvelle solution est actuellement envisagée pour aborder cette question des ARNm d'origine maternelle, avec l'utilisation de la technologie des CRISPR-cas13. Cette technologie cible et délète spécifiquement les ARNs par complémentarité de l'ARN guide

avec la séquence ciblée. Les CRISPR-cas13 ont été utilisés chez le poisson zèbre et diminuent de près de 80% le transcrit ciblé (Kushawah *et al.*, 2020).



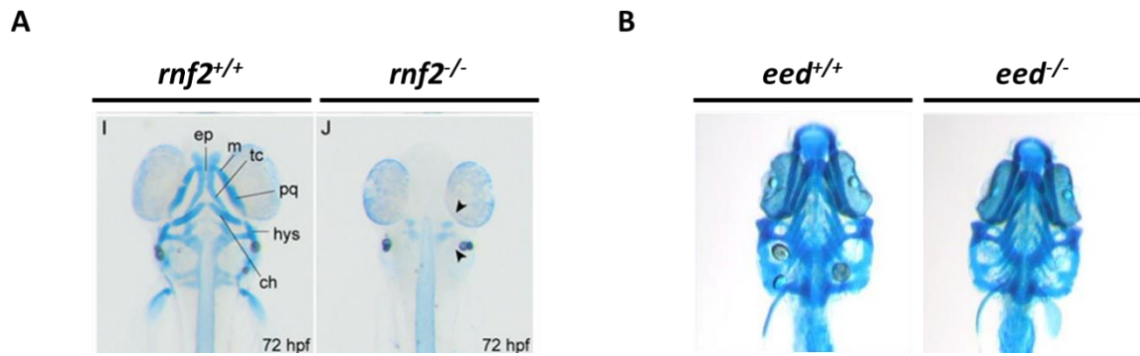
**Figure 23 : L'Injection de morpholinos chez le poisson zèbre.** A. Photographies d'embryons injectés à 72 hpf. La barre d'échelle représente 72 hpf. B. Représentation graphique des phénotypes observés en fonction de chaque condition. Le taux de mortalité de plus de 40% parmi les embryons contrôles témoigne d'une ponte de mauvaise qualité. Tous les embryons qui présentent un phénotype sont classés dans la catégorie « malformés » sans discrimination de sévérité.

Néanmoins, le plan d'organisation du corps chez les mutants *ezh2* maternels zygotiques est correctement mis en place à 24 hpf. Une étude montre d'ailleurs que la marque H3K27me3 n'est déposée qu'après l'activation du génome zygotique, soit après la transition de mid-blastula (Hickey *et al.*, 2022). A la différence des mammifères, le complexe PRC2 et la marque épigénétique H3K27me3 ne semblent pas requis pour les premiers stades de développement chez le poisson zèbre. Ces résultats pourraient marquer une spécialisation d'action de PRC2 au cours de l'évolution. De plus certaines protéines accessoires comme EPOP n'ont pas d'orthologues chez le poisson zèbre. Ces protéines pourraient avoir un rôle particulier dans le recrutement de PRC2 au cours du développement.

Chez les poissons zèbres, la taille est corrélée au développement des larves (Parichy *et al.*, 2009). La mesure de la taille est un critère relevant de mesure du développement sur les stades tardifs. Les larves mutantes sont significativement plus petites que les larves sauvages. La perte de fonction de *eed* retarde le développement général des larves. Chez l'homme des altérations du développement sont associées avec des mutations de *EED* (Cohen & Gibson, 2016). Les patients présentent un tableau clinique large, qui comprend par exemple un retard intellectuel. En revanche, les patients ont tous une croissance hors-norme, au-dessus des courbes de croissance normales. La perte de fonction de *Eed* ne semble donc pas affecter le développement de l'homme et du poisson zèbre de la même manière. Les mutations chez l'homme sont très rares et sont des substitutions ponctuelles dans les domaines WD40 n°3, 4 et 5. Il est donc très difficile de savoir dans quelle mesure ces mutations affectent la conformation de la protéine et l'activité du complexe PRC2.

**Le paysage épigénétique.** La déposition de plusieurs marques épigénétiques est modifiée par la perte de fonction de Eed. D'une part les marques H3K27me3/2 sont drastiquement réduites chez les larves *eed*<sup>-/-</sup> à 9 jpf. L'apport du transcrit d'origine maternelle semble là encore jouer un rôle important dans le maintien d'une activité minimale du PRC2 au cours du développement précoce. La majorité des transcrits sont dégradés après la transition mid-blastula. Certains ARNs sont conservés plus tardivement, par leurs interactions avec des protéines (Shi *et al.*, 2020) et leurs modifications post-transcriptionnelles (Yang *et al.*, 2019). Ces mécanismes sont peut-être à l'origine d'une préservation des transcrits *eed* au cours du développement. La marque H3K27me3 est une marque par ailleurs particulièrement stable dans le temps (Zee *et al.*, 2010). Les marques toujours visibles à 9 jpf pourraient avoir été précocement déposées et être maintenues dans quelques cellules au cours des divisions successives (Jadhav *et al.*, 2020). Dans la littérature, aucune autre méthyltransférase que EZH2/1 n'a été reportée capable de catalyser la di- et triméthylation de la lysine 27 de l'histone H3. Néanmoins cela pourrait suggérer que le complexe PRC2 garde une activité minimale en absence de Eed. Les domaines WD de Eed sont requis pour reconnaître la marque H3K27me3, recruter le complexe PRC2 sur ces marques, stimuler son activité et favoriser la propagation de la répression Polycomb (Margueron *et al.*, 2009). La perte de fonction de EED spécifiquement dans les cardiomyocytes chez les souris provoquent une diminution importante de la marque H3K27me3 sans l'abolir totalement (Ai *et al.*, 2017). Alors que les études structurales révèlent l'importance de la fixation de Eed au sein du complexe afin de modifier la conformation de Ezh2 et autoriser l'activité catalytique (Jiao *et al.*, 2015), il serait intéressant de vérifier *in vitro* si une activité minimale de PRC2 subsiste sans la présence de Eed chez le poisson zèbre.

La répression Polycomb est assurée par deux complexes, PRC1 et PRC2. L'activité des deux complexes est intrinsèquement liée. Les protéines CBX par exemple reconnaissent la marque H3K27me3 et recrutent le PRC1 au niveau de ces marques épigénétiques. Les larves mutantes *eed*<sup>-/-</sup> ont une réduction majeure de la marque H3K27me3, ce qui pourrait inhiber le recrutement du PRC1 et altérer la déposition de la marque H2AK119ub1. Pourtant, la déposition de cette marque n'est pas modifiée entre les larves mutantes et sauvages. Ceci explique pourquoi les mutants *eed* chez les poissons zèbres présentent des phénotypes si différents de ceux des mutants *rnf2*. La perte de fonction de *rnf2* provoque une létalité précoce autour de 4 jpf, avec des phénotypes sévères, comme des malformations cartilagineuses (van der Velden *et al.*, 2013). A 3 jpf, les larves *rnf2*<sup>-/-</sup> sont dépourvues de la structure cartilagineuse à la base de la formation des articulations de la mâchoire (Figure 24 A), ce qui indique un rôle prépondérant de PRC1 dans la mise en place du cartilage chez le poisson zèbre. Chez les mutants *eed*, les cartilages crâniofaciaux sont formés normalement (Figure 24 B).



**Figure 24 : Mise en place du cartilage chez les mutants Polycomb. A.** Coloration au bleu alcian de larves sauvages et *rnf2*<sup>-/-</sup> à 72 hpf. ch : ceratohyal ; ep : ethmoid plate ; hys : hyosymplectic ; m : Meckel's cartilage ; pc : parachordal ; pq palatoquadrate ; tc : trabeculae. Figure de van der Veldern *et al.*, 2013. **B.** Coloration au bleu alcian de larves sauvages et *eed*<sup>-/-</sup> à 9 dpf. Les cartilages sont correctement mis en place chez les larves mutantes.

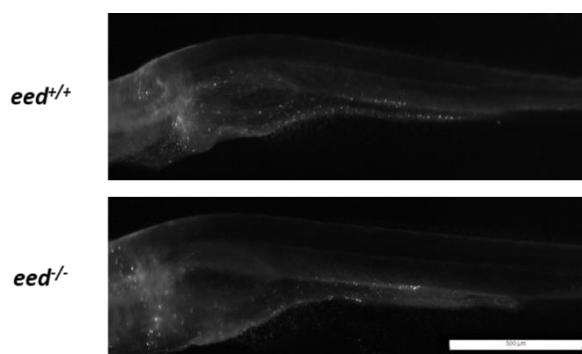
A cet égard, le complexe PRC1 et la marque H2AK119ub1 semblent être particulièrement importants dans le développement précoce des embryons. Les phénotypes sont plus sévères et apparaissent plus tôt chez les embryons mutants *rnf2*<sup>-/-</sup>. La marque H2AK119ub1 est d'ailleurs appliquée avant l'activation du génome zygotique, contrairement à la marque H3K27me3 (Hickey *et al.*, 2022). Notre modèle d'inactivation du PRC2 semble confirmer la temporalité d'action des complexes PRC1 et PRC2 chez le poisson zèbre : Le complexe PRC1 initie la répression Polycomb au cours du développement précoce, puis le complexe PRC2 affirme la répression de programmes géniques au sein de cellules préalablement spécifiées et maintient leur statut de différenciation. Néanmoins, si la quantité de H2AK119ub1 n'est pas modifiée chez les larves *eed*<sup>-/-</sup>, rien n'indique que ces marques soient distribuées de façon similaire chez les larves sauvages et mutantes. L'inhibition du recrutement par les protéines Cbx pourrait modifier le profil d'action des complexes PRC1c, remanier l'expression de certains gènes et ainsi participer à l'apparition de certains phénotypes. Il apparaît donc nécessaire de vérifier la distribution de la marque H2AK119ub1 dans le génome des larves *eed* mutantes.

La perte d'activité du complexe PRC2 impacte la déposition d'autres marques épigénétiques. La marque épigénétique répressive H4K20me3 est diminuée de près de moitié chez les larves mutantes. Les marques H3K27me3 et H4K20me3 sont enrichies au niveau de loci communs, dans des progéniteurs neuronaux. Elles assurent le contrôle transcriptionnel de gènes impliqués dans le développement et la différenciation cellulaire (Rhodes *et al.*, 2016). Chez les larves *eed*<sup>-/-</sup>, la marque activatrice H3K27ac est également enrichie. L'environnement chromatinien est donc largement modifié chez les larves mutantes. Notre modèle reproduit les changements épigénétiques observés dans certains cancers dans lesquels l'activité du PRC2 est réduite (Ueda *et al.*, 2016). Dans les gliomes pédiatriques de haut grade avec une mutation H3.3-K27M, la perte des marque H3K27me2 et H3K27me3 est associée avec une augmentation de la marque H3K27ac. L'acétylation de la lysine 27 augmente l'expression d'éléments répétitifs, ce qui augmente l'instabilité génomique des cellules

cancéreuses (Krug *et al.*, 2019). La perte de contrôle sur l'expression est également renforcée par une hypométhylation de l'ADN (Bender *et al.*, 2013). Il serait intéressant de vérifier si notre lignée mutante reproduit ces phénotypes observés dans les cancers.

**Les organes digestifs.** Les larves mutantes *eed*<sup>-/-</sup> et *ezh2*<sup>-/-</sup> ont des altérations du système digestif identiques. L'épaisseur de la paroi intestinale des larves est largement réduite et les villosités sont absentes. Le pancréas est également de taille réduite, dès 5 jpf. La similarité des phénotypes observés après délétion de *eed* et *ezh2* suggère que les fonctions de Ezh2 en dehors du complexe PRC2 sont minimes, voire inexistantes chez le poisson zèbre, au moins au cours du développement.

Les altérations intestinales sont potentiellement responsables de la létalité des larves autour de 11 jpf. A ce stade, les larves ont épuisé leurs réserves vitellines et leur survie dépend de leur capacité à s'alimenter et absorber les nutriments. Les articulations de la mâchoire sont correctement mises en place, les larves sont capables de s'alimenter (Figure 24). L'altération de la paroi et la réduction de la surface d'absorption intestinale diminue considérablement les capacités des larves à capter les nutriments nécessaires à leur survie. Ces organes sont correctement mis en place à 5 jpf, mais semblent ne pas être maintenus au cours du développement. La mortalité des cellules épithéliales intestinales n'est pas augmentée chez les larves mutantes en comparaison des sauvages (Figure 25). Chez la souris, un KO de *Eed* au sein de l'intestin révèle une diminution du nombre de cellules souches intestinales (Koppens *et al.*, 2016). Ainsi, l'altération intestinale des larves *eed*<sup>-/-</sup> est très certainement le résultat d'un non-renouvellement de l'épithélium intestinal.



**Figure 25 : Observation de l'apoptose par immunofluorescence.** Les cellules apoptotiques sont marquées à l'aide d'un anticorps dirigé contre la forme active de la caspase 3. La barre d'échelle représente 500  $\mu$ m.

Le pancréas des larves mutantes est de taille significativement réduite, dès 5 jpf. La formation du pancréas est dépendante de l'expression de nombreux facteurs. Le gène *neuroD* est exprimé dans le primordium pancréatique et marque les précurseurs endocriniens. Après 48 hpf, l'expression transitoire de *ngn3* est nécessaire pour la différenciation des cellules endocrines (Tiso *et al.*, 2009). Par hybridation *in situ*, nous avons montré que l'expression de certains gènes neurogéniques, tel que *neurod1*, est réduite chez les embryons *eed*<sup>-/-</sup>. Un suivi temporel de l'expression de ce gène par exemple pourrait mettre en évidence son implication dans la réduction de la taille du pancréas à 5 jpf.

Les colorations à l'Oil Red O, qui marquent les lipides neutres, révèlent que l'inactivation du PRC2 provoque une accumulation de ces lipides au niveau hépatique. La génération de mutants maternels chez la souris provoque une augmentation du stockage des acides gras (Prokopuk *et al.*, 2018). Chez le poisson zèbre, la perte de fonction de Ezh2 favorisent l'apparition de phénotypes similaires au niveau du foie (Dupret *et al.*, 2017 ; den Broeder *et al.*, 2020). La répression Polycomb médiée par le complexe PRC2 a donc un rôle particulier dans le métabolisme lipidique, et l'homéostasie hépatique en particulier. La stéatose hépatique rappelle une maladie humaine, la NAFLD (Non alcoholic fatty liver disease). Avec une mauvaise alimentation et hygiène de vie, le foie des patients stocke massivement les lipides. Ce stockage n'est pas sans effets puisqu'il altère la morphologie et les fonctions des hépatocytes. Le foie des larves *eed*<sup>-/-</sup> a été analysé sur coupes histologiques. Des grosses vacuoles vides sont visibles en grande quantité sur le foie des larves mutantes. L'inclusion en paraffine ne préserve pas les lipides, mais la présence de ces vacuoles corrélée à la coloration Oil Red O indique un foie stéatosé. De plus les structures que nous avons identifiées sont mises en évidence dans les modèles de poisson zèbre de NAFLD induite par une régime riche en cholestérol (Ma *et al.*, 2020). L'expression de plusieurs marqueurs a été quantifiée, mais ces résultats ne nous permettent pas de comprendre comment apparait ce phénotype et à quel niveau la répression Polycomb intervient. Pour compléter nos connaissances et en se penchant sur le développement de ces maladies du foie non-alcoolique, il semble que l'apparition de cette maladie soit multifactorielle. L'accumulation de lipides peut être le résultat d'une augmentation de la lipogenèse comme une diminution de la  $\beta$ -oxydation (Postic & Girard, 2008). D'autres mécanismes ont par ailleurs été mis en évidence dans l'apparition et l'évolution de cirrhose non-alcoolique (Schuppan *et al.*, 2017). Ces pathologies sont caractérisées par des successions de phases, inflammatoires qui permettent d'éliminer les tissus trop endommagés suivies de phases de fibroses pour réparer les tissus. Le système immunitaire régule ces deux phases par notamment la polarisation des macrophages. La répression Polycomb intervient à tous les stades de différenciation et maturation des cellules immunitaires (Di Carlo *et al.*, 2018). Ainsi, il sera intéressant d'approfondir l'impact de la perte d'activité du PRC2 sur le système immunitaire du poisson zèbre. Enfin, l'apparition d'une stéatose hépatique est corrélée avec un stress du réticulum endoplasmique (Vacaru *et al.*, 2014). La protéine ATF4, lorsqu'elle est surexprimée chez le poisson zèbre, participe à l'accumulation de lipide au niveau du foie via la surexpression de facteurs et d'enzymes lipogéniques, et la création d'un environnement inflammatoire (Yeh *et al.*, 2017). Les marqueurs de stress du réticulum endoplasmique et de la voie UPR (unfolded protein response) sont utilisés couramment aujourd'hui pour caractériser les stéatoses hépatiques et semblent être des cibles thérapeutiques encourageantes (Flamment *et al.*, 2010). L'inhibition de EZH2 dans les cellules cancéreuses active la voie de réponse au stress du réticulum endoplasmique et la production de stress (Casciello *et al.*, 2022) et pourrait expliquer une partie du phénotype des larves mutantes. Des analyses approfondies sont



nécessaires pour comprendre le rôle de la répression Polycomb dans le contrôle du métabolisme. Le gène *Ezh2* est surexprimé dans de nombreux cancers, et peut être utilisé comme marqueur pronostic (Takawa *et al.*, 2011). Ciblé le complexe PRC2 dans ces cancers afin d'altérer le métabolisme des cellules cancéreuses et les rendre vulnérables aux thérapies pourraient devenir une solution thérapeutique intéressante.

**Le développement neural.** Les larves *eed*<sup>-/-</sup> présentent une réduction drastique des marques H3K27me3/2, associé à des modifications d'autres marques épigénétiques. Ces altérations sont retrouvées dans certains types de cancers, et particulièrement dans les gliomes intrinsèques pontiques diffus porteurs de la mutation H3.3-K27M chez les enfants. Cette mutation particulière inhibe la triméthylation du variant d'histone muté et retient le complexe PRC2, ce qui conduit à une diminution globale de la marque H3K27me3 (Lewis *et al.*, 2013). Les cellules cancéreuses sont différenciées, ce qui les rend particulièrement agressives et résistantes aux traitements. La perte d'activité du PRC2 pourrait donc altérer la différenciation de ces cellules. Chez les larves *eed* mutantes, le nombre de cellules souches semble être augmenté, ce qui pourrait indiquer un blocage de la différenciation neurale. Chez le modèle murin, la perte de la protéine EED inhibe de la même façon la différenciation des progéniteurs neuraux (Obier *et al.*, 2015). Néanmoins, les différences au niveau des cellules différenciées sont plus fines. La neurogénèse est un processus complexe qui commence très tôt dans le développement. Certains facteurs dont l'expression est temporairement réduite, ne sont pas visibles à 5 jpf. L'implication du PRC2 dans le développement neural pourrait avoir lieu dans des stades plus précoces. Des analyses complémentaires au laboratoire vont être menées afin de répondre à cette question. Néanmoins, de la même façon que pour les organes du système digestif, le maintien des structures et du statut différencié des cellules neurales pourraient être à l'origine de l'apparition de ces phénotypes. Dans ce cadre, la quantification de l'expression de ces marqueurs au cours du développement des embryons pourrait être considérablement intéressante.

Les variations de profil d'expression des marqueurs neuraux entre les embryons *eed*<sup>-/-</sup> et *eed*<sup>+/+</sup> sont légères. Toutefois, l'expression de gènes comme *MAG*, *NEUROD*, ou *BMI* est également altérée dans les cellules cancéreuses de DIPGs (Silveira *et al.*, 2019 ; Lewis *et al.*, 2022). Notre lignée mutante pour le gène *eed* apparait donc comme un modèle de choix pour la compréhension de mécanismes moléculaires impliquées dans la tumorigenèse. L'équipe de Haratyunyan a publié récemment une analyse transcriptomique de plusieurs lignées cellulaires de DIPGs. La mutation K27M a été éditée par la technologie des CRISPRs, ce qui a permis d'analyser les variations d'expression génique associées à cette mutation. Certains de ces gènes sont impliqués dans le développement et la maturation du système nerveux. Nous avons voulu vérifier si l'expression de ces gènes était également modifiée dans le cerveau des embryons mutants, mais aucune différence significative n'a été révélée.

Notre lignée mutante *eed* semble donc être un modèle de choix pour représenter et étudier certains aspects des DIPGs mais ne parvient pas à résumer tous les aspects de la pathologie.

**La locomotion.** L'altération de la différenciation neurale chez les embryons mutants est corrélée avec une modification comportementale, une hyperactivité durant les stimuli lumineux. Les hypothèses qui peuvent expliquer ce changement de comportement sont nombreuses (Guo, 2003). Le gène *eed* est exprimé dans la rétine des embryons. La perte de fonction de *eed* provoque la néo-expression du gène *phox2bb* dans la rétine, ce qui peut modifier les capacités visuelles des embryons. A 5 jpf, les motoneurones primaires et secondaires sont impliqués dans la nage et la vitesse de nage des embryons en fonction du stimulus (Berg *et al.*, 2018). La maturation de ces neurones est primordiale pour le contrôle du mouvement. De nombreux acteurs, autant cellulaires que moléculaires peuvent donc être à l'origine de l'apparition de ce phénotype. Ce résultat nécessite une meilleure caractérisation de l'altération du développement neural chez les larves mutantes. Néanmoins, ce test mesure l'activité de 48 larves simultanément et est couramment utilisé pour tester l'effet de molécules neuroactives sur le comportement des embryons (Basnet *et al.*, 2019). Ce test de l'activité locomotrice pourrait nous permettre de tester à large échelle l'effet de molécules sur le phénotype et identifier des composés à potentiel thérapeutique.

## 2. Répression Polycomb et oncologie sur le modèle poisson zèbre : la modélisation de la mutation H3.3K27M des DIPGs

En France, chaque année près de 2000 enfants sont diagnostiqués pour un cancer. Les types de cancers sont variables et sont majoritairement de bon pronostic avec une survie moyenne à 5 ans supérieure à 80% (Lacour *et al.*, 2014). La survie des patients a augmenté ces dernières années, mais le traitement de certains cancers comme les leucémies myéloïdes et lymphoïdes ou certaines tumeurs du système nerveux central reste compliqué. Les tumeurs du système nerveux central représentent la majorité des cancers solides de l'enfant, et englobent une large diversité de cancer, sur les plans histologiques, génétiques et moléculaires. Cette diversité a conduit l'OMS à repenser sa classification des gliomes et prendre en compte les caractéristiques moléculaires de ces tumeurs (Louis *et al.*, 2016). La plupart des gliomes sont bénins et évoluent très peu, comme les astrocytomes pilocytiques. Il existe également des gliomes malins, dits de haut grade. Les gliomes de la ligne médiane, qui incluent les gliomes infiltrants du tronc cérébral (DIPGs) font toujours l'objet de nombreuses études. La survie moyenne est inférieure à 1 an, et malgré des essais cliniques prometteurs, aucun traitement ne s'est révélé efficace (Bailleul *et al.*, 2021). Les DIPGs sont associés à un profil moléculaire particulier. Près de 78% des DIPGs présentent une mutation récurrente dans les variants d'histone H3.3 ou H3.1 codés respectivement par les gènes *H3F3A* et *HIST13B* (Wu *et al.*, 2012). C'est une mutation faux-sens qui remplace la lysine 27 de ces variants d'histone par une méthionine. Cette mutation est un événement primaire dans l'initiation tumorale et la fréquence d'apparition de cette altération génique dans les DIPGs a amené l'Organisation Mondiale de la Santé à ajouter la mention « gliome diffus de la ligne médiane, K27M mutants » (Louis *et al.*, 2016). C'est une mutation gain de fonction qui inhibe l'activité du complexe PRC2 et retient la protéine EZH2 au niveau de la mutation (Lewis *et al.*, 2013). Le niveau de triméthylation globale de la lysine 27 de l'histone H3 est considérablement réduit *et altère* la répression transcriptionnelle de nombreux gènes. Ainsi des gènes impliqués dans l'auto-renouvellement de cellules souches ou dans la prolifération cellulaire sont fortement exprimés. Les mutations des variants H3.3 et H3.1 possèdent quelques spécificités. Les mutations H3.1-K27M ont un phénotype mésenchymateux/astrocytique tandis que les mutations H3.3-K27M sont associées à un profil proneural/oligodendrogial. Les deux mutations activent des programmes oncogéniques distincts. Les patients qui portent une mutation H3.3-K27M répondent faiblement à la radiothérapie et le temps de rechute est réduit (Castel *et al.*, 2015). Le caractère faiblement différencié des cellules cancéreuses promeut un caractère particulièrement agressif et une résistance aux thérapies existantes. L'âge médian du diagnostic des DIPGs est de 6 ans (Patil *et al.*; 2021). La maturation cérébrale et des cellules neurales semblent donc jouer un rôle clé dans l'initiation tumorale. Le

décryptage des mécanismes moléculaires mis en jeu dans la tumorigenèse pourraient offrir de nouvelles cibles thérapeutiques.

Dans ce but, la génération d'une nouvelle lignée de poisson zèbre, qui porte la mutation oncogénique humaine, a été initiée. La finalité de ce projet est de pouvoir exprimer et suivre l'expression de la mutation H3.3-K27M dans plusieurs types cellulaires neuraux.

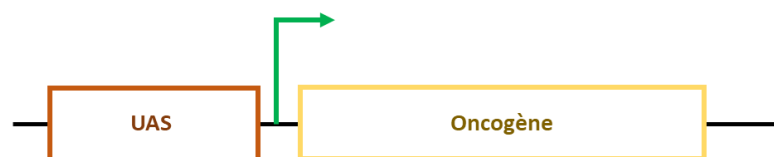
Pour générer cette nouvelle lignée de poisson zèbre, nous avons tiré profit du système GAL4/UAS (Scheer & Campos-Ortega, 1999). Ce système, initialement découvert chez la levure a été amélioré pour garantir un niveau d'expression conséquent, stable et spécifique d'un tissu ou un type cellulaire précis (Asakawa & Kawakami, 2008). Le facteur de transcription GAL4 est inséré sous le contrôle d'un promoteur spécifique d'un type cellulaire. Lorsque ce facteur de transcription est exprimé, il va se fixer sur les séquences activatrices UAS et activer l'expression de l'insert (

Figure 26).

### Construction activateur



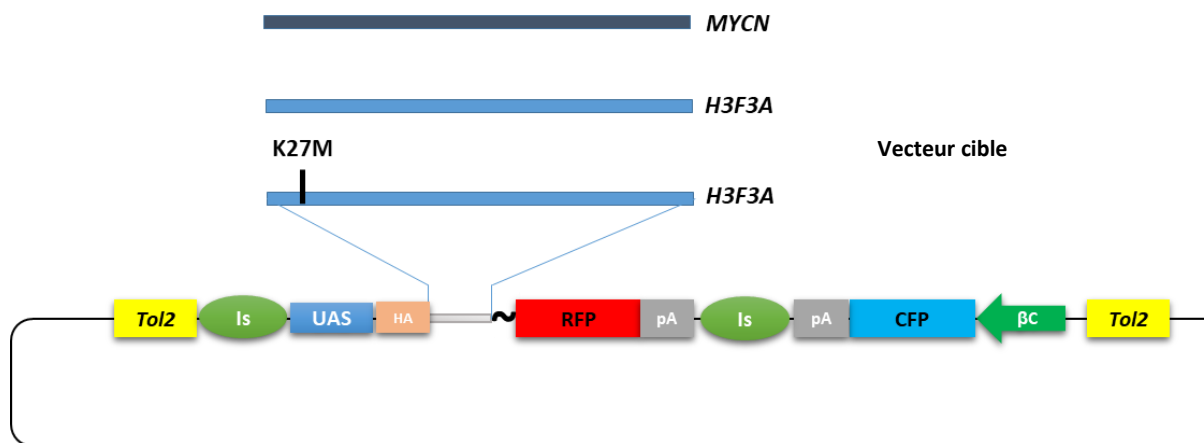
### Construction effecteur



**Figure 26 : Construction schématique du système GAL4/UAS.** Le facteur GAL4 est inséré sous le contrôle d'un promoteur spécifique de la lignée neurale. L'expression de GAL4 permet l'expression de l'oncogène uniquement dans cette lignée cellulaire. Figure adaptée de Scheer & Ortega, 1999.

En plaçant le variant humain H3F3A avec la mutation K27M sous le contrôle de l'UAS, ce système permet un contrôle spatio-temporel de l'expression de la mutation oncogénique. Nous avons donc construit des vecteurs plasmidiques pour faire cette construction. Sous le contrôle de l'UAS, nous avons inséré les gènes *H3F3A*, *H3F3A-K27M* et le gène *MYCN*. Le gène *H3F3A* sera notre lignée contrôle afin de vérifier que l'insertion du variant H3.3 humain ne soit source de problématique. Le gène *MYCN* est surexprimé dans les cellules tumorales (Silveira *et al.*, 2019). Les rôles de MYC dans la prolifération des cellules cancéreuses sont multiples (Dang *et al.*, 2009). Les gènes insérés sous le contrôle de l'UAS sont fusionnés à la RFP par un peptide auto-clivable. La co-expression de la RFP et des gènes d'intérêt permettra de suivre l'intégration des variants H3.3 dans la chromatine, et le développement tumoral. Une étiquette HA a également été ajoutée et sera utilisée pour identifier les régions génomiques dans lesquelles les variants H3.3 sont intégrés par ChIP-Seq. Ce vecteur sera injecté dans des œufs de poissons sauvage. Pour sélectionner les embryons qui ont intégré le vecteur, le gène codant pour la

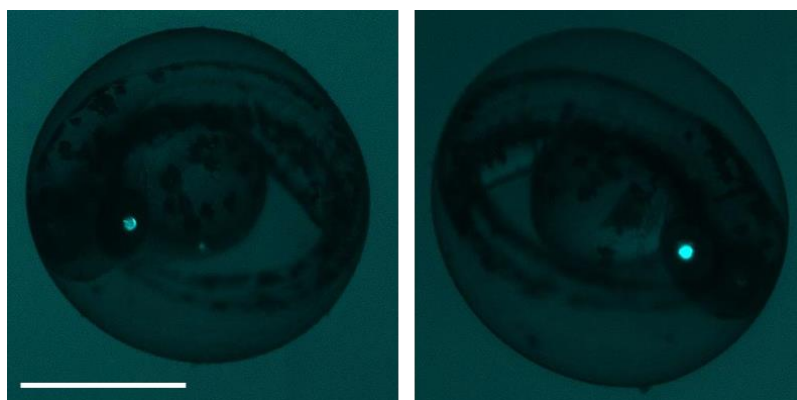
CFP est inséré sous le contrôle du promoteur de la  $\beta$ -cristalline. Les embryons qui ont intégré le vecteur effecteur sont facilement sélectionnés par fluorescence dans les yeux. L'ensemble de la construction est encadré par deux pieds de transposition, et la transposase Tol2 permettra d'intégrer la construction dans le génome des embryons (Fig 27). La transposase Tol2 crée des insertions aléatoires dans le génome des poissons, de manière très efficace (Suster *et al.*, 2009).



**Figure 27 : Construction du vecteur cible.** Tol2 : Pied de transposition. Is : Isolateur. UAS : Upstream Activating Sequence. HA : Hémagglutinine. RFP : Red Fluorescent Protein. CFP : Cyan Fluorescent Protein. BC : Promoteur de la Beta-cristalline.

Les trois vecteurs générés ont été vérifiés par séquençage puis injectés dans les œufs, au stade 1 cellule avec l'ARNm de la transposase Tol2. À partir de 2 jpf, les embryons ont été examinés sous loupe binoculaire à fluorescence et triés en fonction de la présence ou non de fluorescence au niveau des yeux des embryons (Figure 28).

Certains embryons ne possédaient qu'un œil fluorescent, ce qui indique un mosaïcisme important des embryons suite à la latence entre le temps d'injection et le temps d'intégration de la construction.

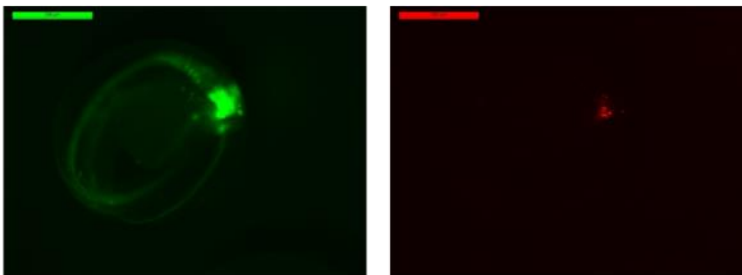


**Figure 28 : Embryon de 2 jpf après injection du vecteur contenant le variant H3.3 humain.** L'expression de la CFP dans le cristallin permet une sélection rapide des embryons. La barre d'échelle représente 500 $\mu$ m

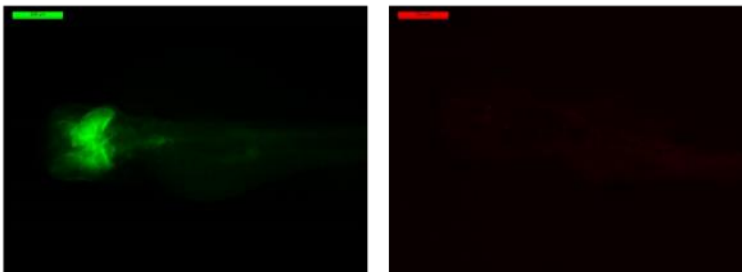
Après leur croissance, les poissons transgéniques ont été croisés afin de sélectionner les individus qui ont intégré les transgènes dans leur lignée germinale et sont capables de les transmettre à leur descendance. Les poissons ont également été croisés avec une lignée activatrice de poisson zèbre qui exprime GAL4 au niveau du système nerveux central, principalement au niveau optique. Le croisement de ces deux lignées active la transcription des transgènes et de la RFP. Parmi la descendance, quelques individus expriment la CFP, mais aucune fluorescence rouge n'est visible.

Afin de vérifier la validité de notre construction vectorielle, les vecteurs ont été injectés directement et de façon transitoire dans les embryons de poisson Tg(GAL4 ; UAS:GFP). L'expression de la RFP a été observée par fluorescence (Figure 29).

(A)



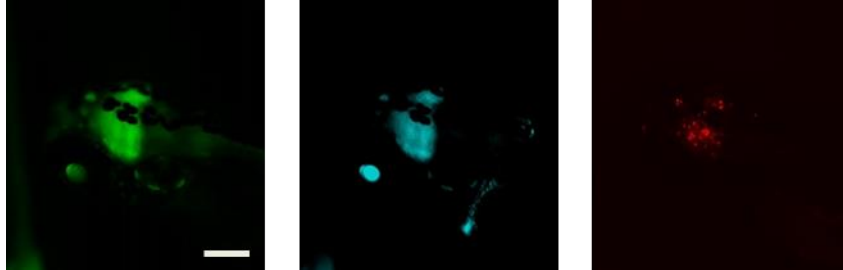
(B)



**Figure 29 : Expression transitoire du vecteur contenant *H3F3A*.** (A). Expression de GAL4 (vert) et du transgène *H3F3A* (rouge) 2 jours après l'injection. Les barres d'échelle représentent 500µm. (B) Expression de GAL4 (vert) 4 jours après l'injection. L'expression de *H3F3A* n'est plus visible. Les barres d'échelle représentent 200 µm.

Nous avons constaté l'expression de la RFP dès le 2<sup>ème</sup> jour après injection des vecteurs, et cette expression est restreinte au profil d'expression de GAL4 (Figure 29 A). Les vecteurs sont fonctionnels et autorisent l'expression des transgènes spécifiquement au niveau des cellules neurales du cerveau moyen des embryons. Seulement 4 jours après injection, l'expression des transgènes est fortement diminuée (Figure 29 B). L'expression de la CFP perdurent chez quelques individus, et pourraient être le résultat d'un évènement rare de recombinaison homologue et intégrerait seulement ce transgène dans le génome des poissons. Nos vecteurs sont donc fonctionnels mais n'ont probablement pas été intégrés dans les embryons. Une efficacité sous-optimale de la transposase Tol2 pourrait expliquer ce résultat.

Après production d'un nouvel ARNm *To2* et vérification de la qualité, de nouvelles lignées ont été générées, mais cette fois par injection dans les œufs de Tg(GAL4 ; UAS:GFP). L'expression de la RFP a été suivie au cours du développement pour vérifier que les transgènes restent exprimés activement et ne sont pas perdus (Figure 30).



**Figure 30** : Embryons Tg(GAL 4 ; UAS :GFP) 4 jours après injection du vecteur contenant le gène *MYCN*. L'expression de GAL 4 (vert), de la CFP dans le cristallin (bleu) et de MYCN (rouge) sont toujours visibles 4 jours après l'injection. LA barre d'échelle représente 200  $\mu$ m.

La présence de la fluorescence bleue au niveau du cristallin des embryons ainsi qu'un nombre important de cellules fluorescentes rouges 4 jours après l'injection indique une bonne intégration des transgènes dans le génome des embryons. Ces embryons ont été sélectionnés et ont été croisés après leur croissance afin d'identifier les lignées transgéniques fondatrices. Dans la génération F2, l'expression de la RFP est visible, mais disparaît après 5 jours de développement.

Si la mutation oncogénique H3.3-K27M est initiatrice du développement tumoral, la tumorigénèse nécessite des événements mutationnels supplémentaires. Plus de la moitié des patients avec la mutation H3.3-K27M ont également une mutation dans le gène suppresseur de tumeur *TP53* (Gröbner *et al.*, 2018). Un croisement supplémentaire avec une lignée de fond génétique *tp53<sup>nué</sup>* sera nécessaire pour initier la transformation des cellules neurales en cellules cancéreuses. Néanmoins la disparition de la RFP dans la lignée F2 reste inexplicée. Ces cellules pourraient être reconnues comme anormales par le système immunitaire et avoir été éliminées. Le suivi de la microglie par marquage fluorescent pourra nous permettre de vérifier ce point.

Ces résultats préliminaires sur ces nouvelles lignées transgéniques sont prometteurs. La co-expression de la RFP permettra de suivre le développement tumoral très précisément sans procédures invasives pour les poissons. Les tumeurs seront analysées par histologie et comparées aux données histopathologiques disponibles. Les cellules tumorales pourront être prélevées, triées et transplantées par greffe orthotopiques afin de vérifier leur capacité tumorigène. La caractérisation de ce modèle de DIPGs chez le poisson zèbre pourrait permettre de décrypter les mécanismes à l'origine du développement tumoral. Il sera également possible de tester l'effet de molécules à visée

thérapeutiques et leurs effets secondaires sur le développement et le comportement du poisson zèbre.



# CONCLUSIONS

---

La répression Polycomb est un mécanisme épigénétique conservé chez de nombreuses espèces. Les protéines de la famille Polycomb du poisson zèbre présentent un degré d'homologie très élevées avec celles des mammifères. Nous avons montré que tous les gènes de la famille *CBX* chez l'homme ont tous au moins un orthologues chez le poisson zèbre. Ces gènes sont exprimés au cours du développement et au stade adulte. Le poisson zèbre pourra être utilisé dans l'étude fonctionnelle des protéines *cbx*.

L'étude approfondie du complexe PRC2 par la caractérisation des lignées mutantes *ezh1<sup>ul3</sup>* et *eed<sup>ul4</sup>* a démontré plusieurs points importants. Les fonctions de *ezh1* et *ezh2* ne sont que faiblement redondantes et les deux protéines possèdent des rôles propres. Le complexe PRC2 n'est pas indispensable au développement précoce du poisson zèbre, mais apparaît particulièrement important dans le maintien des structures au cours du temps. Les modifications de l'environnement chromatinien et la dérégulation de certains gènes sont similaires à aux altérations observées dans les cancers, en cas de perte d'activité du complexe PRC2. La lignée mutante *eed<sup>ul4</sup>* pourra ainsi être utilisée d'étudier certains aspects de ces cancers.

Le poisson zèbre est utilisé en recherche dans de nombreux domaines, et de plus en plus fréquemment en cancérologie. Les expériences de transplantation de cellules cancéreuses permettent d'aborder plusieurs problématiques telles que la dissémination métastatique ou le test de molécules. De plus, l'essor des techniques de séquençage et d'édition du génome chez le poisson zèbre est particulièrement intéressant pour la recherche sur le cancer. Il est désormais possible d'insérer une mutation précise dans un oncogène ou un gène suppresseur de tumeur. Dans ce cadre, un nouveau modèle transgénique portant la mutation H3.3-K27M a été généré avec succès au laboratoire et permettra l'étude des gliomes de haut grade. Ce modèle interrogera le rôle de la répression Polycomb dans le cancer, sur le modèle poisson zèbre.



## RÉFÉRENCES :

---

- Abe S, Yamashita SI, Miyahara S, Wakahara J, Yamamoto L, Mori R, Imamura N, Yoshida Y, Waseda R, Hiratsuka M, Shiraishi T, Nabeshima K, Iwasaki A. Pronotic Significance of BMI1 But Not MEL18 Expression in Pulmonary Squamous Cell Carcinoma. *Anticancer Research*. 2017; 37(4): 1923-1929.
- Abel KJ, Brody LC, Valdes JM, Erdos MR, McKinley DR, Castilla LH, Merajver SD, Couch FJ, Friedman LS, Ostermeyer EA, Lynch ED, King M-C, Welcsh PL, Osborne-Lawrence S, Spillman M, Bowcock AM, Collins FS, Weber BL. Characterization of *EZH1*, a Human Homolog of *Drosophila Enhancer of zeste* near *BRCA1*. *Genomics*. 1996; 37: 161-171.
- Adriaenssens E, Lottin S, Dugimont T, Fauquette W, Coll J, Dupouy JP, Boilly B, Curgy JJ. Steroid hormones modulate H19 gene expression in both mammary gland and uterus. *Oncogene*. 1999; 18, 4460–4473.
- Ai S, Peng Y, Li C, Gu F, Yu X, Yue Y, Ma Q, Chen J, Lin Z, Zhou P, Xie H, Prendiville TW, Zheng W, Liu Y, Orkin SH, Wang DZ, Yu J, Pu WTP, He A. EED orchestration of heart maturation through interaction with HDACs is H3K27me3-independent. *eLife*. 2017; 6: e24570.
- Ai S, Yu X, Li Y, Peng Y, Li C, Yue Y, Tao G, Li C, Pu WT, He A. Divergent Requirements for EZH1 in Heart Development Versus Regeneration. *Circulation Research*. 2017; 121(2): 106-112.
- Akasaka T, Kanno M, Balling R, Mieza MA, Taniguchi M, Koseki H. A role for mel-18, a Polycomb group-related vertebrae gene, during the anteroposterior specification of the axial skeleton. *Development*. 1996; 122(5): 1513-22.
- Akasaka T, Tsuji K, Kawahira H, Kanno M, Harigaya K, Hu L, Ebihara Y, Nakahata T, Tetsu O, Taniguchi M. The role of mel-18, a mammalian Polycomb group gene, during IL-7-dependent proliferation of lymphocyte precursors. *Immunity*. 1997; 7: 135–146.
- Alekseyenko AA, Gorchakov AA, Kharchenko PV, Kuroda MI. Reciprocal interactions of human C10orf12 and C17orf96 with PRC2 revealed by BioTAP-XL cross-linking and affinity purification. *PNAS*. 2014; 111(7): 2488-2493.
- Alkema MJ, Wlegant J, Raap AK, Berns A, van Lohuizen. Characterization and chromosomal localization of the human proto-oncogene BMI-1. *Human Molecular Genetics*. 1993; 2(10): 1597-1603.
- Almeida M, Bowness JS, Brockdorff N. The many faces of Polycomb regulation by RNA. *Current Opinion in Genetics & Development*. 2020; 61: 53-61.
- Almeida M, Pintacuda G, Masui O, Koseki Y, Gdula M, Cerase A, Brown D, Mould A, Innocent C, Nakayama M, Schermelleh L, Nesterova T, Koseki H, Brockdorff N. PCGF3/5 initiates Polycomb recruitment in X chromosome inactivation. *Science*. 2017; 356(6342): 1081-1084.
- Al-Raawi D, Kanhere A. Autoregulation of JARID2 through PRC2 interaction with its antisense ncRNA. *BMC Research Notes*. 2020; 13(1): 501.
- Amores A, Force A, Yan YL, Joly L, Amemiya C, Fritz A, Ho RK, Langeland J, Prince V, Wang YL, Westerfield M, Ekker M, Postlethwait JH. Zebrafish hox clusters and vertebrate genome evolution. *Science*. 1998; 282:1711–1714.

- Anderson RM, Bosch JA, Goll MG, Hesselson D, Dong PDS, Shin D, Chi NC, Shin CH, Schlegel A, Halpern M, Stainier DYR. Loss of Dnmt1 catalytic activity reveals multiple roles for DNA methylation during pancreas development and regeneration. *Developmental Biology*. 2009; 334(1): 213-223.
- Andre M, Ando S, Ballagny C, Durliat M, Poupard G, Briancon C, Babin PJ. Intestinal fatty acid binding protein gene expression reveals the cephalocaudal patterning during zebrafish gut morphogenesis. *International Journal of Developmental Biology*. 2000; 44: 249–252.
- Angrand PO, Venni C, Le Bourhis X, Adriaenssens E. The role of long non-coding RNAs in genome formatting and expression. *Frontiers Genetics*. 2015; 6: e165.
- Arnold P, Schöler A, Pachkov M, Balwierz PJ, Jorgensen H, Stadler MB, van Nimwegen E, Schübeler D. Modeling of epigenome dynamics identifies transcription factors that mediate Polycomb targeting. *Genome Research*. 2013; 23: 60-73.
- Arrigoni R, Alam SL, Wamstad JA, Bardwell VJ, Sundquist WI, Schreiber-Agus N. The Polycomb-associated protein Rybp is a ubiquitin binding protein. *FEBS Letters*. 2006; 580(26): 6233-41.
- Asakawa K, Kawakami K. Targeted gene expression by the Gal4-UAS system in zebrafish. *Development, Growth & Differentiation*. 2008; 50(6): 391-399.
- Astolfi A, Fiore M, Melchionda F, Indio V, Bertuccio AN, Pession A. BCOR involvement in cancer. *Epigenomics*. 2019; 11(7): 835-855.
- Astolfi A, Fiore M, Melchionda F, Indio V, Bertuccio SN, Pession A. BCOR involvement in cancer. *Epigenomics*. 2019; 11(7): 835-855.
- Bae WK, Kang K, Yu JH, Yoo KH, Factor VM, Kaji K, Matter M, Thorgeirsson S, Hennighausen L. The methyltransferases enhancer of zeste homolog (EZH)1 and EZH2 control hepatocyte homeostasis and regeneration. *The FASEB Journal*. 2015; 29(5): 1653-62.
- Bae WK, Kang K, Yu JH, Yoo KH, Factor VM, Kaji K, Matter M, Thorgeirsson S, Hennighausen L. The methyltransferases enhancer of zeste homolog (EZH)1 and EZH2 control hepatocyte homeostasis and regeneration. *FASEB Journal*. 2015; 29(5): 1653-62.
- Bailleul Q, Rakotomalala A, Ferry I, Leblond P, Meignan S, Furlan A. The art of war as applied to pediatric gliomas: Know your enemy. *Médecine Sciences*. 2021; 37(2): 159-166.
- Balasubramanian S, Raghunath A, Perumal E. Role of epigenetics in zebrafish development. *Gene*. 2019; 718: 144049.
- Ballaré C, Lange M, Lapinaite A, Mas Martin G, Morey L, Pascual G, Liefke R, Simon B, Shi Y, Gozani O, Carlomagno T, Benotah SA, Di Croce L. Phf19 links methylated Lys36 of histone H3 to regulation of Polycomb activity. *Nature Structural & Molecular Biology*. 2012; 19: 1257-1265.
- Bally-Cuif L, Hammerschmidt M. Induction and patterning of neuronal development, and its connection to cell cycle control. *Current Opinion*. 2003; 13(1): 16-25.
- Barman RP. A taxonomic revision of the Indo-Burmese species of *Danio rerio*. *Record of the Zoological Survey of India Occasional Papers*. 1991; 10: 1351-1358.
- Basnet RM, Zizioli D, Taweedet S, Finazzi D, Memo M. Zebrafish Larvae as a Behavioral Model in Neuropharmacology. *Biomedicines*. 2019; 7:23.

- Bauer M, Trupke J, Ringrose L. The quest for mammalian Polycomb response elements: are we there yet? *Chromosoma*. 2016; 125: 471-496.
- Bender S, Tang Y, Lindroth AM, Hovestadt V, Jones DTW, Kool M, Zapatka M, Northcott PA, Sturm D, Wang W, Radlwimmer B, Højfeldt JW, Truffaux N, Castel D, Schubert S, Ryzhova M, Seker-Cin H, Gronych J, Johann PD, Stark S, Meyer J, Milde T, Schuhmann M, Ebinger M, Monoranu CM, Ponnuswami A, Chen S, Jones C, Witt O, Collins VP, von Deimling A, Jabado N, Puget S, Grill J, Helin K, Korshunov A, Lichter P, Monje M, Plass C, Cho YJ, Pfister SM. Reduced H3K27me3 and DNA Hypomethylation Are Major Drivers of Gene Expression in K27M Mutant Pediatric High-Grade Gliomas. *Cancer Cell*. 2013; 24(5): 660-672.
- Benveniste D, Sonntag HJ, Sanguinetti G, Sproul D. Transcription factor binding predicts histone modifications in human cell lines. *Proceedings of the National Academy of Sciences USA*. 2014; 111(37): 13367-13372.
- Berg EM, Björnfors ER, Pallucchi I, Picton LD, El Manira A. Principles Governing Locomotion in Vertebrates: Lessons From Zebrafish. *Frontiers Neural Circuits*. 2018; 12.
- Beringer M, Pisano P, Di Carlo V, Blanco E, Chammas P, Vizán P, Gutiérrez A, Aranda S, Payer B, Wierer M, Di Croce L. EPOF Functionally Links Elongin and Polycomb in Pluripotent Stem Cells. *Molecular Cell*. 2016; 64(4): 645-658.
- Bernstein BE, Mikkelsen TS, Xie X, Kamal M, Huebert DJ, Cuff J, Fry B, Meissner A, Wernig M, Plath K, Jaenisch R, Wagschal A, Feil R, Schreiber SL, Lander ES. A Bivalent Chromatin Structure Marks Key Developmental Genes in Embryonic Stem Cells. *Cell*. 2006; 125: 315-326.
- Bernstein E, Duncan EM, Masui O, Gil J, Heard E, Allis CD. Mouse Polycomb Proteins Bind Differentially to Methylated Histone H3 and RNA and Are Enriched in Facultative Heterochromatin. *Molecular and Cellular Biology*. 2006; 26(7): 2560-9.
- Betancur RR, Wiley EO, Arratia G, Acero A, Bailly N, Miya M, Lecointre G, Ortí. Phylogenetic classification of bony fishes. *BMC Evolutionary Biology*. 2017; 17: 162.
- Bian F, Gao F, Kartashov AV, Jegga AG, Barski A, Das SK. Polycomb repressive complex 1 controls uterine decidualization. *Scientific Reports*. 2016; 6: 2606.
- Biemar F, Argenton F, Schmidtke R, Epperlein S, Peers B, Driever W. Dispersed Appearance of Endocrine Hormone Expressing Cells and their Convergence to Form the Definitive Islet. *Developmental Biology*. 2001; 230(2): 189-203.
- Blackledge NP, Farcas AM, Kondo T, King HW, McGouran JF, Hanssen LLP; Ito S, Cooper S, Kondo K, Koseki Y, Ishikura T, Long HK, Sheahan TW, Brockdorff N, Kessler BM, Koseki H, Klose RJ. Variant PRC1 complex-dependant H2A ubiquitylation drives PRC2 recruitment and polycomb domain formation. *Cell*. 2014; 157(6):1445-1459.
- Blader P, Fischer N, Gradwohl G, Guillemont F, Strahle U. The activity of neurogenin1 is controlled by local cues in the zebrafish embryo. *Development*. 1997; 124(22): 4557-4569.
- Boch J, Scholze H, Schornack S, Landgraf A, Hahn S, Kay S, Lahaye T, Nickstadt A, Bonas U. Breaking the code of DNA binding specificity of TAL-type III effectors. *Science* 2009; 326: 1509–1512.
- Bodega B, Marasca F, Ranzani V, Cherubini A, Della Valle F, Neguembor MV, Wassef M, Zippo A, Lanzuolo C, Pagani M, Orlando V. A cytosolic Ezh1 isoform modulates a PRC2-Ezh1 epigenetic

- adaptive response in postmitotic cells. *Nature Structural & Molecular Biology*. 2017; 24(5): 435-437.
- Bordeira-Carriço R, Teixeira J, Duque M, Glhardo M, Ribeiro D, Acemel RD, Firbas PN, Tena JJ, Eufrásio A, Marques J, Ferreira FJ, Freitas T, Carneiro F, Gómez-Skarmeta JL, Bessa J. Multidimensional chromatin profiling of zebrafish pancreas to uncover and investigate disease-relevant enhancers. *Nature Communications*. 2022; 13: 1945.
- Boros J, Arnoult N, Stroobant V, Collet JF, Decottignies A. Polycomb Repressive Complex 2 and H3K27me3 Cooperate with H3K9 Methylation To Maintain Heterochromatin Protein 1 $\alpha$  at Chromatin. *Molecular and Cellular Biology*. 2014; 34(19): 3662-3674.
- Boulard M, Edwards JR, Bestor TH. Abnormal X chromosome inactivation and sex-specific gene dysregulation after ablation of FBXL10. *Epigenetics Chromatin*. 2016; 9:22.
- Bousard A, Raposo AC, Zyllicz JJ, Picard C, Borges Pires V, Qi Y, Gil C, Syx L, Chang HY, Heard E, da Rocha ST. The role of Xist-mediated Polycomb recruitment in the initiation of X-chromosome inactivation. *EMBO Reports*. 2019; 20:e48019.
- Boyer LA, Plath K, Zeitlinger J, Brambrink T, Medeiros LA, Ihn Lee T, Levine SS, Wernig M, Tajonar A, Ray MK, Bell GW, Otte AP, Vidal M, Gifford DK, Young RA, Jaenisch R. Polycomb complexes repress developmental regulators in murine embryonic stem cells. *Nature*. 2006; 441: 349-353.
- Brannan CI, Dees EC, Ingram RS, Tilghman SM. The Product of the H19 Gene May Function as an RNA. *Molecular and Cellular Biology*. 1990; 10: 28–36.
- Brunet FG, Crollius HR, Paris M, Aury JM, Gibert P, Jaillon O, Laudet V, Robinson-Rechavi M. Gene loss and evolutionary rates following whole-genome duplication in teleost fishes. *Molecular Biology and Evolution*. 2006; 23(9): 1808-16.
- Butler MG, Iben JR, Marsden KC, Epstein JA, Granato M, Weinstein BM. SNPfisher:tools for probing genetic variation in laboratory-reared zebrafish. *Development*. 2015; 142(8): 1542-1552.
- Caganova MC, Carrisi C, Varano G, Mainoldi F, Zanardi F, Germain PL, George L, Alberghini F, Ferrarini L, Talukder AK, Ponzoni M, Testa G, Nojima T, Doglioni C, Daisuke Kitamura D, Toellner KM, Su I, Casola S. Germinal center dysregulation by histone methyltransferase EZH2 promotes lymphomagenesis. *Journal of Clinical Investigation*. 2013 ; 123(12) : 5009-22.
- Cai L, Rothbart SB, Lu R, Xu B, Chen WY, Tripathy A, Shira R, Zheng D, Patel DJ, Allis D, Strahl BD, Song J, Wang GG. Polycomb-like Proteins Mediates PRC2 Complex Targeting. *Molecular Cell*. 2013; 49(3):571-582.
- Cai X, Cullen BR. The imprinted H19 noncoding RNA is a primary microRNA precursor. *RNA*. 2007; 13: 313–316.
- Calebiro D, Grassi ES, Eszlinger M, Ronchi CL, Godbole A, Bathon K, Guizzardi F, de Filippis T, Krohn K, Jaeschke H, Schwarzmayr T, Bircan R, Gozu HI, Sancak S, Niedziela M, Strom TM, Fassnacht M, Persani L, Paschke R. Recurrent EZH1 mutations are a second hit in autonomous thyroid adenomas. *Journal of Clinical Investigation*. 2016 Sep 1;126(9):3383-8.
- Calés C, Pavón L, Starowicz K, Pérez C, Bravo M, Ikawa T, Koseki H, Vidal M. Role of Polycomb RYBP in Maintaining the B-1-to-B-2 B-Cell Lineage Switch in Adult Hematopoiesis. *Molecular & Cellular Biology*. 2016; 36: 900–912.

- Carrington EA, Jones RS. The *Drosophila Enhancer of zeste* gene encodes a chromosomal protein: examination of wild-type and mutant protein distribution. *Development*. 1996; 122: 4073-4093.
- Casciello F, Kellu GM, Ramarao-Milne P, Kamal N, Stewart TA, Mukhopadhyay P, Hazakoff SH, Mjranda M, Kim D, Davis FM, Hayward NK, Vertino PM, Waddell N, Gannon F, Lee JS. Combined Inhibition of G9a and EZH2 Suppresses Tumor Growth via Synergistic Induction of IL24-Mediated Apoptosis. *Cancer Research*. 2022; 82(7): 1208-1221.
- Castel D, Philippe C, Calmon R, Le Dret L, Truffaux N, Boddaert N, Pagès M, Taylor KR, Saulnier P, Lacroix L, Mackay A, Jones C, Sainte-Rose C, Blauwblomme T, Andreiuolo F, Puget S, Grill J, Varlet P, Debily MA. Histone H3F3A and HIST1H3B K27M mutations define two subgroups of diffuse intrinsic pontine gliomas with different prognosis and phenotypes. *Acta Neuropathology*. 2015; 130(6): 815-27.
- Chan CS, Rastelli L, Pirrota V. A Polycomb response element in the *ubx* gene that determines an epigenetically inherited state of repression. *EMBO Journal*. 1994; 13(11):2553-64.
- Chapouton P, Godinho L. Chapter 4 - Neurogenesis. In: Dietrich HW, Westerfield M, Zon LI. *Methods in Cell Biology*. Elsevier. 2010; vol 100 : pp. 72–126.
- Chen H, Zhou J, Chen H, Liang J, Xie C, Gu X, Wang R, Mao Z, Zhang Y, Li Q, Zuo G, Miao D, Jin J. Bmi-1-RING1B prevents GATA4-dependent senescence-associated pathological cardiac hypertrophy by promoting autophagic degradation of GATA4. *Clinical & Translational Medicines*. 2022; 12(4): e574.
- Chen MJ, Deng J, Chen C, Hu W, Yuan YC, Xia ZK. LncRNA H19 promotes epithelial mesenchymal transition and metastasis of oesophageal cancer via STAT3/EZH2 axis. *International Journal of Biochemistry & Cell Biology*. 2019; 113: 27-36.
- Chen S, Jiao L, Shubbar M, Yang X, Liu X. Unique Structural Platforms of Suz12 Dictate Distinct Classes of PRC2 for chromatin Binding. *Molecular Cell*. 2018; 69(5): 840-852.
- Chrispijn ND, Andralojc KM, Castenmiller C, Kamminga LM. Gene expression profile of a selection of Polycomb Group genes during zebrafish embryonic and germ line development. *PLOS ONE*. 2018; 13: e02003316.
- Chrispijn ND, Elurbe DM, Mickoleit M, Aben M, de Bakker DEM, Andralojc KM, Huisken J, Bakkers J, Kamminga LM. Loss of the Polycomb group protein Rnf2 results in derepression of *tbx*-transcription factors and defects in embryonic and cardiac development. *Scientific Reports*. 2019; 9: 4327.
- Chu J, Sadler KC. New school in liver development: Lessons from zebrafish. *Hepatology*. 50(5): 1656-1663.
- Ciferri C, Lander GC, Maiolica A, Herzog F, Aebersold R, Nogales E. Molecular architecture of human polycomb repressive complex 2. *Structural Biology and Molecular Biophysics*. 2012; 1: e00005.
- Ciferri C, Lander GC, Maiolica A, Herzog F, Aebersold R, Nogales E. Molecular architecture of human polycomb repressive complex 2. *Elife*. 2012; 1: e00005.
- Cifuentes-Rojas C, Hernandez AJ, Sarma K, Lee JT. Regulatory interactions between RNA and polycomb repressive complex 2. *Molecular Cell*. 2014; 55: 171-185.

- Clarke J. Role of polarized cell divisions in zebrafish neural tube formation. *Current Opinion Neurobiology*. 2009; 19: 134-138.
- Cohen ASA, Gibson WT. EED-associated overgrowth in a second male patient. *Journal of Human Genetics*. 2016; 61: 831-4.
- Connelly KE, Dykhuizen EC. Compositional and functional diversity of canonical PRC1 complexes in mammals. *Biochimica et Biophysica Acta*. 2017; 1860(2): 233-245.
- Conway E, Jerman E, Healy E, Ito S, Holoch D, Oliviero G, Deevy O, Glancy E, Fitzpatrick DJ, Mucha M, Watson A, Rice AM, Chammas P, Huang C, Pratt-Kelly I, Koseki Y, Nakayama M, Ishikura T, Streubel G, Wynne K, Holamp K, McLysagt A, Ciferri C, Di Croce L, Cagney G, Margueron R, Koseki H, Bracken AP. A family of Vertebrate-Specific Polycombs Encoded by the LCOR-LCORL Genes Balance PRC2 Subtype Activities. *Molecular Cell*. 2018; 70(3):408-421.
- Conway EM, Bracken AP. "Chapter 4 : Unraveling the Roles of Canonical and Noncanonical PRC1 Complexes" *Polycomb Group Proteins*. Pirrotta V. 2017; *Academic Press*, 57-80.
- Coolen M, Thieffry D, Drivenes O, Becker TS, Bally-Cuif L: miR-9 controls the timing of neurogenesis through the direct inhibition of antagonistic factors. *Developmental Cell*. 2012, 22: 1052-1064.
- Coré N, Bel S Gaunt SJ, Aurrand-Lions M, Pearce J, Fisher A, Djabali M. Altered cellular proliferation and mesoderm patterning in Polycomb-M33-deficient mice. *Development*. 1997; 124: 721–729.
- Corley M, Kroll KL. The roles and regulation of Polycomb complexes in neural development. *Cell and Tissue Research*. 2015; 359: 65-85.
- Corredor-Adámez M, Welten MCM, Spaink HP, Jeffery JE, Schoon RT, De Bakker MAG. Bagowski CP, Meijer AH, Verbeek FJ, Richardson MK. Genomic annotation and transcriptome analysis of the zebrafish (*Danio rerio*) *hox* complex with description of a novel member, *hoxb13a*. *Evolution & Development*. 2005; 7(5): 362-375.
- Costa B, Estrada MF, Barroso MT, Fior R. Zebrafish Patient-Derived Avatars from Digestive Cancers for Anti-cancer Therapy Screening. *Current Protocols*. 2022; 2(4): e415.
- Costa B, Estrada MF, Mendes RV, Fior R. Zebrafish Avatars towards Personalized Medicine-A Comparative Review between Avatar Models. *Cells*. 2020; 9(2):293.
- Cotney J, Leng J, Oh S, DeMare LE, Reilly SK, Gerstein MB, Noonan JP. Chromatin state signatures associated with tissue-specific gene expression and enhancer activity in the embryonic limb. *Genome Research*. 2012; 22: 1069-1080.
- Cox BJ, Vollmer M, Tamplin O, Lu Mei, Biechele S, Gertsenstein M, van Campenhout C, Floss T, Kühn R, Wurst W, Lickert H, Rossant J. Phenotypic annotation of the mouse X chromosome. *Genome Research*. 2010; 20: 1154-1164.
- Crea F, Paolicchi E, Marquez VE, Danesi R. Polycomb genes and cancer: Time for clinical application? *Oncology Hematology*. 2012; 83(2): 184-193.
- Creaser CW. The technique of handling the zebrafish (*Brachydanio rerio*) for the production of eggs which are favourable for embryological research and are available at any specified time throughout the year. *Copeia*. 1934. 159-161.



- Creyghton MP, Cheng AW, Welstead GG, Kooistra T, Carey BW, Steine EJ, Hanna J, Lodato MA, Frampton GM, Sharp PA, Boyer LA, Young RA, Jaenisch R. H3K27ac separates active from poised enhancers and predicts developmental state. *Proceedings of the National Academy of Sciences USA*. 2010; 107(50): 21931-21936.
- Cunliffe, V.T., and Casaccia-Bonneli, P. Histone deacetylase 1 is essential for oligodendrocyte specification in the zebrafish CNS. *Mechanisms of Development*. 2006; 123: 24–30.
- Cyrus S, Burkardt D, Weaver DD, Gibson WT. PRC2-complex related dysfunction in overgrowth syndromes: A review of *EZH2*, *EED*, and *SUZ12* and their syndromic phenotypes. *American Journal of Medical Genetics part C : Seminars in Medical Genetics (Am J Med Genet C Semin Med Genet)*. 2019. 181(4): 519-531.
- da Rocha ST, Boeva V, Escamilla-Del-Arenal M, Ancelin K, Granier C, Matias NR, Sanulli S, Chow J, Schulz E, Picard C, Kaneko S, Helin K, Reinberg D, Stewart AF, Wutz A, Margueron R, Heard E. Jarid2 Is Implicated in the Initial Xist-Induced Targeting of PRC2 to the Inactive X Chromosome. *Molecular Cell*. 2014; 53: 301-316.
- Dametto FS, Fior D, Idalencio R, Rosa JGS, Fagundes M, Marqueze A, Barreto RE, Piato A, Barcellos LJG. Feeding regimen modulates zebrafish behavior. *PeerJ*. 2018; 6: e5343.
- Dang CV, Le A, Gao P. MYC-induced Cancer Cell Energy Metabolism and Therapeutic Opportunities. *Clinical Cancer Research*. 2009;15(21): 6479-6483.
- Dauber KL, Perdigoto CN, Valdes VJ, Santoriello FJ, Cohen I, Ezhkova E. Dissecting the Roles of Polycomb Repressive Complex 2 Subunits in the Control of the Skin Development. *Journal of Investigative Dermatology*. 2016; 136: 1647-1655.
- del Mar Lorente M, Marcos-Gutiérrez C, Pérez C, Schoorlemmer J, Ramírez A, Magin T, Vidal M. Loss- and gain-of-function mutations show a polycomb group function for Ring1A in mice. *Development*. 2000; 127: 5093–5100.
- den Broeder MJ, Ballangby J, Kamminga LM, Aleström P, Legler J, Lindeman LC, Kamstra JH. Inhibition of methyltransferase activity of enhancer of zeste 2 leads to enhanced lipid accumulation and altered chromatin status in zebrafish. *Epigenetics & Chromatin*. 2020; 13(1): 5.
- Denisenko ON, Bomsztyk K. The Product of the Murine Homolog of the *Drosophila extra sex combs* Gene Displays Transcriptional Repressor Activity. *Molecular and Cellular Biology*. 1997; 17(8): 47007-4717.
- Denk D, Nebral K, Bradtke J, Pass G, Möricke A, Attarbaschi A, Strehl S. PAX5-AUTS2: a recurrent fusion gene in childhood B-cell precursor acute lymphoblastic leukemia. *Leukemia Research*. 2012; 36(8): e178-81.
- Desai D, Khanna A, Pethe P. Inhibition of RING1B alters lineage specificity in human embryonic stem cells. *Cell Biology International*. 2020; 44(6): 1299-1311.
- Desai D, Pethe P. Polycomb repressive complex 1: Regulators of neurogenesis from embryonic to adult stage. *Journal of Cell Physiology*. 2020; 235(5): 4031-4045.
- Di Carlo V, Mocavini I, Di Croce L. Polycomb complexes in normal and malignant hematopoiesis. *Journal of Cell Biology*. 2018; 218(1): 55-69.
- Dickinson ME, Flennikin AM, Ji X, Teboul L, Wong MD, White JK, Meehan TF, Weninger WJ, Westerberg H, Adissu H, Baker CN, Bower L, Brown JM, Caddle LB, Chiani F, Clary D, Cleak J, Daly MJ,

- Denegre JM, Doe B, Dolan ME, Edie SM, Fuchs H, Gailus-Durner V, Galli A, Gambadoro A, Gallegos J, Guo S, Horner NR, Hsu CW, Johnson SJ, Kalaga S, Keith LC, Lanoue L, Lawson TN, Lek M, Mark M, Marshall S, Mason J, McElwee ML, Newbigging S, Nutter LMJ, Peterson KA, Ramirez-Solis R, Rowland DJ, Ryder E, Samocha KE, Seavitt JR, Selloum M, Szoke-Kovacs Z, Tamura M, Trainor AG, Tudose I, Wakana S, Warren J, Wendling O, West DB, Wong L, Yoshiki A, MacArthur DG, Tocchini-Valentini GP, Gao X, Flicek P, Bradley A, Skarnes WC, Justice MJ, Parkinson HE, Moore M, Wells S, Braun RE, Svenson KL, de Angelis MH, Herault Y, Mohun T, Mallon AM, Henkelman RM, Brown SDM, Adams DJ, Lloyd KCK, McKerlie C, Beaudet *al*, Bucan M, Murray SA. High-throughput discovery of novel developmental phenotypes. *Nature*. 2016; 537(7621): 508-514.
- Dietrich N, Lerdrup M, Landt E, Agrawal-Singh S, Bak M, Tommerup N, Rappsilber J, Södersten E, Hansen K. REST-Mediated Recruitment of Polycomb Repressor Complexes in Mammalian Cells. *PLOS Genetics*. 2012; 8(3):e1002494.
- Dong PDS, Munson CA, Norton W, Crosnier C, Pan XF, Gong ZY, Neumann CJ, Stainier DYR. Fgf10 regulates hepatopancreatic ductal system patterning and differentiation. *Nature Genetics*. 2007; 39: 397-402.
- Dong Z, Yang N, Yeo SY, Chitnis A, Guo S. Intralineaage directional Notch signaling regulates self-renewal and differentiation of asymmetrically dividing radial glia. *Neuron*. 2012; 74: 65-78.
- Doudna JA, Charpentier E. Genome editing. The new frontier of genome engineering with CRISPR-Cas9. *Science*. 2014; 346: 1258096.
- Driever W, Solnica-Krezel L, Schier AF, Neuhauss SCF, Malicki J, Stemple DL, Stainer DYR, Zwartkruis F, Abdelilah S, Rangini Z, Belak J, Boggs C. A genetic screen for mutations affecting embryogenesis in zebrafish. *Development*. 1996; 123: 37-46.
- Duchatel RJ, Jackson ER, Alvaro F, Nixon B, Hondermarck H, Dun MD. Signal transduction in Diffuse Intrinsic Pontine Glioma. *Proteomics*. 2019; 19(21-22): 1800479.
- Dumasia NP, Pethe PS. Pancreas development and the Polycomb group protein complexes. *Mechanisms of Development*. 2020; 164: 103647.
- Dupret B, Völkel P, Le Bourhis X, Angrand PO. The Polycomb Group Protein Pcgf1 Is Dispensable in Zebrafish but Involved in Early Growth and Aging. *PLOS ONE*. 2016; 11(7): e0158700.
- Dupret B, Völkel P, Le Bourhis X, Angrand PO. The Polycomb Group Protein Pcgf1 is Dispensable in Zebrafish but Involved in Early Growth and Aging. *PLOS ONE*. 2016; 11(7): e0158700.
- Dupret B, Völkel P, Vennin C, Toillon RA, Le Bourhis X, Angrand PO. The histone lysine methyltransferase Ezh2 is required for maintenance of the intestine integrity and for caudal fin regeneration in zebrafish. *Biochimica and Biophysica Acta - Gene Regulatory Mechanisms*. 2017; 1860(10): 1079-1093.
- Dupret B, Völkel P, Vennin C, Toillon RA, Le Bourhis X, Angrand PO. The histone lysine methyltransferase Ezh2 is required for maintenance of the intestine integrity and for caudal fin regeneration in zebrafish. *Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms*. 2017; 1860(10): 1079-93.
- Dura J-M, Brock HW, Santamaria P. *Polyhomeotic* : A gene of *Drosophila melanogaster* required for correct expression of segmental identity. *Molecular and General Genetics MGG*. 1985; 198: 213-220.

- Dutta SPS. Food and feeding habits of *Danio rerio* inhabiting Gadigargh stream, Jamu. *Journal of Freshwater Biology*. 1993; 5: 165-168.
- Eaton RC, Farley RD. Spawning cycle and egg production of zebrafish, *Brachydanio rerio*, in the laboratory. *Copeia*. 1974; 1974: 195-204.
- Eid W, Opitz L, Biason-Lauber A. Genome-wide identification of CBX2 targets: insights in the human sex development network. *Molecular Endocrinology*. 2015; 29(2): 247-57.
- Endoh M, Endo TA, Shinga J, Hayashi K, Farcas A, Ma K-W, Ito S, Sharif J, Endoh T, Onaga N, Nakayama M, Ishikura T, Masui O, Kessler BM, Suda T, Ohara O, Okuda A, Klose R, Kosek H. PCGF6-PRC1 suppresses premature differentiation of mouse embryonic stem cells by regulating germ cell-related genes. *Elife*. 2017; 6.
- Engeszer RE, Patterson LB, Rao AA, Parichy DM. Zebrafish in the Wild: A Review of Natural History And New Notes from The Field. *Zebrafish*. 2007; 4(1): 21-40.
- Ercoskun P, Yuce Kahraman C, Adanur Saglam K, Kanjee M, Tatar A. A new case of Turnpenny-Fry syndrome. *Am J Med Genet A*. 2022 Feb;188(2):688-691. doi: 10.1002/ajmg.a.62560. Epub 2021 Nov 9. PMID: 34750959.
- Ericson J, Rashbass P, Schedl A, Brenner-Morton S, Kawakami A, van Heyningen V, Jessell TM, Briscoe J. Pax6 Controls Progenitor Cell Identity and Neuronal Fate in Response to Graded Shh Signaling. *Cell*. 1997; 90: 169-180.
- Eroles P, Bosch A, Pérez-Fidalgo JA, Lluch A. Molecular biology in breast cancer: Intrinsic subtypes and signaling pathways. *Cancer Treatments Reviews*. 2012; 38(6): 698-707.
- Ezhkova E, Lien WH, Stokes N, Pasolli HA, Silva JM, Fuchs E. EZH1 and EZH2 cogovern histone H3K27 trimethylayion and are essential for hair follicle homeostasis and wound repair. *Genes & Development*. 2011; 25(5): 485-498.
- Ezhkova E, Pasolli HA, Parker JS, Stokes N, Su IH, Hannon G, Tarakhovsky A, Fuchs E. Ezh2 orchestrates gene expression for the stepwise differentiation of tissue-specific stem cells. *Cell*. 2009; 136: 1122-1135.
- Faheem M, Naseer MI, Rasool M, Chaudhary AG, Kumosani TA, Ilyas AM, Pushparaj P, Ahmed F, Algahtani HA, Al-Qahtani MH, Saleh Jamal H. Molecular genetics of human primary microcephaly: an overview. *BMC Medecine Genomics*. 2015; 8 Suppl 1(Suppl 1): S4.
- Fang F. Phylogenetic analysis of the Asian cyprinid genus *Danio* (Teleostei, Cyprinidae). *Copeia*. 2003; 4: 714-728.
- Farcas AM, Blackledge NP, Sudbery I, Long AK, McGouran JF, Rose NR, Lee S, Sims D, Cerase A, Sheahan TW, Koseki H, Brockdorff N, Ponting CP, Kessler BM, Klose RJ. KDM2B links the Polycomb Repressive Complex 1 (PRC1) to recognition of CpG islands. *eLIFE*. 2012; 1:e00205.
- Fares J, Fares MY, Khachfe HH, Salhab HA, Fares Y. Molecular principles of metastasis: a hallmark of cancer revisited. *Signal Transduction and Targeted Therapy*. 2020; 5:8.
- Farooq M, Sulochana KN, Pan XP, To J, Sheng D, Gong Z, Ge R. Histone deacetylase 3 (hdac3) is specifically required for liver development in zebrafish. *Developmental Biology*. 2008; 317(11): 336-53.

- Faust C, Lawson KA, Schork NJ, Thiel B, Magnuson T. The Polycomb-group gene *eed* is required for normal morphogenetic movements during gastrulation in the mouse embryo. *Development*. 1998; 125(22): 4495-506.
- Fazi B, Garbo S, Toschi N, Mangiola A, Lombardi M, Sicari D, Battistelli C, Galardi S, Michienzi A, Trevisi G, Harari-Steinfeld R, Cicchini C, Ciafrè SA. The lncRNA H19 positively affects the tumorigenic properties of glioblastoma cells and contributes to NKD1 repression through the recruitment of EZH2 on its promoter. *Oncotarget*. 2018; 9(21): 15512-15525.
- Federico A, Pallante P, Bianco M, Ferraro A, Esposito F, Monti M, Cozzolino M, Keller S, Fedele M, Leone V, Troncone G, Chiriotti L, Pucci P, Fusco A. Chromobox protein homologue 7 protein, with decreased expression in human carcinomas, positively regulates E6cadherin expression by interacting with the histone deacetylase 2 protein. *Cancer Research*. 2009; 69(17): 7079-87.
- Feng G, Sun Y. The Polycomb group gene *rnf2* is essential for central and enteric neural system development in zebrafish. *Frontiers in Neurosciences* . 2022; 16: 960149.
- Feng X, Juan A, Wang HA, Dae Ko K, Zare H, Sartorelli V. Polycomb Ezh2 controls the fate of GABAergic neurons in the embryonic cerebellum. *Development*. 2016; 143(11): 1971-80.
- Ferrari KJ, Scelfo A, Jammula S, Cuomo A, Barozzi I, Stützer A, Fischle W, Bonaldi T, Pasini D. Polycomb-dependent H3K27me1 and H3K27me2 regulate active transcription and enhancer fidelity. *Molecular Cell*. 2014; 53(1): 49-62.
- Field HA, Ober EA, Roeser T, Stainier DYR. Formation of the digestive system in zebrafish. I. Liver morphogenesis. *Developmental Biology*. 2003; 253: 279-290.
- Fischle W, Wang Y, Jacobs SA, Kim Y, Allis CD, Khorasanizadeh S. Molecular basis for the discrimination of repressive methyl-lysine marks in histone H3 by Polycomb and HP1 chromodomains. *Gene and Development*. 2003; 17(15): 1870-81.
- Flamier A, El Hajjar J, Adjaye J, Fernandes KJ, Abdouh M, Bernier G. Modeling Late-Onset Sporadic Alzheimer's Disease through BMI1 Deficiency. *Cell Reports*. 2018 ; 23(9): 2653-2666.
- Flamment M, Kammoun HL, Hainault I, Ferré P, Foufelle F. Endoplasmic reticulum stress: a new actor in the development of hepatic steatosis. *Current Opinion in Lipidology*. 2010; 21(3): 239-246.
- Flavahan WA, Gaskell E, Bernstein BE. Epigenetic plasticity and the hallmarks of cancer. *Science*. 2017; 357(6348): eaal2380.
- Flores EM, Nguyen AT, Odem MA, Eisenhoffer GT, Krachler AM. The zebrafish as a model for gastrointestinal tract-microbe interactions. *Cellular Microbiology*. 2019; 22(3): e13152.
- Forzati F, Federico A, pallante P, Abbate A, Esposito F, Malapelle U, Sepe R, Palma G, Troncone G, Scarfò M, Arra C, Fedele M, Fusco A. CBX7 is a tumor suppressor in mice and humans. *Journal of Clinical Investigations*. 2012; 122: 612–623.
- Francis NJ, Kingston RE, Woodcock CL. Chromatin Compaction by a Polycomb Group Protein Complex. *Science*. 2004; 306(5701): 1574-1577.
- Francke A, DeCamillis M, Zink D, Cheng N, Brock HW, Paro R. *Polycomb* and *polyhomeotic* are constituents of a multimeric protein complex in chromatin of *Drosophila melanogaster*. *The EMBO Journal*. 1992; 11(8): 2941-2950.

- Fritsch C, Brown JL, Kassi JA, Muller J. The DNA-binding polycomb group protein pleiohomeotic mediates silencing of a Drosophila homeotic gene. *Development*. 1999; 126(17): 3905-3913.
- Fukuda T, Tokunaga A, Sakamoto R, Yoshida N. Fbxl10/Kdm2b deficiency accelerates neural progenitor cell death and leads to exencephaly. *Molecular & Cellular Neurosciences*. 2011; 46; 614–624.
- Gao Z, Lee P, Stafford JM, von Schimmelmann, Schaefer A, Reinberg D. AUTS2 confers gene activation to Polycomb group proteins in the CNS. *Nature*. 2014; 516(7531): 349-354.
- Gao Z, Zhang J, Bonasio R, Strino F, Sawai A, Parisi F, Kluger Y, Reinberg D. PCGF homologs, CBX proteins, and RYBP define functionally distinct PRC1 family complexes. *Molecular Cell*. 2012; 45(3) : 344-56.
- García E, Marcos-Gutiérrez C, del Mar Lorente M, Moreno JC, Vidal M. RYBP, a new repressor protein that interacts with components of the mammalian Polycomb complex, and with the transcription factor YY1. *The EMBO Journal*. 1999; 18(12): 3404-3418.
- Gaubatz S, Wood JG, Livingston DM. Unusual proliferation arrest and transcriptional control properties of a newly discovered E2F family member, E2F-6. *Proceedings of the National Academy of Sciences of the United States of America*. 1998; 95:9190-9195.
- Geng Z, Gao Z. Mammalian PRC1 Complexes: Compositional Complexity and Diverse Molecular Mechanisms. *International Journal of Molecular Sciences*. 2020; 21: 8594.
- Gerhard GS, Kauffman EJ, Wang X, Stewart R, Moore JL, Kasales CJ, Demidenko E, Cheng KC. Life spans and senescent phenotypes of zebrafish (*Danio rerio*). *Experimental Gerontology*. 2002; 37: 1055-1068.
- Gerlach G, Lysiak N. Kin recognition and inbreeding avoidance in zebrafish, *Danio rerio*, is based on phenotype matching. *Animal behaviour*. 2006; 71:1371-1377.
- Gil J, O’Loughlen A. PRC1 complex diversity : where is it taking us? *Trends in Cellular Biology*. 2014; 24(11): 632-41.
- Gil J, Peters G. Regulation of the INK4b-ARF-INK4a tumour suppressor locus: all for one or one for all. *Molecular & Cellular Biology*. 2006; 7(9): 667-677.
- Glasauer SMK, Neuhauss SCF. Whole-genome duplication in teleost fishes and its evolutionary consequences. *Molecular Genetics and Genomics*. 2014; 289: 1045-1060.
- Goessling W, Sadler KC. Zebrafish: An Important Tool for Liver Disease Research. *Gastroenterology*. 2015; 149(6): 1361-1377.
- Goldstone JV, McArthur AG, Kubota A, Zanette J, Parente T, Jönsson ME, Nelson DR, Stegeman JJ. Identification and developmental expression of Cytochrome P450 genes in Zebrafish: the full complement of Cytochrome P450 genes in Zebrafish. *BMC Genomics*. 2010; 643.
- Grijzenhout A, Godwin J, Koseki H, Gdula MR, Szumska D, McGouran JF, Bhattacharya S, Kessler BM, Brockdorff N, Cooper S. Functional analysis of AEBP2, a PRC2 Polycomb protein, reveals a Trithorax phenotype in embryonic development and in ESCs. *Development*. 2016; 143: 2716-2723.
- Gröbner SN, Worst BC, Weischenfeldt J, Buchhalter I, Kleinheinz K, Rudneva VA, Johann PD, Balasubramanian GP, Segura-Wang M, Brabetz S, Bender S, Hutter B, Sturm D, Pfaff E, Hübschmann D, Zipprich G, Heinold M, Eils J, Lawrenz C, Erkek S, Lambo S, Waszak S,

- Blattmann C, Borkhardt A, Kuhlen M, Eggert A, Fulda S, Gessler M, Wegert J, Kappler R, Baumhoer D, Burdach S, Kirschner-Schwabe R, Kontny U, Kulozik AE, Lohmann D, Hettmer S, Eckert C, Bielack S, Nathrath M, Niemeyer C, Richter GH, Schulte J, Siebert R, Westermann F, Molenaar JJ, Vassal G, Witt H; ICGC PedBrain-Seq Project; ICGC MMML-Seq Project, Burkhardt B, Kratz CP, Witt O, van Tilburg CM, Kramm CM, Fleischhack G, Dirksen U, Rutkowski S, Frühwald M, von Hoff K, Wolf S, Klingebiel T, Koscielniak E, Landgraf P, Koster J, Resnick AC, Zhang J, Liu Y, Zhou X, Waanders AJ, Zwijnenburg DA, Raman P, Brors B, Weber UD, Northcott PA, Pajtler KW, Kool M, Piro RM, Korbel JO, Schlesner M, Eils R, Jones DTW, Lichter P, Chavez L, Zapatka M, Pfister SM. The landscape of genomic alterations across childhood cancers. *Nature*. 2018; 555(7696): 321-327.
- Grunwald DJ, Eisen JS. Headwater of the zebrafish – emergence of a new model vertebrate. *Nature Reviews Genetics*; 3(9): 717-724.
- Gu W, Cheng Y, Wang S, Sun T, Li Z. PHD Finger Protein 19 Promotes Cardiac Hypertrophy via Epigenetically Regulating SIRT2. *Cardiovascular Toxicology*. 2021; 21(6): 451-461.
- Guo S. Linking genes to brain, behavior and neurological diseases: what can we learn from zebrafish. *Genes, Brain and Behavior*. 2003; 3: 63-74.
- Haffter P, Granato M, Brand M, Mullins MC, Hammerschmidt M, Kane DA, Odenthal J, Van Eeden FJM, Jiang YJ, Heisenberg CP, Kelsh RN, Furutani-Seiki M, Vogelsang E, Beuchle D, Schach U, Fabian C, Nüsslein-Volhard C. The identification of genes with unique and essential function in the development of the zebrafish, *Danio rerio*. *Development*. 1996; 123: 1-36.
- HAMILTON-BUCHANAN, F. An account of the fishes found in the River Ganges and its branches. *Archibald Constable and Co*. Edinburgh. 1822.
- Han Z, Xing X, Hu M, Zhang Y, Liu P, Chai J. Structural basis of EZH2 Recognition by EED. *Structure*. 2007; 15(10): 1306-1315.
- Hanahan D, Weinberg RA. Hallmarks of Cancer: The Next Generation. *Cell*. 2011; 144(5):646-674.
- Harshbarger JJ, Slatick MS. Lesser known aquarium fish tumor models. *Marine Biotechnology*. 2001; 3: S115–S129.
- Hasegawa K, Sin HS, Maezawa S, Broering TJ, Kartashov AV, Alaattam KG, Ichijima Y, Zhang F, Bacon WC, Greis KD, Andreassen PR, Barski A, Namekawa SH. SCML2 Establishes the Male Germline Epigenome through Regulation of Histone H2A Ubiquitination. *Developmental Cell*. 2015; 32(5): 574-588.
- Hauenschild A, Ringrose L, Altmutter C, Paro R, Rehmsmeier M. Evolutionary plasticity of polycomb/trithorax response elements in Drosophila species. *PLOS Biology*. 2008; 6(10):e261.
- Hauri S, Comoglio F, Seimiya M, Gerstung M, Glatzer T, Hansen K, Aebersold R, Paro R, Gstaiger M, Beisel C. A High-Density Map for Navigating the Human Polycomb Complexome. *Cell Reports*. 2016; 17(2): 583-595.
- He J, Shen L, Wan M, Taranova O, Wu H, Zhang Y. Kdm2b maintains murine embryonic stem cell status by recruiting PRC1 complex to CpG islands of developmental genes. *Nature Cell Biology*. 2013; 15(4): 373-384.

- Healy E, Mucha M, Glancy E, Fitzpatrick DJ, Conway E, Neikes HK, Monger C, Van Mierlo G, Baltissen MP, Koseki Y, Vermeulen M, Koseki H, Bracken AP. PRC2.1 and PRC2.2 Synergize to coordinate H3K27 Trimethylation. *Molecular Cell*. 2019; 76: 437-452.
- Hemenway CS, de Erkenez AC, Gould GCD. The polycomb protein MPc3 interacts with AF9, an MLL fusion partner in t(9;11)(p22;q23) acute leukemias. *Oncogene*. 2001; 20:3798-3805.
- Henriquez B, Bustos FJ, Aguilar R, Becerra A, Simon F, Montecino M, van Zundert. Ezh1 and Ezh2 differentially regulate PSD-95 gene transcription in developing hippocampal neurons. *Molecular and Cellular Neuroscience*. 2013; 57: 130-143.
- Her GM, Chiang CC, Chen WY, Wu JL. *In vivo* studies of liver-type fatty acid binding protein (L-FABP) gene expression in liver of transgenic zebrafish (*Danio rerio*). *FEBS Letters* 2003; 538(1-3): 125-133.
- Hickey GJM, Wike CL, Nie X, Guo Y, Tan M, Murphy PJ, Cairns BR. Establishment of developmental gene silencing by ordered Polycomb complex recruitment in early zebrafish embryos. *eLife*. 2022; 11: e67738.
- Hidalgo I, Herrera-Merchan A, Ligos JM, Carramolino L, Nunez J, Martinez F, Dominguez O, Torres M, Gonzalez S. Ezh1 Is Required for Hematopoietic Stem Cell Maintenance and Prevents Senescence-like Cell Cycle Arrest. *Cell Stem Cell*. 2012; 11(5):649-662.
- Hobert O, Jallal B, Ullrich A. Interaction of Vav with ENX-1, a putative transcriptional regulator of homeobox gene expression. *Molecular & Cellular Biology*. 1996a; 16, 3066-3073.
- Hobert O, Sures I, Ciossek T, Fuchs M, Ullrich A. Isolation and developmental expression analysis of *Enx-1* a novel mouse Polycomb group gene. *Mechanisms of Development*. 1996b; 55: 171-184.
- Højfeldt JW, Laugesen A, Willumsen BM, Damhofer H, Hedehus L, Tvardovskiy A, Mohammad F, Jensen ON, Helin K. Accurate H3K27 methylation can be established *de novo* by SUZ12-directed PRC2. *Nature Structural & Molecular Biology*. 2018; 25:225-232.
- Højfeldt JW, Laugesen A, Willumsen BM, Damhofer H, Hedehus L, Tvardovsky A, Mohammad F, Jensen ON, Helin K. Accurate H3K27 methylation can be established *de novo* by SUZ12-directed PRC2. *Nature Structural & Molecular Biology*. 2018; 25(3): 225-232.
- Holden LA, Brown KH. Baseline mRNA expression differs widely between common laboratory strains of zebrafish. *Scientific Reports*. 2017; 8: 4780.
- Holland PWH, Garcia-Fernández J, Williams NA, Sidow A. Gene duplications and the origins of vertebrate development. *Development*. 1994; Suppl: 215-133.
- Holoch D, Margueron R. Mechanisms Regulating PRC2 Recruitment and Enzymatic Activity. *Trends in Biochemical Sciences*. 2017; 42(7): 531-542.
- Howe K, Clark MD, Torroja CF, Tarrant J, Berthelot C, Muffato M, Collins JE, Humphray S, McLaren K, Mattwes L, McLaren S, Sealy I, Caccamo M, Churcher C, Scott C, Barrett JC, Koch R, Rauch GJ, White S, Chow W, Kilian B, Quintais LT, Guerra-Assunção JA, Zhou Y, ... , Stemple DL. The zebrafish reference genome sequence and its relationship to the human genome. *Nature*. 2013; 496: 498-503.
- Hrzenjak A. JAZF1/SUZ12 gene fusion in endometrial stromal sarcomas. *Orphanet Journal of Rare Diseases*. 2016; 11:15.

- Hu XM, Nie XY, Xu KL, Wang Y, Tang F, Du ZG, Xiong J. H3K27M Mutation Doesn't Mean Worse Prognosis in Old Patients. *Frontiers in Oncology*. 2022 ; 12 : 912166.
- Hu Y, Cheng Y, Jiang X, Zhang Y, Wang H, Ren H, Xu Y, Jiang J, Wang Q, Su H, Zhang B, Qiu X. PCGF3 promotes the proliferation and migration of non-small cell lung cancer cells via the PI3K/AKT signaling pathway. *Experimental Cell Research*. 2021; 400(2): 112496.
- Huang J, Gou H, Yao J, Yi K, Jin Z, Masuoka M, Zhao T. The noncanonical role of EZH2 in cancer. *Cancer Science*. 2021; 112(4): 1376-1382.
- Huang Z, Wu Q, Guryanova OA, Cheng L, Shou W, Rich JN, Bao S. Deubiquitylase HAUSP stabilizes REST and promotes maintenance of neural progenitor cells. *Nature Cellular Biology*. 2011; 13: 142–152.
- Huillard E, Ziercher L, Blond O, Wong M, Deloulme JC, Souchelnytskyi S, Baudier J, Cochet C, Buchou T. Disruption of CK2beta in embryonic neural stem cells compromises proliferation and oligodendrogenesis in the mouse telencephalon. *Molecular & Cellular Biology*. 2010; 30: 2737–2749.
- Hussain M, Lu Y, Liu YQ, Su K, Zhang J, Liu J, Zhou GB. Skp1: Implications in cancer and SCF-oriented anti-cancer drug discovery. *Pharmacological Research*. 2016;111: 34-42.
- Huynh KD, Fischle W, Verdin E, Bardwell VJ. BCoR, a novel corepressor involved in BCL-6 repression. *Genes & Development*. 2000; 14: 1810–1823.
- Ikeda K, Ueda T, Yamasaki N, Nakata Y, Sera Y, Nagamachi A, Miyama T, Kobayashi H, Takubo K, Kanai A, Oda H, Wolff L, Honda Z, Ichinohe T, Matsubara A, Suda T, Inaba T, Honda H. Maintenance of the functional integrity of mouse hematopoiesis by EED and promotion of leukemogenesis by EED haploinsufficiency. *Scientific Reports*. 2016; 19(6):29454.
- Imagawa E, Seyama R, Aoi H, Uchiyama Y, Marcarini BG, Furquim I, Honjo RS, Bertola DR, Kim CA, Matsumoto N. Imagawa-Matsumoto syndrome: SUZ12-related overgrowth disorder. *Clinical Genetics*. 2023; 103(4): 383-391.
- Inoue JG, Miya M, Lam K, Tay BH, Danks JA, Bell Justin, Walker TI, Venkatesh B. Evolutionary Origin and Phylogeny of the Modern Holocephalans (Chondrichthyes: Chimaeriformes): A Mitogenomic Perspective. *Molecular Biology and Evolution*. 2010; 27(11): 2576-2586.
- Ishimaru N, Arakaki R, Omotehara F, Yamada K, Mishima K, Saito I, Hayashi Y. Novel role for RbAp48 in tissue-specific, estrogen deficiency-dependent apoptosis in the exocrine glands. *Molecular Cell Biology*. 2006; 26(8): 2924-35.
- Isono K, Endo TA, Ku M, Yamada D, Suzuki R, Sharif J, Ishikura T, Toyoda T, Bernstein BE, Koseki H. SAM Domain polymerization Links Subnuclear Clustering of PRC1 to Gene Silencing. *Developmental Cell*. 2013; 26(6): 565-577.
- Isono K, Endo TA, Ku M, Yamada D, Suzuki R, Sharif J, Ishikura T, Toyoda T, Bernstein BE, Koseki H. SAM domain polymerization links subnuclear clustering of PRC1 to gene silencing. *Developmental Cell*. 2013; 26(6): 565-77.
- Isono KI, Fujimura YI, Shinga J, Yamaki M, O-Wang J, Takihara Y, Murahashi Y, Takada Y, Mizutani-Koseki Y, Koseki H. Mammalian polyhomeotic homologues Phc2 and Phc1 act in synergy to mediate polycomb repression of Hox genes. *Molecular Cell Biology*. 2005; 25(15): 6694-706.
- Iwagawa T, Watanabe S. Molecular mechanisms of H3K27me3 and H3K4me3 in retinal development. *Neurosciences Research*. 2019; 138: 43-48.



- Iwama A, Oguro H, Negishi M, Kato Y, Tsukui H, Ema H, Kamijo T, Katoh-Fukui Y, Koseki H, van Lohuizen M, Nakauchi H. Enhanced self-renewal of hematopoietic stem cells mediated by the polycomb gene product Bmi1. *Immunity*. 2004; 21: 843-851.
- Jacobs JJ, Kieboom K, Marino S, DePinho RA, van Lohuizen M. The oncogene and Polycomb group gene bmi-1 regulates cell proliferation and senescence through the ink4a locus. *Nature*. 1999; 397: 164-168.
- Jadhav U, Manieri E, Nalapareddy K, Madha S, Chakrabarti S, Wucherpfennig K, Barefoot M, Shivdasani RA. Replicational Dilution of H3K27me3 in Mammalian Cells and the Role of Poised Promoters. *Molecular Cell*. 2020; 78(1): 1414-151.
- Jain P, Ballare C, Blanco E, Vizan P, Di Croce L. PHF19 mediated regulation of proliferation and invasiveness in prostate cancer cells. *Elife*. 2020; 9: e51373.
- Jalaguier S, Teyssier C, Nait Achour T, Lucas A, Bonnet S, Rodriguez C, Elarouci N, Lapierre M, Cavallès V. Complex regulation of LCoR signaling in breast cancer cells. *Oncogene*. 2017; 36(33): 4790-4801.
- Ji G, Zhou W, Du J, Zhou J, Wu D, Zhao M, Yang L, Hao A. PCGF1 promotes epigenetic activation of stemness markers and colorectal cancer stem cell enrichment. *Cell Death Diseases*. 2021;12(7):633.
- Jiao L, Liu X, Structural basis of histone H3K27 trimethylation by an active polycomb repressive complex 2. *Science*. 2015; 350(6258): aac4383-aac4383.
- Jimenez-Gonzalez A, Garcà-Concejo A, Leon-Lobera F, Rodriguez R. Morphine delays neural stem cells differentiation by facilitating Nestin overexpression. *Biochimica et Biophysica Acta - General Subjects*. 2018; 1862(3): 474-484.
- Jin Q, Yu LR, Wang L, Zhang Z, Kasper LH, Jee JE, Wang C, Brindle PK, Dent SY, Ge K. Distinct roles of GCN5/PCAF-mediated H3K9ac and CBP/p300-mediated H3K18/27ac in nuclear receptor transactivation. *EMBO Journal*. 2011; 30(2): 249-62.
- Johansson ALV, Trewin CB, Fredriksson I, Reinertsen KV, Russnes H, Ursin G. In modern times, how important are breast cancer stage, grade and receptor subtype for survival: a population-based cohort study. *Breast Cancer Research*. 2021; 23:17.
- Johnstone AL, O'Reilly JJ, Patel AJ, Guo Z, Andrade NS, Magistri M, Nathanson L, Esanov R, Miller BH, Turecki G, Brothers SP, Zeier Z, Wahlestedt C. EZH1 is an antipsychotic-sensitive epigenetic modulator of social and motivational behavior that is dysregulated in schizophrenia. *Neurobiology Diseases*. 2018; 119: 149-158.
- Jones CA, Ng J, Peterson AJ, Morgan K, Simon J, Jones RS. The Drosophila esc and E(z) proteins are direct partners in polycomb group-mediated repression. *Molecular & Cellular Biology*. 1998; 18: 2825Y2834.
- Juan AH, Wang S, Dae Ko K, Zare H, Tsai PF, Feng X, Vivanco KO, Ascoli AM, Gutierrez-Cruz G, Krebs J, Sidoli S, Knight AL, Pedersen RA, Garcia BA, Casellas R, Zou J, Sartorelli V. Roles of H3K27me2 and H3K27me3 Examined during Fate Specification of Embryonic Stem Cells. *Cell Reports*. 2016; 17(5): 1369-1382.
- Jukam D, Shariati SAM, Skotheim JM. Zygotic genome activation in vertebrates. *Developmental cell*. 2017; 42(4): 316-332.

- Junco SE, Wang R, Gaipa JC, Taylor AB, Schirf V, Gearhart M, Bardwell VJ, Demeler B, Hart PJ, Kim CA. Structure of the polycomb group protein PCGF1 in complex with BCOR reveals basis for binding selectivity of PCGF homologs. *Structure*. 2013; 21(4): 665-71.
- Jürgens G. A group of genes controlling the spatial expression of bithorax complex in *Drosophila*. *Nature*. 1985; 316(11):153-155.
- Kajiume T, Ohno N, Sera Y, Kawahara Y. Reciprocal expression of Bmi1 and Mel-18 is associated with functioning of primitive hematopoietic cells. *Experimental Hematology*. 2009; 37(7): 857-866.
- Kalb R, Latwiel S, Baymaz HI, Jansen PWTC, Müller CW, Vermeulen M, Müller J. Histone H2A monoubiquitination promotes histone H3 methylation in Polycomb repression. *Nature Structural & Molecular Biology*. 2014; 21(6): 569-571.
- Kalisch WE, Rasmuson B. Changes of *zeste* phenotype induced by autosomal mutations in *Drosophila melanogaster*. *Hereditas*. 1974; 78: 97-104.
- Kane DA, Kimmel CB. The zebrafish midblastula transition. *Development*. 1993; 119(2): 447-56.
- Kapranow P, Cheng J, Dike S, Dutttagupta R, Willingham A, Stadler PF, Hertel J, Hackermüller J, Hofacker IL, Bell I, Cheung E, Drenkow J, Dumais E, Patel S, Helt S, Ganesh M, Ghosh S, Piccolboni A, Sementchenko V, Tammana H, Gingers TR. RNA Maps Reveal New RNA Classes and a Possible Function for Pervasive Transcription. *Science*. 316(5830): 1484-1488.
- Katoch S, Patial V. Zebrafish: An emerging model system to study liver diseases and related drug discovery. *Journal of Applied Toxicology*. 2020; 1-19.
- Katoh-Fukui Y, Miyabayashi K, Komatsu T, Owaki A, Baba T, Shima Y, Kidokoro T, Kanai Y, Schedl A, Wilhelm D, Koopman P, Okuno Y, Mohashi KI. *Cbx2*, a polycomb group gene, is required for *Sry* Gene Expression in Mice. *Endocrinology*. 2012; 153(2): 913-924.
- Katoh-Fukui Y, Tsuchiya R, Shiroishi T, Nakahara Y, Hashimoto N, Noguchi K, Higashinakagawa T. Male-to-female sex reversal in M33 mutant mice. *Nature*. 1998; 393: 688–692.
- Kawamura A, Yamada K, Fujimori K, Higashinakagawa T. Alternative transcripts of a polyhomeotic gene homolog are expressed in distinct regions of somites during segmentation of zebrafish embryos. *Biochemical and Biophysical Research Communications*. 2002a; 291(2): 245-254.
- Kawamura A, Yokota S, Yamada K, Inoue H, Inohaya K, Yamazaki K, Yasumasu I, Higashinakagawa. *Biochemical and Biophysical Research Communications*. 2002b; 294(2): 456-463.
- Kim H, Kang K, Kim J. AEBP2 as a potential targeting protein for Polycomb Repression Complex PRC2. *Nucleic Acids Research*. 2009; 37(9): 2940-2950.
- Kim J, Kingston RE. The CBX family of proteins in transcriptional repression and memory. *Journal of Biosciences*. 2020; 45:16.
- Kim TG, Kraus JC, Chen J, Lee Y. JUMONJI, a critical factor for cardiac development, functions as a transcriptional repressor. *Journal of Biological Chemistry*. 2003; 278: 42247-42255.
- Kimmel CB, Ballard WW, Kimmel SR, Ullmann B, Schilling TF. Stages of embryonic development of the zebrafish. *Developmental Dynamics*. 1995; 203: 253-310.
- Kimmel CB, Warga RM, Schilling TF. Origin and organization of the zebrafish fate map. *Development*. 1990 ; 108: 581-594.

- Kimmel CB. Genetics and early development of zebrafish. *Trends in Genetics*. 1989; 5: 283-288.
- Kimmel CB. Patterning the brain of the zebrafish embryo. *Annual Review of Neuroscience*. 1993; 16: 707-732.
- Kimmel CB. Reticulospinal and vestibulospinal neurons in the young larva of a teleost fish, *Brachydanio rerio*. *Progress in Brain Research*. 1982; 57: 1–23.
- Kinkel MD, Prince VE. On the diabetic menu: Zebrafish as a model for pancreas development and function. *BioEssays*. 2009; 31(2): 139-152.
- Klauke K, Radulović V, Broekhuis M, Weersing E, Zwart E, Olthof S, Ritsema M, Bruggeman S, Wu X, Helin K, Byrstrykh L, de Haan G. Polycomb Cbx family members mediate the balance between haematopoietic stem cell self-renewal and differentiation. *Nature Cell Biology*. 2013; 15(4): 353-362.
- Kloet SL, Makowski MM, Baymaz HI, van Voorthuysen L, Karemaker ID, Santanach A, Jansen PWTC, Di Croce L, Vermeulen M. The dynamic interactome and genomic targets of Polycomb complexes during stem-cell differentiation. *Nature Structural & Molecular Biology*. 2016; 23: 682-690.
- Komoike Y, Kawamura A, Shindo N, Sato C, Satoh J, Shiurba R, *et al.* Zebrafish Polycomb group gene *ph2a* is required for epiboly and tailbud formation acting downstream of FGF signaling. *Biochemical and Biophysical Research Communications* 2005; 328(4):858-66.
- Kon N, Kobayashi Y, Lji M, Brooks CL, Ludwig T, Gu W. Inactivation of HAUSP in vivo modulates p53 function. *Oncogene*. 2010; 29: 1270–1279.
- Koppens MAJ, Bounova G, Gargiulo G, Tanger E, Janssen H, Cornelissen-Steijger P, Blom M, Song JY, Wessels LFA, van Lohuizen. Deletion of Polycomb Repressive Complex 2 From Mouse Intestine Causes Loss of Stem Cells. *Gastroenterology*. 2016; 151(4): 684-697.
- Korz V, Sleptsova I, Liao J, He J, Gong Z. Expression of zebrafish bHLH genes *ngn1* and *nrd* defines distinct stages of neural differentiation. *Developmental Dynamics*. 1998; 213(1): 92-104.
- Kouzaridse T. Chromatin modifications and their function. *Cell*. 2007; 128(4): 693-705.
- Koyen AE, Madden MZ, Park D, Minten EV, Kapoor-Vazirani P, Werner E, Pfister NT, Haji-Seyed-Javadi R, Zhang H, Xu J, Deng N, Duong DM, Peecen TJ, Frazier Z, Nagel Z, Lazaro JB, Moouw KW, Seyfried NT, Moreno CS, Owonikoko TK, Deng X, Yu DS. EZH2 has a non-catalytic and PRC2-independent role in stabilizing DDB2 to promote nucleotide excision repair. *Oncogene*. 2020; 39(25): 4798-4813.
- Krug B, De Jay N, Harutyunyan A S, Deshmukh S, Marchione D M, Guilhamon P, Bertrand K C, Mikael L G, McConechy M K, Chen C, Khazaei S, Koncar R F, Agnihotri S, Faury D, Ellezam B, Weil A G, Ursini-Siegel J, De Carvalho D, Dirks P B, Lewis P W, Salomoni P, Lupien M, Arrowsmith C, Lasko P F, Garcia B A, Kleinman C L, Jabado N, Mack S C. Pervasive H3K27 acetylation leads to ERV expression and a therapeutic vulnerability in H3K27M gliomas. *Cancer Cell*. 2019; 35(5): 782-797.e8.
- Ku M, Koche RP, Rheinbay E, Mendenhall EM, Endoh M, Mikkelsen TS, Presser A, Nusbaum C, Xie X, Chi AS, Adli M, Kasif S, Ptaszek LM, Cowan CA, Lander ES, Koseki H, Bernstein BE. Genomewide Analysis of PRC1 and PRC2 Occupancy Identifies Two Classes of Bivalent Domains. *PLoS Genetics*. 2008; 4(10):e1000242.

- Kushawah G, Hernandez-Huertas L, del Prado JAN, Martinez-Morales JR, DeVore ML, Hassan H, Moreno-Sanchez I, Tomas-Gallardo L, Diaz-Moscoso A, Monges DE, Guelfo JR, Theune WC, Brannan EO, Wang W, Corbin TJ, Moran AM, Alvarado AS, Málaga-Trillo E, Takacs CM, Bazzini AA, Moreno-Mateos MA. CRISPR-Cas13d Induces Efficient mRNA Knockdown in Animal Embryos. *Developmental Cell*. 2020; 54(6): 805-817.
- Kuzmichev A, Nishioka K, Erdjument-Bromage H, Tempst p, Reinberg D. Histone methyltransferase activity associated with a human multiprotein complex containing the Enhancer of Zeste protein. *Genes & Development*. 2002; 16: 2893-2905.
- Laale HW. Ethanol induced notochord and spinal cord duplications in the embryo of the zebrafish, *Brachydanio rerio*. *Journal of Experimental Zoology*. 1971; 177: 51-64.
- Lacour B, Goujon S, Guissou S, Guyot-Goubin A, Desmée S, Désandes E, Clavel J. Childhood cancer survival in France, 2000-2008. *European Journal of Cancer Prevention*. 2014; 23(5): 449-57.
- Laible G, Wolf A, Dorn R, Reuter G, Nislow C, Lebersorger A, Popkin D, Pillus L, Jenuwein T. Mammalian homologues of the Polycomb-group gene Enhancer of zeste mediate gene silencing in *Drosophila* heterochromatin and at *A. cerevisiae* telomeres. *EMBO Journal*. 1997; 16(11): 33219-3232.
- Lakstygai AM, de Abreu MS, Kalueff AV. Zebrafish models of epigenetic regulation of CNS functions. *Brain Research Bulletin*. 2018; 142: 344-351.
- Laugesen A, Højfeldt JW, Helin K. Molecular mechanisms directing PRC2 recruitment and H3K27 methylation. *Molecular Cell*. 2019; 74(1): 8-18.
- Lavarone E, Barbieri CM, Pasini D. Dissecting the role of H3K27 acetylation and methylation in PRC2 mediated control of cellular identity. *Nature Communications*. 2019; 10: 1679.
- Lawrence C. The husbandry of zebrafish (*Danio rerio*): A review. *Journal of Aquaculture*. 2007; 269(1-4): 0-20.
- Le Faou P, Völkel P, Angrand PO. The zebrafish genes encoding the Polycomb repressive complex (PRC) 1. *Gene*. 2011; 475: 10-21.
- Lecerf C, Le Bourhis X, Adriaenssens E. The long non-coding RAN *H19*: an active player with multiple facets to sustain the hallmarks of cancer. *Cellular and Molecular Life Sciences*. 2019; 76: 4673-4687.
- Lee J, Lee BK, Gross JM. Bcl6a function is required during optic cup formation to prevent p53-dependent apoptosis and colobomata. *Human Molecular Genetics*. 2013; 22(17): 3568-82.
- Lee KW, Webb SE, Miller AL. Awake of free cytosolic calcium traverses zebrafish eggs on activation. *Developmental Biology*. 1999; 214: 168-180.
- Lee SC, Miller S, Kauppi M, Lebois M, Di Rago L, Metcalf D, Kinkel SA, Josefsson EC, Blewitt ME, Majewski IJ, Alexander WS. Polycomb repressive complex 2 component Suz12 is required for hematopoietic stem cell function and lymphopoiesis. *Blood*. 2015 ; 126(2): 167-75.
- Lee TI, Jenner RG, Boyer LA, Guenther MG, Levine SS, Kumar RM, Chevalier B, Johnstone SE, Cole MF, Isono KI, Koseki H, Fuchikami T, Abe K, Murray HL, Zucker JP, Yuan B, Otte AP, Volkert TL, Bartel DP, Melton DA, Gifford DK, Jaenisch R, Young RA. Control of Developmental Regulators by Polycomb in Human Embryonic Stem Cells. *Cells*. 2006; 125: 301-313.

- Lee W, Teckie S, Wiesner T, Ran L, Prieto Granada CN, Lin M, Zhu S, Cao Z, Liang Y, Sboner A, Tap WD, Fletcher JA, Huberman KH, Qin LX, Viale A, Singer S, Zheng D, Berger MF, Chen Y, Antonescu CR, Chi P. PRC2 is recurrently inactivated through EED or SUZ12 loss in malignant peripheral nerve sheath tumors. *Nature Genetics*. 2014; 46(11): 1227-32.
- Lee W, Teckie S, Wiesner T, Ran Leili, Granada CNP, Lin M, Zhu S, Cao Z, Liang Y, Sboner A, Tap WD, Fletcher JA, Huberman KH, Qin LX, Viale A, Singer S, Zheng D, Berger MF, Chen Y, Antonescu CR, Chi P. PRC2 is recurrently inactivated through EED or SUZ12 loss in malignant peripheral nerve sheath tumors. *Nature Genetics*. 2014; 46: 1227-1232.
- Lewis EB. A gene complex controlling segmentation in *Drosophila*. *Nature* 1978; 276: 565-570
- Lewis NA, Klein RH, Kelly C, Yee J, Knoepfler PS. Histone H3.3 K27M chromatin functions implicate a network of neurodevelopmental factors including ASCL1 and NEUROD1 in DIPG. *Epigenetics & Chromatin*. 2022; 15: 18.
- Lewis PH. New mutants report. *Drosophophila Information Service*. 1947; 21: 69.
- Lewis PW, Muller MM, Koletsky MS, Cordero F, Lin S, Banaszynski LA, Garcia BA, Muir TW, Becher OJ, Allis CD. Inhibition of PRC2 Activity by a Gain-of-Function H3 Mutation Found in Pediatric Glioblastoma. *Science* 2013; 340(6134): 857-861.
- Li G, Margueron R, Ku M, Chambon P, Bernstein BE, Reinberg D. Jarid2 and PRC2, partners in regulating gene expression. *Genes & Development*. 2010; 24:368-380.
- Li H, Liefke R, Jiang J, Kurland JV, Tian W, Deng P, Zhang W, He Q, Patel DJ, Bulyk ML, Shi Y, Wang Z. Polycomb-like proteins link the PRC2 complex to CpG islands. *Nature*. 2017; 549:287–291.
- Li J, Li D, Zhang X, Li C, Zhu F. Long noncoding RNA SLC9A3 AS1 increases E2F6 expression by sponging microRNA 486 5p and thus facilitates the oncogenesis of nasopharyngeal carcinoma. *Oncology Reports*. 2021; 46(2): 165.
- Li J, Xu Y, Long XD, Wang W, Jiao HK, Mei Z, Yin QQ, Ma LN, Zhou AW, Wang LS, Yao M, Xia Q, Chen GQ. Cbx4 governs hif-1 $\alpha$  to potentiate angiogenesis of hepatocellular carcinoma by its sumo e3 ligase activity. *Cancer Cell*. 2014; 25: 547–548
- Li Q, Pan Z, Cao Z, Zhao S. Comprehensive Analysis of Prognostic Value and Immune Infiltration of Chromobox Family Members in Colorectal Cancer. *Frontiers in Oncology*. 2020; 10:582667.
- Li X, Bian WJ, Liu XR, Wang J, Luo S, Li BM, Yi YH, Wu QY, Zhai QX, Gao LD, Zhang HF, He N, Liao WP; China Epilepsy Gene 1.0 Project. BCOR variants are associated with X-linked recessive partial epilepsy. *Epilepsy Research*. 2022; 187: 107036.
- Li X, Gou J, Yang X. Bioinformatic analysis of the expression and prognostic value of chromobox family proteins in human breast cancer. *Scientific Reports*. 2020; 10: 17739.
- Li X, Ji G, Zhou J, Du J, Li X, Shi W, Hu Y, Zhou W, Hao A. Pcgf1 Regulates Early Neural Tube Development Through Histone Methylation in Zebrafish. *Frontiers in Cell and Developmental Biology*. 2021; 8: 581636.
- Li X, Lin Y, Yang X, Wu X, He X. Long noncoding RNA H19 regulates EZH2 expression by interacting with miR-630 and promotes cell invasion in nasopharyngeal carcinoma. *Biochemical and Biophysical Research Communications*. 2016; 473(4): 913-919.

- Li Y, Huang W, Huang S, Du J, Huang C. Screening of anti-cancer agent using zebrafish: comparison with the MTT assay. *Biochemical and Biophysical Research Communications*. 2012; 422(1): 85-90.
- Li YD, Lv Z, Zhu WF. RBBP4 promotes colon cancer malignant progression via regulating Wnt/ $\beta$ -catenin pathway. *World Journal of Gastroenterology*. 2020; 26(35): 5328-5342.
- Liang S, Zhou Z, Zhou Z, Liang J, Lin W, Zhang C, Zhou C, Zhao H, Meng X, Zou F, Yu C, Cai S. Blockade of CBX4-mediated  $\beta$ -catenin SUMOylation attenuates airway epithelial barrier dysfunction in asthma. *International Immunopharmacology*. 2022; 113(Pt A):109333.
- Liang YK, Lin HY, Chen CF, Zeng D. Prognostic values of distinct CBX family members in breast cancer. *Oncotarget*. 2017; 8(54): 92375-92387.
- Liu B, Liu YF, Du YR, Mardaryev AN, Yang W, Chen H, Xu ZM, Xu CQ, Zhang XR, Botchkarev VA, Zhang Y, Xu GL. Cbx4 regulates the proliferation of thymic epithelial cells and thymus function. *Development*. 2013; 140: 780-788.
- Liu T, Sun Y, Bai W. The Role of Epigenetics in the Chronic Sinusitis with Nasal Polyp. *Current Allergy and Asthma Reports*. 2020; 21(1): 1.
- Loh C, Veenstra GJC. The Role of Polycomb Proteins in Cell Lineage Commitment and Embryonic Development. *Epigenomes*. 2022; 6(3): 23.
- Long HK, Sims D, Heger A, Blackledge NP, Kutter C, Wright ML, Grützner F, Odom DT, Patient R, Ponting CP, Kloise RJ. Epigenetic conservation at gene regulatory elements revealed by non-methylated DNA profiling in seven vertebrates. *eLIFE*. 2013; 2: e00348.
- Lorzadeh A, Romero-Wolf M, Goel A, Jadhav U. Epigenetic Regulation of Intestinal Stem Cells and Disease: A Balancing Act of DNA and Histone Methylation. *Gastroenterology*. 2021; 160(7): 2267-2282.
- Lottin S, Adriaenssens E, Dupressoir T, Berteaux N, Montpellier C, Coll J, Dugimont T, Cury JJ. Overexpression of an ectopic H19 gene enhances the tumorigenic properties of breast cancer cells. *Carcinogenesis*. 2002; 23: 1885-1895.
- Lou DY, Dominguez I, Toselli P, Landesman-Bollag E, O'Brien C, Seldin DC. The Alpha Catalytic Subunit of Protein Kinase CK2 Is Required for Mouse Embryonic Development. *Molecular & Cellular Biology*. 2008; 28: 131-139.
- Louis D, Perry A, Reifenberger G, von Deimling A, Figarella-Branger D, Cavenee W, Ohgaki H, Wiestler O, Kleihues P, Ellison D. The 2016 World Health Organization Classification of Tumors of the Central Nervous System: a summary. *Acta Neuropathology*. 2016 ; 131: 803-820.
- Lovejoy DA, Michalec OM, Hogg DW, Wosnick DI. Role of elasmobranchs and holocephalans in understanding peptide evolution in the vertebrates: Lessons learned from gonadotropin releasing hormone (GnRH) and corticotropin releasing factor (CRF) phylogenies. *General and Comprehensive Endocrinology*. 2018; 264: 78-83.
- Lovering R, Hanson IM, Borden KL, Martin S, O'Reilly N, Evan GI, Rahman D, Pappin DJC, Trowsdale J, Freemont PS. Identification and preliminary characterization of a protein motif related to the zinc finger. *Proceedings of the National Academy of Sciences*. 1993; 90: 2112-2116.
- Luis NM, Morey L, Mejetta S, Pascual G, Janich P, Kuebler B, Roma G, Nascimento E, Frye M, Di Croce L, Benitah SA. Regulation of human epidermal stem cell proliferation and senescence requires polycomb-dependent and -independent functions of cbx4. *Cell Stem Cell*. 2011; 9: 233-246.

- Luo X, Schoch K, Jangam SV, Bhavana VH, Graves HK, Kansagra S, Jasien JM, Stong N, Keren B, Mignot C, Ravelli C; Undiagnosed Diseases Network; Bellen HJ, Wangler MF, Shashi V, Yamamoto S. Rare deleterious de novo missense variants in Rnf2/Ring2 are associated with a neurodevelopmental disorder with unique clinical features. *Human Molecular Genetics*. 2021; 30(14):1283-1292.
- Lyons DA, Talbot WS. Glial cell development and function in zebrafish. *Cold Spring Harbor Laboratory Press*. 2014; 7(2): a020586.
- Ma J, Yin H, Li M, Deng Y, Ahamd O, Qin G, He Q, Li J, Gao K, Wang B, Wu S, Wang T, Shang J. A Comprehensive Study of High Cholesterol Diet-Induced Larval Zebrafish Model: A Short-Time *In Vivo* Screening Method for Non-Alcoholic Fatty Liver Disease Drugs. *International Journal of Biological Sciences*. 2019; 15(5): 973-983.
- Maertens GN, El Messaoudi-Aubert S, Racek T, Stock JK, Nicholls J, Rodriguez-Niedenführ M, Gil J, Peters G. Several Distinct Polycomb Complexes Regulate and Co-Localize on the INK4a Tumor Suppressor Locus. *PLOS ONE*. 2009; 4(7):e6380.
- Maethner E, Garcia-Cuellar MP, Breiting C, Takacova S, Divoky V, Hess JL, Slany RK. MLL-ENL inhibits polycomb repressive complex 1 to achieve efficient transformation of hematopoietic cells. *Cell Reports*. 2013; 3: 1553-66.
- Mallen-St. Clair J, Soydaner-Azeloglu R, Lee KE, Taylor L, Livanos A, Pylayeva-Gupta Y, Miller G, Margueron R, Reinberg D, Bar-Sagi D. Ezh2 couples pancreatic regeneration to neoplastic progression. *Genes & Development*. 2012; 26(5): 439-444.
- Mandel SA, Fishman-Jacob T, Youdim MB. Targeting SKP1, an ubiquitin E3 ligase component found decreased in sporadic Parkinson's disease. *Neurodegenerative Diseases*. 2012; 10(1-4): 220-3.
- Mandel SA, Fishman-Jacob T, Youdim MBH. Modeling sporadic Parkinson's disease by silencing the ubiquitin E3 ligase component, SKP1A. *Parkinsonism Relational Disorders*. 2009; 15: 148-151.
- Marcon M, Mocelin R, Benvenuto R, Costa T, Herrmann AP, de Oliveira DL, Koakoski G, Barcellos LJG, Piato A. Environmental enrichment modulates the response to chronic stress in zebrafish. *Journal of Experimental Biology*. 2018; 221: jeb176735.
- Margueron R, Reinberg D. The Polycomb complex PRC2 and its mark life. *Nature*. 2011; 469: 343-349.
- Margueron R, Li G, Sarma K, Blais A, Zavadil J, Woodcock CL, Dynlacht BD, Reinberg D. Ezh1 and Ezh2 Maintain Repressive Chromatin through Different Mechanisms. *Molecular Cell*. 2008; 32: 503-518.
- Massner MP, Raocy J, Losordo T. Recirculating aquaculture tank production systems: management of recirculating systems. *Southern Regional Aquaculture Publication*. 1999. 452.
- Mathsyaraja H, Catchpole J, Freie B, Eastwood E, Babaeva E, Geuenich M, Cheng PF, Ayers J, Yu M, Wu N, Moorthi S, Poudel KR, Koehne A, Grady W, Houghton AM, Berger AH, Shii Y, MacPherson D, Eisenman RN. Loss of MGA repression mediated by an atypical polycomb complex promotes tumor progression and invasiveness. *Elife*. 2021; 10: e64212.
- McClure MM, McIntyre PB, McCune AR. Notes on the natural diet and habitat of eight danionin fishes , including the zebrafish, *Danio rerio*. *Journal of Fish Biology*. 2006; 69: 553-570.
- McGhee JD, Felsenfeld G. Nucleosome Structure. *Annual Review of Biochemistry*. 1980; 49: 1115-56.

- Meyer A, Ritchie PA, Witte KE. Predicting Developmental Processes from Evolutionary Patterns: A molecular Phylogeny of the Zebrafish (*Danio rerio*) and its Relatives. *Philosophical Transactions: Biological Sciences*. 349(1327): 103-111.
- Michalak EM, Milevskiy MJG, Joyce RM, Dekkers JF, Jamieson PR, Pal B, Dawson CA, Hu Y, Orkin SH, Alexander WS, Lindeman GJ, Smyth GK, Visvader JE. Canonical PRC2 function is essential for mammary gland development and affects chromatin compaction in mammary organoids. *PLOS Biology*. 2018; 16(8): e2004986.
- Mitra S, Dunphy PS, Das S, Zhu B, Luo T, McBride JW. Ehrlichia chaffeensis TRP120 Effector Targets and Recruits Host Polycomb Group Proteins for Degradation To Promote Intracellular Infection. *Infection & Immunity*. 2018; 86(4): e00845-17.
- Mladek AC, Yan H, Tian S, Decker PA, Burgenske DM, Bakken K, Hu Z, He L, Connors MA, Carlson BL, Wilson J, Bommi-Reddy A, Conery A, Eckel-Passow JE, Sarkaria JN, Kitange GJ. RBBP4-p300 axis modulates expression of genes essential for cell survival and is a potential target for therapy in glioblastoma. *Neurology & Oncology*. 2022; 24(8): 1261-1272.
- Mohammad F, Weissmann S, Leblanc B, Pandey DP, Højfeldt JW, Comet I, Zheng C, Johansen JV, Rapin N, Porse BT, Tvardovskiy A, Jensen ON, Olaciregui NG, Lavarino C, Suñol M, de Torres C, Mora J, Carcaboso AM, Helin K. EZH2 is a potential therapeutic target for H3K27M-mutant pediatric gliomas. *Nature Medicine*. 2017; 23(4): 483-492.
- Mohn F, Weber M, Rebhan M, Roloff TC, Richter J, Stadler MB, Bibel M, Schübeler D. Lineage-specific polycomb targets and de novo DNA methylation define restriction and potential of neuronal progenitors. *Molecular Cell*. 2008; 30: 755-766.
- Montgomery ND, Yee D, Chen A, Kalantry S, Chamberlain SJ, Otte AP, Magnuson T. The murine polycomb group protein Eed is required for global histone H3 lysine-27 methylation. *Current Biology*. 2005; 15: 942-947.
- Montuenga LM, Guembe L, Burrell MA, Bodegas ME, Calvo A, Sola JJ, Sesma P, Villaro AC. The diffuse endocrine system : from embryogenesis to carcinogenesis. *Progress in Histochemistry and Cytochemistry*. 2003; 38: 155-272.
- Moody RR, Lo MC, Meagher JL, Li CC, Stevers NO, Tinsley SL, Jung I, Matvekas A, Stuckey JA, Sun D. Probing the interaction between the histone methyltransferase/deacetylase subunit RBBP47 and the transcription factor BCL11A in epigenetic complexes. *Protein Structure and Folding*. 2018; 293(6): 2125-2136.
- Moore LD, Le T, Fan G. DNA Methylation and Its Basic Function. *Neuropsychopharmacology*. 2013; 38: 23-38.
- Morey L, Aloia L, Cozzuto L, Benitah SA, Di Croce L. RYBP and Cbx7 define specific biological functions of polycomb complexes in mouse embryonic stem cells. *Cell Reports*. 2013;3(1): 60-9.
- Morey L, Santanach A, Blanco E, Aloia L, Nora EP, Bruneau BG, Di Croce L. Polycomb regulates mesoderm cell fate-specification in embryonic stem cells through activation and repression mechanisms. *Cell Stem Cells*. 2015; 17: 300-315.
- Morin RD, Mendez-Lago M, Mungall AJ, Goya R, Mungall KL, Corbett RD, Johnson NA, Severson TM, Chiu R, Field M, Jackman S, Krzywinski M, Scott DW, Trinh DL, Tamura-Wells J, Li S, Firme MR, Rogic S, Griffith M, Chan S, Yakovenko O, Meryer IM, Zhao EY, Smailus D, Moksa M, Chittaranjan S, Rimsza L, Brooks-Wilson A, Spinelli JJ, Ben-Neriah S, Meissner B, Woolcock B,



- Boyle M, McDonald H, Tam A, Zhao Y, Delaney A, Zeng T, Tse K, Butterfield Y, Birol I, Holt R, Schein J, Horsman DE, Moore R, Jones SJM, Connors JM, Hirst M, Gascoyne RD, Marra MA. Frequent mutation of histone-modifying genes in non-Hodgkin lymphoma. *Nature*. 2011; 476: 298-303.
- Moscou MJ, Bogdanove AJ. A simple cipher governs DNA recognition by TAL effectors. *Science*. 2009, 326, 1501.
- Mozetta C, Pontis J, Fritsch L, Robin P, Portoso M, Proux C, Margueron R, Ait-Si-Ali S. The Histone H3 Lysine 9 Methyltransferases G9a and GLP Regulate Polycomb Repressive Complex 2-Mediated Gene Silencing. *Molecular Cell*. 2014; 53: 277-289.
- Mu W, Starmer J, Shibata Y, Yee D, Magnuson T. EZH1 in germ cells safeguards the function of PRC2 during spermatogenesis. *Developmental Biology*. 2017; 424(2): 198-207.
- Mueller T, Wullmann MF. Anatomy of neurogenesis in the early zebrafish brain. *Developmental Brain Research*. 2003; 140:137-155.
- Murzina NV, Pei XY, Zhang W, Vicente-Garcia J, Pratap JV, McLaughlin SH, Ben-Shahar TR, Verreault A, Luisi BF, Laue ED. Structural Basis for the Recognition of Histone H4 by the Histone-Chaperone RbAp46. *Structure*. 2008; 16(7): 1077-1085.
- Musselman CA, Avvkumov N, Watanabe R, Abraham CG, Lalonde ME, Hong Z, Allen C, Roy S, Nuñez JK, Nickoloff J, Kulesza CA, Yasui A, Côté J, Kutateladze TG. Molecular basis for H3K36me3 recognition by the Tudor domain of PHF1. *Nature Structural Molecular Biology*. 2012; 19: 1266-1272.
- Nakatani Y, Takeda H, Kohara Y, Morishita S. Reconstruction of the vertebrate ancestral genome reveals dynamic genome reorganization in early vertebrates. *Genome Research*. 2007; 14: 1254-1265.
- Naqvi AAT, Rizvi SAM, Hassan I. Pan-cancer analysis of Chromobox (CBX) genes for prognostic significance and cancer classification. *Biochimica et Biophysica Acta – Molecular Basis of Disease*. 2023; 1869(1): 166561.
- Nasevicius A, Ekker SC. Effective targeted gene 'Knockdown' in zebrafish. *Nature genetics*. 2000; 26: 216-220.
- Nawaji T, Yamashita N, Umeda H, Zhang S, Mizoguchi N, Seki M, Kitazawa T, Teraoka H. Cytochrome P450 Expression and Chemical Metabolic Activity before Full Liver Development in Zebrafish. *Pharmaceuticals*. 2020; 13(12): 456.
- Ng ANY, de Jong-Curtain TA, Mawdsley DJ, White SJ, Shin J, Appel B, Dong PDS, Stainier DYR, Heath JK. Formation of the digestive system in zebrafish: III. Intestinal epithelium morphogenesis. *Developmental Biology*. 2005; 286:114-135.
- Ng D, Thakker N, Corcoran CM, Donnai D, Perveen R, Schneider A, Hadley DW, Tiftt C, Zhang L, Wilkie AOM, van der Smagt, JJ, Gorlin RJ, Burgess SM, Bardwell VJ, Black GCM, Biesecker LG. Oculofaciocardiodental and Lenz microphthalmia syndromes result from distinct classes of mutations in BCOR. *Nature Genetics*. 2004; 36(4): 411-6.
- Ng J, Hart CM, Morgan K, Simon JA. A Drosophila ESCE(Z) protein complex is distinct from other polycomb group complexes and contains covalently modified ESC. *Molecular & Cellular Biology*. 2000 ; 20: 3069Y3078.

- Noel ES, Casal-Sueiro A, Busch-Nentwich E, Verkade H, Dong PD, Stemple DL, Ober EA. Organ-specific requirements for Hdac1 in liver and pancreas formation. *Developmental Biology*. 2008; 322: 237-250.
- Noël ES, Casal-Sueiro A, Busch-Nentwich E, Verkade H, Dong PD, Stemple DL, Ober EA. Organ-specific requirements for Hdac1 in liver and pancreas formation. *Developmental Biology*. 2008; 322(2): 237-250.
- Nomura M, Takihara Y, Shimada K. Isolation and characterization of retinoic acid-inducible cDNA clones in F9 cells: One of the early inducible clones encodes a novel protein sharing several highly homologous regions with a *Drosophila Polyhomeotic* protein. *Differentiation*. 1994; 57: 39-50.
- O'Carroll D, Erhardt S, Pagani M, Barton SC, Surani MA, Jenuwein T. The Polycomb-Group *Eh2* Is Required for Early Mouse Development. *Molecular and Cellular Biology*. 2001; 21(13): 4330-4336.
- Obier N, Lin Q, Cauchy P, Hornich V, Zenke M, Becker M, Muller AM. Polycomb protein EED is required for silencing of pluripotency genes upon ESC differentiation. *Stem Cell Reviews and Reports*. 2015 ; 11(1): 50-61.
- Oguro H, Yuan J, Ichikawa H, Ikawa T, Yamazaki S, Kawamoto H, Nakauchi H, Iwama A. Poised Lineage Specification in Multipotential Hematopoietic Stem and Progenitor Cells by the Polycomb Protein Bmi1. *Cell Stem Cell*. 2010; 6: 279–286.
- Ohno S. Evolution by gene duplication. *Springer-Verlag*. New York. 1970; 160 pp.
- Oliviero G, Brien GL, Watson A, Streubel G, Jerman E, Andrews D, Doyle B, Munawar N, Wynne K, Crean J, Bracken AP, Cagney G. Dynamic protein interactions of the Polycomb repressive complex 2 during differentiation of pluripotent cells. *Molecular Cell Proteomics*. 2016; 14: 3450-3460.
- Oppel F, Ki DH, Zimmermann MW, Ross KN, Tao T, Shi H, *et al.* *suz12* inactivation in p53 and nf1 deficient zebrafish accelerates the onset of MPNSTs and expands the spectrum of tumor types to include adenocarcinoma, leukemia, and soft tissue sarcoma. *Disease Models & Mechanisms*. 2020; 13(8): dmm.042341.
- Pack M, Solnica-Krezel L, Malicki  $\mu$ J, Neuhauss SCF, Schier AF, Stemple DL, Driever W, Fishman MC. Mutations affecting development of zebrafish digestive organs. *Development*. 1996; 123: 312-318.
- Pan L, Hu Y, Qian C, Yao Y, Wang S, Shi W, Xu T. RNF2 mediates pulmonary fibroblasts activation and proliferation by regulating mTOR and p16-CDK4-Rb1 signaling pathway. *Inflammation Research*. 2022; 71(10-11): 1283-1303.
- Pan QS, Fang ZP, Zhao YX. Immunocytochemical identification and localization of APUD cells in the gut of seven stomachless teleost fishes. *World Journal of Gastroenterology*. 2000 ; 6 : 96-101.
- Papan C, Campos-Ortega JA. On the formation of the neural keel and neural tube in the zebrafish *Danio (Brachydanio) rerio*. *Roux's Arch Developmental Biology*. 1994; 203: 178-186.
- Parichy DM, Elizondo MR, Mills MG, Gordon TN, Engeszer RE. Normal Table of Post-Embryonic Zebrafish Development: Staging by Externally Visible Anatomy of the Living Fish. *Developmental Dynamics*. 2009; 238(12): 2975-3015.
- Park SH, Fong KW, Mong E, Martin MC, Schiltz GE, Yu J. Going beyond polycomb: EZH2 functions in prostate cancer. *Oncogene*. 2021; 40: 5788-5798.

- Pasini D, Bracken AP, Jensen MR, Lazzerini Denchi E, Helin K. Suz12 is essential for mouse development and for EZH2 histone methyltransferase activity. *The EMBO Journal*. 2004; 23: 4061-4071.
- Pasini D, Malatesta M, Jung HR, Walfridsson J, Willer A, Olsson, Skotte J, Wutz A, Porse B, Jensen ON, Helin K. Characterization of an antagonistic switch between histone H3 lysine 27 methylation and acetylation in the transcriptional regulation of Polycomb group target genes. *Nucleic Acids Research*. 2010; 38(15): 4958-4969.
- Patil N, Kelly ME, Yeboa DN, Buerki RA, Cioffi G, Balaji S, Ostrom QT, Kruchko C, Barnholtz-Sloan JS. Epidemiology of brainstem high-grade gliomas in children and adolescents in the United States, 2000-2017. *Neur-Oncology*. 2021; 23(6): 990-998.
- Patton EE, Mitchell DL, Nairn RS. Genetic and environmental melanoma models in fish. *Pigment Cell & Melanoma Research*. 2010; 23(3): 314-337.
- Pearce JJ, Singh PB, Gaunt SJ. The mouse has a Polycomb-like chromobox gene. *Development*. 1992; 114(4): 921-929.
- Peng X, Feng G, Zhang Y, Sun Y. PRC1 Stabilizes Cardiac Contraction by Regulating Cardiac Sarcomere Assembly and Cardiac Conduction System Construction. *International Journal of Molecular Sciences*. 2021; 22(21): 11368.
- Pereira JD, Sansom SN, Smith J, Dobenecker M-W, Tarakhovsky A, Livesey FJ. Ezh2, the histone methyltransferase of PRC2, regulates the balance between self-renewal and differentiation in the cerebral cortex. *Proceeding of the National Academy of Sciences U S A*. 2010; 107: 15957-15962.
- Perino M, van Mierlo G, Karemaker ID, van Genesen S, Vermeulen M, Marks H, van Heeringen SJ, Veenstra GJC. MTF2 recruits Polycomb Repressive Complex 2 by helical-shape-selective DNA binding. *Nature Genetics*. 2018; 50: 1002–1010.
- Pesavento PA, Agnew D, Keel MK, Woolard KD. Cancer in wildlife: patterns of emergence. *Nature Reviews Cancers*. 2018; 18: 646—661.
- Peters AHS, Kubicek S, Mechtler K, O'Sullivan RJ, Derijck AA, Perez-Burrigos L, Kohlmaier A, Opravil S, Tachibana M, Shinkai Y, Martens JH, Jenuwein T. Partitioning and plasticity of repressive histone methylation states in mammalian chromatin. *Molecular Cell*. 2003; 12:1577-1589.
- Peterson AJ, Ktba M, Bornemann D, Morgan K, Brock HW, Simon J. A Domain Shared by the Polycomb Group Proteins Scm and ph Mediates Heterotypic and Homotypic Interactions. *Molecular and Cellular Biology*. 1997; 17(11): 6683-6692.
- Pierce SB, Stewart MD, Gulsuner S, Walsh T, Dhall A, McClellan JM, Klevit RE, King MC. De novo mutation in *RING1* with epigenetic effects on neurodevelopment. *Proceeding of the National Academy of Sciences U S A*. 2018; 115(7): 1558-1563.
- Pirity MK, Locker J, Schreiber-Agus N. Rybp/DEDAF is required for early postimplantation and for central nervous system development. *Molecular & Cellular Biology*. 2005; 25: 7193–7202.
- Pirity MK, Wang WL, Wolf LV, Tamm ER, Schreiber-Agus N, Cvekl A. Rybp, a polycomb complex-associated protein, is required for mouse eye development. *BMC Developmental Biology*. 2007; 7:39.
- Pitcher TJ, Parrish JK. Functions of shoaling behaviour in teleosts. In: *Behaviour of teleost fishes*. (ed. T. J. Pitcher), pp. 363-439. Chapman and Hall, London. 1993.

- Piunti A, Shilatifard A. The roles of Polycomb repressive complexes in mammalian development and cancer. *Molecular Cell Biology*. 2021; 22: 326-345.
- Plehn M. Ueber geschwülste bei kaltblütern. *Z. Krebs*. 1906; 4: 525–564
- Pliss GB, Zabezhinski MA, Petrov AS, Khudoley VV. Peculiarities of N-nitramines carcinogenic action. *Arch Geschwulstforsch*. 1982; 52: 629-634.
- Poepsel S, Kasinath V, Nogales E. Cryo-EM structures of PRC2 simultaneously engaged with two functionally distinct nucleosomes. *Nature Structural and Molecular Biology*. 2018; 25: 154-162.
- Pohlers M, Truss M, Frede U, Scholz A, Strehle M, Kuban R-J, Hoffmann B, Morkel M, Birchmeier C, Hagemeyer C. A Role for E2F6 in the Restriction of Male-Germ-Cell-Specific Gene Expression. *Current Biology*. 2005; 15: 1051–1057.
- Polyak K, Haviv I, Campbell IG. Co-evolution of tumor cells and their microenvironment. *Trends in Genetics*. 2009; 25(1): 30-38.
- Postic C, Girard J. The role of the lipogenic pathway in the development of hepatic steatosis. *Diabetes & Metabolism*. 2008; 34(6): 643-648.
- Prokopuk L, Stringer JM, White CR, Vossen RH, White SJ, Cohen ASA, Gibson WT, Western PS. Loss of maternal EED results in postnatal overgrowth. *Clinical Epigenetics*. 2018; 10(1):95.
- Puda A, Milosevic JD, Berg T, Klampfl T, Harutyunyan AS, Gisslinger B, Rumi E, Pietra D, Malcovati L, Elena C, Doubek M, Steurer M, Tosic N, Pavlovic S, Guglielmelli P, Pieri L, Vannucchi AM, Gisslinger H, Cazzola M, Kralovics R. Frequent deletions of JARID2 in leukemic transformation of chronic myeloid malignancies. *American Journal of Hematology*. 2012; 87(3): 245-50.
- Qin J, Whyte WA, Anderssen E, Apostolou E, Chen HH, Akbarian S, Bronson RT, Hochedlinger K, Ramaswamy S, Young RA, Hock H. The polycomb group protein L3mbtl2 assembles an atypical PRC1-family complex that is essential in pluripotent stem cells and early development. *Cell Stem Cell*. 2012; 11(3): 319-32.
- Qu H, Fang X (2013) A brief review on the human encyclopedia of DNA elements (ENCODE) project. *Genomics, Proteomics & Bioinformatics*. 11:135–141.
- Raaphorst FM, Vermeer M, Fieret E, Blokzijl T, Dukers D, Sewalt RG, Otte AP, Willemze R, Meijer CJ. Site-specific expression of polycomb-group genes encoding the HPC-HPH/PRC1 complex in clinically defined primary nodal and cutaneous large B-cell lymphomas. *American Journal of Pathology*. 2004; 164(2): 533-42.
- Raas MWD, Zijlmans DW, Vermeulen M, Marks H. There is another: H3K27me3-mediated genomic imprinting. *Trends in Genetics*. 2022; 38(1): 82-96.
- Ragge N, Isidor B, Bitoun P, Odent S, Giurgea I, Cogné B, Deb W, Vincent M, Le Gall J, Morton J, Lim D; DDD Study; Le Meur G, Zazo Seco C, Zafeiropoulou D, Bax D, Zwijnenburg P, Arteche A, Swafiri ST, Cleaver R, McEntagart M, Kini U, Newman W, Ayuso C, Corton M, Herenger Y, Jeanne M, Calvas P, Chassaing N. Expanding the phenotype of the X-linked BCOR microphthalmia syndromes. *Human Genetics*. 2019; 138(8-9): 1051-1069.
- Rai K, Chidester S, Zavala CV, Manos EJ, James SR, Karpf AR, Jones DA, Cairns BR. Dnmt2 functions in the cytoplasm to promote liver, brain, and retina development in zebrafish. *Genes & Development*. 2007; 21: 261-266.

- Raj B, Farrell JA, Liu J, El Kholtei J, Carte AN, Acedo JN, Du LY, McKenna A, Relić D, Leslie JM, Schier AF. Emergence of Neuronal Diversity during Vertebrate Brain Development. *Neuroresource*. 2020; 108(6): 1058-1074.
- Reingi J, Ruge F, Howard M, Ringrose LA. A theoretical model of Polycomb/Trithorax action unites stable epigenetic memory and dynamic regulation. *Nature Communications*. 2020; 11: 4782.
- Ren X, Hu B, Song M, Ding Z, Dang Y, Liu Z, Zhang W, Ji Q, Ren R, Ding J, Chan P, Jiang C, Ye K, Qu J, Tang F, Liu GH. Maintenance of Nucleolar Homeostasis by CBX4 Alleviates Senescence and Osteoarthritis. *Cell Rep*. 2019 Mar 26;26(13):3643-3656.e7. doi: 10.1016/j.celrep.2019.02.088. PMID: 30917318.
- Ren X, Kerppola TK. REST Interacts with Cbx Proteins and Regulates Polycomb Repressive Complex 1 Occupancy at RE1 Elements. *Molecular and Cellular Biology*. 2011; 31(10):2100-10.
- Ren Z, Ahn JH, Liu H, Tsai YH, Bhanu NV, Koss B, Allison DF, Ma A, Storey AJ, Wang P, Mackintosh SG, Edmondson RD, Groen RWJ, Martens AC, Garcia BA, Tackett AJ, Jin J, Cai L, Zheng D, Wang GG. PHF19 promotes multiple myeloma tumorigenicity through PRC2 activation and broad H3K27me3 domain formation. *Blood*. 2019; 134(14): 1176-1189.
- Rhodes CT, Sandstrom R, Huang SA, Wang Y, Schotta G, Berger M, Lin CA. Cross-species analyses unravel the complexity of H3K27me3 and H4K20me3 in the context of neural stem progenitor cells. *Neuroepigenetics*. 2016; 6:10-25.
- Richards JG. Chapter : BONY FISHES In: Farrell AP, Cech JJ, Richards JG, Stevens ED. Encyclopedia of Fish Physiology. *Elsevier*. London, UK. 2011; 2163 pp - p1819-1823.
- Rikin A, Evans T. The tbx/bHLH transcription factor *mga* regulates *gata4* and organogenesis. *Developmental Dynamics*. 2010; 239(2): 535-47.
- Rinn JL, Kertesz M, Wang JK, Squazzo SL, Xu X, Bruggmann SA, Goodnough LH, Helms JA, Farnham PJ, Segal E, Chang HY. Functional Demarcation of Active and Silent Chromatin Domains in Human *HOX* Loci by Noncoding RNAs. *Cell*. 2007; 129: 1311-1323.
- Roberts AC, Bill BR, Glanzman DL. Learning and memory in zebrafish larvae. *Frontiers in Neural Circuits*. 2013; 7:126.
- Rodin SN, Riggs AD. Epigenetic Silencing May Aid Evolution by Gene Duplication. *Journal of Molecular Evolution*. 2003; 56: 718-729.
- Roffey SE, Litchfield DW. CK2 Regulation: Perspectives in 2021. *Biomedicines*. 2021; 9(10): 1361.
- Román-Trufero, Méndez-Gómez HR, Pérez C, Hijikata A, Fujimura Y-I, Endo T, Koseki H, Vicario-Abejón C, Vidal M. Maintenance of Undifferentiated State and Self-Renewal of Embryonic Neural Stem Cells by Polycomb Protein Ring1B. *Stem Cells*. 2009; 27(7): 1559-1570.
- Rose NR, King HW, Blackledge NP, Fursova NA, Ember KJI, Fischer R, Kessler BM, Klose RJ. RYBP stimulates PRC1 to shape chromatin-based communication between Polycomb repressive complexes. *eLife*. 2016; 5:e18591.
- Ross K, Sedello AK, Todd GP, Paszkowski-Rogacz M, Bird AW, Ding L, Grinenko T, Behrens K, Hubner N, Mann M, Waskow C, Stocking C, Buhholz F. Polycomb group ring finger 1 cooperates with Runx1 in regulating differentiation and self-renewal of hematopoietic cells. *Blood*. 2012; 119: 4152-4161.

- Rothberg JLM, Maganti HB, Jrade H, Porter CJ, Palidwor GA, Cafariello C, Battaion HL, Khan ST, Perkin TJ, Paulson RF, Ito CY, Stanford WL. Mtf2-PRC2 control of canonical Wnt signaling is required for definitive erythropoiesis. *Cell Discovery*. 2018; 4:21.
- Rougeot J, Chrispijn ND, Aben M, Elurbe DM, Andralojc KM, Murphy PJ, Jansen PWTC, Vermeulen M, Cairns BR, Kamminga LM. Maintenance of spatial gene expression by Polycomb-mediated repression after formation of a vertebrate bodyplan. *Development*. 2019; 146(19): dev.178590.
- San B, Chrispijn ND, Wittkopp N, van Heeringen SJ, Lagendijk AK, Abe M, Bakkers J, Ketting RF, Kamminga LM. Normal formation of a vertebrate body plan and loss of tissue maintenance in the absence of ezh2. *Scientific Reports*. 2016; 6(1): 24658.
- Sánchez-Beato M, Sánchez E, García JF, Pérez-Rosado A, Montoya MC, Fraga M, Artiga MJ, Navarrete M, Abraira V, Morente M, Esteller M, Koseki H, Vidal M, Piris MA. Abnormal PcG protein expression in Hodgkin's lymphoma. Relation with E2F6 and NFkappaB transcription factors. *Journal of Pathology*. 2004; 204(5): 528-37.
- Sánchez-Beato M, Sánchez E, González-Carreró J, Morente M, Díez A, SánchezVerde L, Martín MC, Cigudosa JC, Vidal M, Piris MA. Variability in the expression of polycomb proteins in different normal and tumoral tissues. A pilot study using tissue microarrays, *Modern Pathology*. 2006; 19: 684–694.
- Sanchez-Jimeno C, Blanco-Kelly F, López-Grondona F, Losada-Del Pozo R, Moreno B, Rodrigo-Moreno M, Martinez-Cayuelas E, Riveiro-Alvarez R, Fenollar-Cortés M, Ayuso C, Rodríguez de Alba M, Lorda-Sanchez I, Almoguera B. Attention Deficit Hyperactivity and Autism Spectrum Disorders as the Core Symptoms of AUTS2 Syndrome: Description of Five New Patients and Update of the Frequency of Manifestations and Genotype-Phenotype Correlation. *Genes (Basel)*. 2021; 12(9): 1360.
- Sandhya P, Danda D. Role of vacuolar ATPase and Skp1 in Sjögren's syndrome. *Medical Hypotheses*. 2014; 82(3): 319-25.
- Santoro SW, Dulac C. Histone variants and cellular plasticity. *Trends in Genetics*. 2015; 31(9): 516-527.
- Sanulli S, Justin N, Teissandier A, Ancelin K, Portoso M, Caron M, Michaud A, Lombard B, da Rocha ST, Offer J, Leow D, Servant N, Wassef M, Burlina F, Gamblin SJ, Heard E, Margueron R. Jarid2 Methylation via the PRC2 Complex Regulates 3K27me3 Deposition during Cell Differentiation. *Molecular Cell*. 2015; 57: 769-783.
- Sanulli S, Justin N, Teissandier A, Ancelin K, Portoso M, Caron M, Michaud A, Lombard B, da Rocha ST, Offer J, Leow D, Servant N, Wassef M, Burlina F, Gamblin SJ, Heard E, Margueron R. Jarid2 Methylation via the PRC2 Complex Regulates H3K27me3 Deposition during Cell Differentiation. *Molecular Cell*. 2015; 57(5): 769-783.
- Sasai Y. Identifying the Missing Links: Genes that Connect Neural Induction and Primary Neurogenesis in Vertebrate Embryos. *Neuron*. 1998; 21: 455-458.
- Scelfo A, Fernández-Pérez D, Tamburri S, Zanotti M, Lavarone E, Soldi M, Bonaldi T, Ferrari KJ, Pasini D. Functional Landscape of PCGF Proteins Reveals Both RING1A/B-Dependent-and RING1A/B-Independent-Specific Activities. *Molecular Cell*. 2019; 74(5):1037-1052.

- Schaaf CA, Misulovin Z, Gause M, Koenig A, Gohara DW, Watson A, Dorsett D. Cohesin and Polycomb Proteins Functionally Interact to Control Transcription at Silenced and Active Genes. *PLoS Genetics*. 2013; 9: e1003560.
- Scheer N, Campos-Ortega JA. Use of the Gal4-UAS technique for targeted gene expression in the zebrafish. *Mechanisms of Development*. 1999, 80, 153–158.
- Schmidt R, Strähle U, Scholpp S. Neurogenesis in zebrafish – from embryo to adult. *Neural Development*. 2013; 8: 3.
- Schoeftner S, Sengupta AK, Kubicek S, Mechtler K, Spahn L, Koseki H, Jenuwein T, Wutz A. Recruitment of PRC1 function at the initiation of X inactivation independent of PRC2 and silencing. *EMBO Journal*. 2006; 25(13): 3110-22.
- Schuettengruber B, Bourbon H-M, Di Croce L, Cavalli G. Genome Regulation by Polycomb and Trithorax: 70 Years and Counting. *Cells*. 2017; 171: 34-57
- Schultz LE, Haltom JA, Almeida MP, Wierson WA, Solin SL, Weiss TJ, *et al*. Epigenetic regulators Rbbp4 and Hdac1 are overexpressed in a zebrafish model of RB1 embryonal brain tumor, and are required for neural progenitor survival and proliferation. *Disease Models & Mechanisms*. 2018; 11(6): dmm034124.
- Schumacher A, Faust C, Magnuson T. Positional cloning of a global regulator of anterior-posterior patterning in mice. *Nature*. 1996; 383 : 250-253.
- Schuppan D, Surabattula R, Wang XY. Determinants of Fibrosis Progression and Regression in NASH. *Journal of Hepatology*. 2018; 68(2): 238-250.
- Scott CL, Gil J, Hernando E, Teruya-Feldstein J, Narita M, Martinez D, Visakorpi T, Mu D, Cordon-Cardo C, Peters G, Beach D, Lowe SW. Role of the chromobox protein CBX7 in lymphomagenesis. *Proceeding of the National Academy of Science*. 2007; 104(13); 5389-5394.
- Segovia C, José-2nériz ES, Munera-Maravilla E, Martínez-Fernández M, Garate L, Miranda E, Vilas-Zornoza A, Lodewijk I, Rubio C, Segrelles C, Valcárcel V, Rabal O, Casares N, Bernardini A, Suarez-Cabrera C, López-Calderón FF, Fortes P, Casado JA, Dueñas M, Villacampa F, Lasarte JJ, Guerrero-Ramos F, de Velasco G, Oyarzabal J, Castellano D, Agirre X, Prósper F, Paramio JM. Inhibition of a G9a/DNMT network triggers immune-mediated bladder cancer regression. *Nature medicine*. 2019; 25: 1073-1081.
- Senthilkumar R, Mishra rK. Novel motifs distinguish multiple homologues of *Polycomb* in vertebrates: expansion and diversification of the epigenetic toolkit. *BMC Genomics*. 2009; 10:549.
- Serittrakul P, Gross JM. Genetic and epigenetic control of retinal development in zebrafish. *Current Opinion in Neurobiology*. 2019; 59: 120-127.
- Shao Z, Raible F, Mollaaghababa R, Guyon JR, Wu CT, Bender W, Kingstone RE. Stabilization of chromatin structure by PRC1, a Polycomb complex. *Cell*. 1999; 98: 37Y46.
- Shema E, Jones D, Shores N, Donohue L, Ram o, Bernstein BE. Single-molecule decoding of combinatorially modified nucleosomes. *Sciences*. 2016 ; 352 : 717-721.).
- Shen X, Liu Y, Hsu YJ, Fujiwara Y, Kim J, Mao X, Yuan GC, Orkin SH. EZH1 mediates methylation on histone H3 lysine 27 and complements EZH2 in maintaining stem cell identity and executing pluripotency. *Molecular Cell*. 2008; 32(4): 491-502.

- Shen-Li H, O'Hagan RC, Hou H, Horner JW, Lee HW, DePinho RA. Essential role for Max in early embryonic growth and development. *Genes & Development*. 2000; 14: 17–22.
- Sher F, Boddeke E, Olah M, Copray S. Dynamic changes in Ezh2 gene occupancy underlie its involvement in neural stem cell self-renewal and differentiation towards oligodendrocytes. *PLOS ONE*. 2012; 7:e40399.
- Shi B, Zhang J, Heng J, Gong J, Zhang T, Li P, Sun BF, Yang Y, Zhang N, Zhao YL, Wang HL, Liu F, Zhang QC, Yang YG. RNA structural dynamics regulate early embryogenesis through controlling transcriptome fate and function. *Genome Biology*. 2020; 21: 20.
- Shin D, Shin CH, Tucker J, Ober EA, Rentzsch F, Poss KD, Hammerschmidt M, Mullins MC, Stainier DYR. Bmp and Fgf signaling are essential for liver specification in zebrafish. *Development*. 2007; 134: 2041-2050.
- Shirai M, Osugi T, Koga H, Kaji Y, Takimoto E, Komuro I, Hara J, Miwa T, Yamauchi-Takahara K, Takihara Y. The Polycomb-group gene *Rae28* sustains *Nkx2.5/Csx* expression and is essential for cardiac morphogenesis. *Journal of Clinical Investigations*. 2002; 110(2): 177-184.
- Shumake J, Ilando A, Scheich H, Wetzel W, Ohl FW. Differential neuromodulation and retrieval of avoidance learning by the lateral habenula and ventral tegmental area. *Journal of Neuroscience*. 2010; 30: 5876-5883.
- Silva J, Garcia V, Garcia JM, Pena C, Dominguez G, Diaz R, Lorenzo Y, Hurtado A, Sanchez A, Bonilla F. Circulating Bmi-1 mRNA as a possible prognostic factor for advanced breast cancer patients. *Breast Cancer Research*. 2007; 9: R55.
- Silveira AB, Kasper LH, Fan Y, Jin H, Wu G, Shaw TI, Zhu X, Larson JD, Easton J, Shao Y, Yergeau DA, Rosencrance C, Boggs K, Rusch MC, Ding L, Zhang J, Finkelstein D, Noyes RM, Russell BL, Xu B, Broniscer A, Wetmore C, Pounds SB, Ellison DW, Zhang J, Baker SJ. H3.3 K27M depletion increases differentiation and extends latency of diffuse intrinsic pontine glioma growth in vivo. *Acta Neuropathology*. 2019; 137: 637–655.
- Simoës da Silva CJ, Simón R, Busturia A. Epigenetic and non-epigenetic functions of the RYBP protein in development and disease. *Mechanisms of Ageing and Development*. 2018;174:111-120.
- Skarnes WC, Rosen B, West AP, Koutourakis M, Bushell W, Iyer V, Mujica AO, Thomas M, Harrow J, Cox T, Jackson D, Severin J, Biggs P, Fu J, Nefedov M, de Jong PJ, Stewart AF, Bradley A. A conditional knockout resource for the genome-wide study of mouse gene function. *Nature*. 2011; 474: 337–342.
- Slack JM. Developmental biology of the pancreas. *Development*. 1995; 121: 1569-1580.
- Smits AH, Jansen PWTC, Poser I, Hyman AA, Vermeulen M. Stoichiometry of chromatin-associated protein complexes revealed by label-free quantitative mass spectrometry-based proteomics. *Nucleic Acids Research*. 2013; 41(1): e28.
- Sneeringer CJ, Scott MP, Kuntz KW, Knutson SK, Pollock RM, Richon RM, Copeland RA. Coordinated activities of wild-type plus mutant EZH2 drive tumor-associated hypertrimethylation of lysine 27 on histone H3 (H3K27) in human B-cell lymphomas. *Proceeding of the National Academy of Sciences USA*. 2010; 107: 20980-20985.
- Soto RA, Najia MAT, Hachimi M, Frame JM, Yette GA, Lummertz da Rocha E, Stankunas K, Daley GQ, North TE. Sequential regulation of hemogenic fate and hematopoietic stem and progenitor cell formation from arterial endothelium by Ezh1/2. *Stem Cell Reports*. 2021; 16(7): 1718-34.



- Spence R, Fatema MK, Ellis S, Ahmed ZF, Smith C. The diet, growth and recruitment of wild zebrafish (*Danio rerio*) in Bangladesh. *Journal of Fish Biology*. 2007; 71: 304-309.
- Spence R, Gerlach G, Lawrence C, Smith C. The behaviour and ecology of the zebrafish. *Danio rerio*. 2008; 83(1): 13-34.
- Sproll P, Eid W, Biason-Lauber A. CBX2-dependent transcriptional landscape: implications for human sex development and its defects. *Scientific Reports*. 2019; 9(1): 16552.
- Stanton MF. Diethylnitrosamine-Induced Hepatic Degeneration and Neoplasia in the Aquarium Fish, *Brachydanio rerio*. *JNCI: Journal of the National Cancer Institute*. 1965; 34(1): 117-130.
- Steiner LA, Schulz Vp, Maksimova Y, Wong C, Gallagher PG. Patterns of histone H3 lysine 27 monomethylation and erythroid cell type-specific gene expression. *Journal of Biological Chemistry*. 2011; 286(45): 39457-65.
- Stevens CH, Reed BT, Hawkins P. Enrichment for Laboratory Zebrafish-A review of the Evidence and the Challenges. *Animals*. 2021; 11(3): 698.
- Stielow B, Finkernagel F, Stiewe T, Nist A, Suske G. MGA, L3MBTL2 and E2F6 determine genomic binning of the non-canonical Polycomb repressive complex PRC1.6. *PLOS Genetics*. 2018; 14(1): e1007193.
- Stielow C, Stielow B, Finkernagel F, Maren Scharfe, Jarek M, Suske G. SUMOylation of the polycomb group protein L3MBTL2 facilitates repression of its target genes. *Nucleic Acids Research*. 2014; 42(5): 3044-58.
- Stigloher C, Chapouton P, Adolf B, Bally-Cuif L. Identification of neural progenitor pools by E(Spl) factors in the embryonic and adult brain. *Brain Research Bulletin*. 2008; 75: 266-273.
- Stock JK, Giadrossi S, Casanova M, Brookes E, Vidal M, Koski H, Brockdorff N, Fisher A, Pombo A. Ring1-mediated ubiquitination of H2A restrains poised RNA polymerase II at bivalent genes in mouse ES cells. *Nature Cell Biology*. 2007; 9(12): 2468-2473.
- Storre J, Elsässer H-P, Fuchs M, Ullmann D, Livingston DM, Gaubatz S. Homeotic transformations of the axial skeleton that accompany a targeted deletion of E2f6. *EMBO Reports*. 2002; 3: 695–700.
- Streisinger G, Walker C, Dower N, Knauber D, Singer F. Production of clones of homozygous diploid zebra fish (*Brachydanio rerio*). *Nature*. 1981; 291: 293–296.
- Streubel G, Watson A, Jammula SG, Scelfo A, Fitzpatrick DJ, Oliviero G, McCole R, Conway E, Glancy E, Negri GL, Dillon E, Wynne K, Pasini D, Krogan NJ, Bracken AP, Cagney G. The H3K36me2 Methyltransferase Nsd1 Demarcates PRC2-Mediated H3K27me2 and H3K27me3 Domains in Embryonic Stem Cells. *Molecular Cell*. 2018; 70(2): 371-379.
- Strohman RC. Linear genetics, non-linear epigenetics: Complementary approaches to understanding complex diseases. *Integrative Physiological and Behavioral Science*. 1995; 30(4): 273-282.
- Struhl G. A gene product required for correct initiation of segmental determination in *Drosophila*. *Nature*. 1981; 283(3):36-41.
- Su M, Xiao Y, Tang J, Wu J, Ma J, Tian B, Zhou Y, Wang H, DesongY, Liao Q, Wang W. Role of lncRNA and EZH2 Interaction/Regulatory Network in Lung Cancer. *Journal of Cancer*. 2018; 9: 4156-4165.

- Su W, Han HH, Wang Y, Zhang B, Zhou B, Cheng Y, Rumandla A, Gurrapu S, Chakraborty G, Su J, Yang G, Liang X, Wang G, Rosen N, Scher HI, Ouerfelli O, Giancotti F. The Polycomb Repressor Complex 1 Drives Double Negative Prostate Cancer Metastasis by Coordinating Stemness and Immune Suppression. *Cancer Cell*. 2019; 36(2): 139-155.
- Sun J, Wang J, He L, Lin Y, Wu J. Knockdown of polycomb-group RING finger 6 modulates mouse male germ cell differentiation in vitro. *Cell. Physiol. Biochem*. 2015; 35: 339–352.
- Sun XJ, Xu PF, Zhou T, Hu M, Fu CT, Zhang Y, Jin Y, Chen Y, Chen SJ, Huang QH, Liu TX, Chen Z. Genome-Wide Survey and Developmental Expression Mapping of Zebrafish SET Domain-Containing Genes. *PLOS ONE*. 2008; 3(1): e1499.
- Suster ML, Kikuta H, Urasaki A, Asakawa K, Kawakami K. Transgenesis in Zebrafish with the Tol2 Transposon System. In: Cartwright, E. (eds) Transgenesis Techniques. *Methods in Molecular Biology*. 2009; vol 561: 41-63. Humana Press.
- Tachibana M, Ueda J, Fukuda M, Takeda N, Ohta T, Iwanari H, Sakihama T, Kodama T, Hamakubo T, Shinkai Y. Histone methyltransferases G9a and GLP form heteromeric complexes and are both crucial for methylation of euchromatin at H3-K9. *Genes & Development*. 2005; 19(7): 815-826.
- Takawa M, Masuda K, Kunizaki M, Daigo Y, Takagi K, Iwai Y, Cho HS, Toyokawa G, Yamane Y, Maejima K, Field HI, Kobayashi T, Akasu T, Sugiyama M, Tsuchiya E, Atomi Y, Ponder BAJ, Nakamura Y, Hamamoto R. Validation of the histone methyltransferase EZH2 as a therapeutic target for various types of human cancer and as a prognostic marker. *Cancer Science*. 2011; 102(7): 1298-1305.
- Takeuchi T, Kojima M, Nakajima K, Kondo S. Jumonji gene is essential for the neurulation and cardiac development of mouse embryos with C3H/He background. *Mechanisms of Development*. 1999 ; 86: 29-38.
- Takeuchi T, Yamazaki Y, Katoh-Fukui Y, Tsuchiya R, Kondo S, Motoyama J, Higashinakagawa T. Gene trap capture of a novel mouse gene, jumonji, required for neural tube formation. *Genes & Development*. 1995 ; 9:1211-1222.
- Takahara Y, Tomotsune D, Shirai M, Katoh-Fukui Y, Nishii K, Motaleb MA, Nomura M, Tsuchiya R, Fujita Y, Shibata Y, Higashinakagawa T, Shimada K. Targeted disruption of the mouse homologue of the Drosophila polyhomeotic gene leads to altered anteroposterior patterning and neural crest defects. *Development*. 1997; 124(19): 3673-82.
- Takke C, Dornseifer P, Weizsacker E, Campos-Ortega JA. Her4, a zebrafish homologue of the Drosophila neurogenic gene E(spl), is a target of NOTCH signalling. *Development*. 1999; 126(9): 1811-1821.
- Tan J, Jones M, Koseki H, Nakayama M, Muntean AG, Maillard I, Hess JL. CBX8, a Polycomb Group Protein, Is Essential for MLL-AF9-Induced Leukemogenesis. *Cancer Cell*. 2011; 20(5): 563-575.
- Tan J-Z, Yan Y, Wang XX, Jiang Y, Xu E. EZH2 : biology, disease, and structure-based drug. *Acta Pharmacologica Sinica*. 2014; 35: 161-174.
- Tang C, Lv Y, Ding K, Cao Y, Ma Z, Yang L, Zhang Q, Zhou H, Wang Y, Liu Z, Cao X. Comprehensive Pan-Cancer Analysis of MTF2 Effects on Human Tumors. *Currents Problematics in Cancer*. 2023; 47(3): 100957.
- Tao T, Peng J. Liver development in zebrafish (*Danio rerio*). *Journal of Genetics and Genomics*. 2009; 36(6): 325-334.

- Thomas MJ, Seto E. Unlocking the mechanisms of transcription factor YY1: are chromatin modifying enzymes the key? *Gene*. 1999; 236(2): 197-208.
- Tien CL, Jones A, Wang H, Gerigk M, Nozell S, Chang C. Snail2/Slug cooperates with polycomb repressive complex 2 (PRC2) to regulate neural crest development. *Development*. 2015; 142: 722-731.
- Tiso N, Moro E, Argenton F. Zebrafish pancreas development. *Molecular and Cellular Endocrinology*. 2009; 312: 24-30.
- Tumes D, Hirahara K, Papadopoulos M, Shinoda K, Onodera A, Kumagai J, Yip KH, Pant H, Kokubo K, Kiuchi M, Aoki A, Obata-Ninomiya K, Tokoyoda K, Endo Y, Kimura MY, Nakayama T. Ezh2 controls development of natural killer T cells, which cause spontaneous asthma-like pathology. *Journal of allergy & Clinical Immunology*. 2019; 144(2):549-560.
- Turner BM. Reading signals on the nucleosome with a new nomenclature for modified histones. *Nature Structural & Molecular Biology*. 2005; 12:110-112.
- Ueda T, Nakata Y, Nagamachi A, Honda H. Propagation of trimethylated H3K27 regulated by polycomb protein EED is required for embryogenesis, hematopoietic maintenance, and tumor suppression. *Proceedings of the National Academy of Sciences*. 2016; 113(37): 10370-10375.
- Ujhelly O, Szabo V, Kovacs G, Vajda F, Mallok S, Prorok J, Acsai K, Hegedus Z, Krebs S, Dinnyes A, Pirty MK. Lack of Rybp in Mouse Embryonic Stem Cells Impairs Cardiac Differentiation. *Stem Cells Development*. 2015; 24: 2193–2205.
- Ura H, Usuda M, Kinoshita K, Sun C, Mori K, Akagi T, Matsuda T, Koide H, Yokota T. STAT3 and Oct-3/4 Control Histone Modification through Induction of EED in Embryonic Stem Cells\*. *Molecular Basis of Cell and Developmental Biology*. 2008; 283(15): 9713-9723.
- Urnov FD, Rebar EJ, Holmes MC, Zhang HS, Gregory PD. Genome editing with engineered zinc finger nucleases. *Nature Reviews of Genetics*. 2010; 11: 636–646.
- Vacaru AM, Di Narzo AF, Howarth DL, Tsedensodnom O, Imrie D, Cinaroglu A, Amin S, Hao K, Sadler KC. Molecularly defined unfolded protein response subclasses have distinct correlations with fatty liver disease in zebrafish. *Disease Model & Mechanisms*. 2014 ; 7(7): 823-35.
- Vacík T, Lađinović D, Raška I. KDM2A/B lysine demethylases and their alternative isoforms in development and disease. *Nucleus*. 2018; 9(1): 431-441.
- Van der Lugt NM, Domen J, Kinders K, van Roon M, Robanus-Maandag E, te Riele H, van der Valk M, Deschamps J, Sofroniew M, van Lohuizen M, Berns A. Posterior transformation, neurological abnormalities, and severe hematopoietic defects in mice with a targeted deletion of the bmi-1 proto-oncogene. *Gene & Development*. 1994; 8(7): 757-69.
- van der Velden YU, Wang L, Querol Cano L, Haramis APG. The Polycomb Group Protein Ring1b/Rnf2 Is Specifically Required for Craniofacial Development. *PLOS ONE*. 2013; 8(9): e73997.
- Van der Velden YU, Wang L, van Lohuizen M, Haramis AP. The Polycomb group protein Ring1b is essential for pectoral fin development. *Development*. 2012; 139(12): 2210-2220.
- van Heeringen SJ, Akkers RC, van Kruijsbergen I, Arif MA, Hanssen LLP, Sharifi N, Veenstra GJC. Principles of nucleation of H3K27 methylation during embryonic development. *Genome Research*. 2014; 24: 401-410.

- van Jaarsveld RH, Reilly J, Cornips MC, Hadders MA, Agolini E, Ahimaz P, Anyane-Yeboah K, Bellanger SA, van Binsbergen E, van den Boogaard MJ, Brischoux-Boucher E, Caylor RC, Ciolfi A, van Essen TAJ, Fontana P, Hopman S, Iacone M, Javier MM, Kamsteeg EJ, Kerkhof J, Kido J, Kim HG, Kleefstra T, Lonardo F, Lai A, Lev D, Levy MA, Lewis MES, Lichty A, Mannens MMAM, Matsumoto N, Maya I, McConkey H, Megarbane A, Michaud V, Miele E, Niceta M, Novelli A, Onesimo R, Pfundt R, Popp B, Prijoles E, Relator R, Redon S, Rots D, Rouault K, Saida K, Schieving J, Tartaglia M, Tenconi R, Uguen K, Verbeek N, Walsh CA, Yosovich K, Yuskaitis CJ, Zampino G, Sadikovic B, Alders M, Oegema R. Delineation of a KDM2B-related neurodevelopmental disorder and its associated DNA methylation signature. *Genetics in Medicine*. 2023; 25(1): 49-62.
- van Kruijsbergen I, Hontelez S, Veenstra GJC. Recruiting polycomb to chromatin. *The International Journal of Biochemistry & Cell Biology*. 2015; 67: 177-187.
- van Lohuizen, Verbeek S, Wientjens E, van der Gulden H, Berns A. Identification of cooperating oncogenes in E mu-myc transgenic mice by provirus tagging. *Cell*. 1991; 65(5): 737-52.
- van Wijnen AJ, Bagheri L, Badreldin AA, Larson AN, Dudakovic A, Thaler R, Paradise CR, Wu Z. Biological Functions of Chromobox (CBX) Proteins in Stem Cell Self-Renewal, Lineage-Commitment, Cancer and Development. *Bone*. 2021; 143: 115659.
- Vandamme J, Völkel P, Rosnoblet C, Le Faou P, Angrand PO. Interaction proteomics analysis of polycomb proteins defines distinct PRC1 complexes in mammalian cells. *Molecular Cell Proteomics*. 2011; 10(4):M110.002642.
- Verreault A, Kaufman PD, Kobayashi R, Stillman B. Nucleosome Assembly by a Complex of CAF-1 and Acetylated Histones H3/H4. *Cell*. 1996; 87: 95-104.
- Vidal M, Starowicz K. Polycomb complexes PRC1 and their function in hematopoiesis. *Experimental Hematology*. 2017; 48: 12-31.
- Vijayanathan M, Trejo-Arellano MG, Mozgova. Polycomb repressive Complex 2 in Eukaryotes -An Evolutionary Perspective. *Epigenomes*. 2022; 6(1): 3.
- Violot S, Hong SS, Rakotobe D, Petit C, Gay B, Moreau K, Billaud G, Priet S, Sire J, Schwartz O, Mouscadet JF, Boulanger P. The human polycomb group EED protein interacts with the integrase of human immunodeficiency virus type 1. *Journal of Virology*. 2003; 77(23): 12507-22.
- Vizàn P, Gutiérrez A, Espejo I, Garcia\_Montolio M, Lange M, Carretero A, Lafzi A, De Andrés-Aguayo, Blanco E, Thambyrajah R, Graf T, Heyn H, Bigas A, Di Croce L. The Polycomb-associated factor PHF19 controls hematopoietic stem cell state and differentiation. *Science Advance*. 2020; 6(32): eabb2745.
- Völkel P, Dupret B, Le Bourhis X, Angrand PO. Diverse involvement of EZH2 in cancer epigenetics. *American Journal of Translational Research*. 2015; 7(2): 175-193.
- Voncken JW, Roelen BAJ, Roefs M, de Vries S, Verhoeven E, Marino S, Deschamps J, van Lohuizen. Rnf2 (Ring1b) deficiency causes gastrulation arrest and cell cycle inhibition. *Proceedings of the National Academy of Sciences*. 2003; 100(5): 2468-73.
- Vuong HG, Ngo TNM, Le HT, Dunn IF. The prognostic significance of HIST1H3B/C and H3F3A K27M mutations in diffuse midline gliomas is influenced by patient age. *Journal of Neurooncology*. 2022; 158(3): 405-412.

- Wachter E, Quante T, Merusi C, Arczewska A, Stewart F, Webb S, Bird A. Synthetic CpG islands reveal DNA sequence determinants of chromatin structure. *eLife*. 2014; 3: e03397.
- Waddington CH. The epigenotype. *Endeavour*. 1942; 1:18-20.
- Wafer LN, Jensen VB, Whitney JC, Gomez TH, Flores R, Goodwin BS. Effects of Environmental Enrichment on the Fertility and Fecundity of Zebrafish (*Danio rerio*). *Journal of the American Association for Laboratory Animal Sciences*. 2016; 55: 291-4.
- Wallace KN, Akther S, Smith EM, Lorent K, Pack M. Intestinal growth and differentiation in zebrafish. *Mechanisms of Development*. 2005; 122: 157-173.
- Wamstad, JA, Corcoran CM, Keating AM, Bardwell VJ. Role of the Transcriptional Corepressor Bcor in Embryonic Stem Cell Differentiation and Early Embryonic Development. *PLOS ONE*. 2008; 3: e2814.
- Wan L, Xu K, Wei Y, Zhang J, Han T, Fry C, Zhang Z, Wang YV, Huang L Yuan M, Xia W, Chang WC, Huang WC, Liu CL, Chng YC, Liu J, Wu Y, Jin VX, Dai X, Guo J, Liu J, Jiang S, Li J, Asara JM, Brown M, Hung MC, Wei W. Phosphorylation of EZH2 by AMPK Suppresses PRC2 Methyltransferase Activity and Oncogenic Function. *Molecular Cell*. 2018; 69: 279-291.
- Wang H, Wang L, Erdjument-Bromage H, Vidal M, Tempst P, Jones RS, Zhang Y. Role of histone H2A ubiquitination in Polycomb silencing. *Nature*. 2004; 431(7010): 873-8.
- Wang L, Brown JL, Cao R, Zhang Y, Kassis JA, Jones RS. Hierarchical Recruitment of Polycomb Group Silencing Complexes. *Molecular Cell*. 2004; 14(5): 637-646.
- Wang X, Hu B, Shen H, Zhou H, Xue X, Chen Y, Chen S, Han Y, Yuan B, Zhao H, Zhi Q, Kuang Y. Clinical and prognostic relevance of EZH2 in breast cancer: A meta-analysis. *Biomedicine Pharmacotherapy*. 2015; 75: 218-25.
- Wang YY, Deng YS, Dai SK, Mi TW, Li RY, Liu PP, Liu C, He BD, He XC, Tang Y, Liu CM, Teng ZQ. Loss of microglial EED impairs synapse density, learning, and memory. *Molecular Psychiatry*. 2022 ; 27(7): 2999-3009.
- Wang Z, Gearhart MD, Lee YW, Kumar I, Ramazanov B, Zhang Y, Hernandez C, Lu AY, Neuenkirchen N, Deng J, Jin J, Kluger Y, Neubert TA, Bardwell VJ, Ivanov NB. A Non-canonical BCOR-PRC1.1 Complex Represses Differentiation Programs in Human ESCs. *Cell Stem Cell*. 2018; 22(2): 235-251.
- Ward AB, Warga RM, Prince VE. Origin of the zebrafish endocrine and exocrine pancreas. *Developmental Dynamics*. 2007; 236(6): 1558-1569.
- Washkowitz A.J, Schall C, Zhang K, Wurst W, Floss T, Mager J, Papaioannou VE. Mga is essential for the survival of pluripotent cells during peri-implantation development. *Development*. 2015; 142: 31-40.
- Weber M, de Beaufort LF. The Fishes of the Indo-Australian Archipelago. *E. J. Brill London*. Leiden. 1911.
- Weigelt B, Geyer FC, Reis-Filho JS. Histological types of breast cancer: How special are they? *Molecular Oncology*. 2010; 4(3): 192-208.
- Westerfield M. The zebrafish Book. A Guide for the Laboratory Use of Zebrafish (*Danio rerio*), 3<sup>rd</sup> edition. *University of Oregon Press*. Eugene, OR. 1994; 385 pp.

- White JK, Gerdin AK, Karp NA, Ryder E, Buljan M, Bussel JN, Salisbury J, Clare S, Ingham NJ, Podrini C, Houghton R, Estabel J, Bottomley JR, Melvin DG, Sunter D, Adams NC, Tannahill D, Logan DW, Macarthur DG, Flint J, Mahajan VB, Tsang SH, Smyth I, Watt FM, Skarnes WC, Dougan G, Adams DJ, Ramirez-Solis, Bradley A, Steel KP. Genome-wide generation and systematic phenotyping of knockout mice reveals new roles for many genes. *Cell*. 2013; 154: 452–464.
- White LJ, Thomson JS, Pounder KC, Coleman RC, Sneddon LU. The impact of social context on behaviour and the recovery from welfare challenges in zebrafish, *Danio rerio*. *Animal Behavior*. 2017; 132: 189-199.
- White R, Rose K, Zon L. Zebrafish cancer: the state of the art and the path forward. *Nature reviews cancer*. 2013; 13: 624-636.
- Wong SJ, Gearhart MD, Taylor AB, Nanyez DR, Ha DJ, Robinson AK, Artigas JA, Lee OJ, Demeler B, Hart PJ, Bardell VJ, Kim CA. KDM2B Recruitment of the Polycomb Group Complex, PRC1.1, Requires Cooperation between PCGF1 and BCORL1. *Structure*. 2016; 24: 1795-1801.
- Wong SJ, Senkovich O, Artigas JA, Gearhart MD, Ilangovan U, Graham DW, Abel KN, Yu T, Hinck AP, Bardwell VJ, Kim CA. Structure and Role of BCOR PUF1 in Noncanonical PRC1 Assembly and Disease. *Biochemistry*. 2020; 59(29): 2718-2728.
- Woo K and Fraser SE. Order and coherence in the fate map of the zebrafish nervous system. *Development*. 1995; 121: 2595-2609.
- Wu C, Jones RS, Lasko PF, Gelbart WM. Homeostasis and the interaction of zeste and white in *Drosophila*. *Molecular Genomics & Genetics*. 1989; 218: 559-564.
- Wu CT, Morris JR. Genes, genetics, and epigenetics: a correspondence. *Science*. 2001; 293(5532): 1103-1105.
- Wu G, Broniscer A, McEachron TA, Lu C, Paugh BS, Beckson J, Qu C, Ding L, Huether R, Parker M, Zhang J, Gajjar A, Dyer MA, Mullighan CG, Gilbertson RJ, Mardis ER, Wilson RK, Downing JR, Ellison DW, Zhang J, Baker SJ. Somatic histone H3 alterations in pediatric diffuse intrinsic pontine gliomas and non-brainstem glioblastomas. *Nature Genetics*. 2012; 44(3): 251-253.
- Wu H, Chen X, Xiong J, Li Y, Li H, Ding X, Liu S, Chen S, Gao S, Zhu B. Histone methyltransferase G9a contributes to H3K27 methylation in vivo. *Cell Research*. 2011; 21(2): 365-367.
- Wu H, Zeng H, Dong A, Li F, He H, Senisterra G, Seitova A, Duan S, Brown PJ, Vedadi M, Arrowsmith CH, Schapira M. Structure of the Catalytic Domain of EZH2 Reveals Conformational Plasticity in Cofactor and Substrate Binding Sites and Explains Oncogenic Mutations. *PLOS ONE*. 2013; 8(12):e83737.
- Wu W, Johansen JV, Helin K. Fbxl10/Kdm2b recruits Polycomb repressive complex 1 to CpG islands and regulates H2A ubiquitinylation. *Molecular Cell*. 2013; 49:1134-1146.
- Wu W, Zhang Z, Wang S, Fu Y, Chen J, Liang X, Le H, Pu WT, Zhang B. Depletion of polycomb repressive complex 2 core component EED impair fetal hematopoiesis. *Cell Death Disease*. 2017; 8(4): e2744.
- Wu XD, Zhang M, Gong YH, Qiang BQ, Yuan JG, Meng AM, Peng XZ. Expression pattern of Polycomb gene Nspc1 at the early developmental stage in zebrafish. *Zhongguo Yi Xue Ke Xue Yuan Xue Bao*. 2008; 30(5): 550-3.

- Wurst W, Bally-Cuif L. Neural plate patterning: upstream and downstream of the isthmus organizer. *Nature Reviews of Neuroscience*. 2001; 2: 99–108.
- Wutz AM. Ring1B is crucial for the regulation of developmental control genes and PRC1 proteins but not X inactivation in embryonic cells. *Journal of Cell Biology*. 2007; 178: 219–229.
- Xiao L, Dang Y, Hu B, Luo L, Zhao P, Wang S, Zhang K. Overlapping functions of RBBP4 and RBBP7 in regulating cell proliferation and histone H3.3 deposition during mouse preimplantation development. *Epigenetics*. 2021; 10: 1-14.
- Xiaoyun S, Yuyuan Z, Jie X, Yingjie N, Qing X, Yuezhen D, Haiguang X. PHF19 activates hedgehog signaling and promotes tumorigenesis in hepatocellular carcinoma. *Experimental Cell Research*. 2021; 406(1): 112690.
- Xu C, Bian C, Yang W, Galka M, Ouyang H, Chen C, Qiu W, Liu H, Jones AE, MacKenzie F, Pan P, Shun-Cheng Li, Wang H, Jinrong Min. Binding of different histone marks differentially regulates the activity and specificity of Polycomb repressive complex 2 (PRC2). *Proceedings of the National Academy of Sciences*. 2010 ; 107(45): 19266-19271.
- Xu K, Wu ZJ, Groner AC, He HH, Cai C, Lis RT, Wu X, Stack EC, Loda M, Liu T, Xu H, Cato L, Thornton JE, Gregory RI, Morrissey C, Vessella RL, Montironi R, Magi-Galluzzi C, Kantoff PW, Balk SP, Liu XS, Brown M. EZH2 oncogenic activity in castration-resistant prostate cancer cells is Polycomb-independent. *Science*. 2012; 338: 1465-1469.
- Yamaguchi H, Hung MC. Regulation and Role of EZH2 in Cancer. *Cancer Research Treatment*. 2014; 46(3): 209-22.
- Yan Q, Chen BJ, Hu S, Qi SL, Li LY, Yang JF, Zhou H, Yang CC, Chen LJ, Du J. Emerging role of RNF2 in cancer: From bench to bedside. *J Cell Physiol*. 2021 Aug;236(8):5453-5465. doi: 10.1002/jcp.30260. Epub 2021 Jan 5. PMID: 33400276.
- Yan Y, Zhao W, Huang Y, Tong H, Xia Y, Jiang Q Qin J. Pcgf1 Severly Compromises Proper Differentiation of Embryonic Stem Cells. *Scientific Reports*. 2017; 7:46276.
- Yang L, Lin C, Liu W, Zhang J, Ohgi KA, Grinstead JD, Dorrestein PC, Rosenfeld MG. ncRNA- and Pc2 Methylation-Dependant Gene Relocation between Nuclear Structures Mediates Gene Activation Programs. *Cell*. 2011; 147(4):773-788.
- Yang Y, Wang L, Han X, Yang WL, Zhang M, Ma HL, Sun BF, Li A, Xia J, Chen J, Heng J, WU B, Chen YS, Xu JW, Yang X, Yao H, Sun J, Lyu C, Wang HL, Huang Y, Sun YP, Zhao YL, Meng A, Ma J, Liu F, Yang YG. RNA 5-Methylcytosine Facilitates the Maternal-to-Zygotic Transition by Preventing Maternal mRNA Decay. *Molecular Cell*. 2019; 75(6): 1188-1202.
- Yap KL, Li S, Munoz-Cabello AM, Raguz S, Zeng L, Mujtaba S, Gil J, Walsh MJ, Zhou MM. Molecular Interplay of the Non-coding RNA *ANRIL* and Methylated Histone H3 Lysine 27 by Polycomb CBX7 in Transcriptional Silencing of *INK4a*. *Molecular Cell*. 2011; 38(5):662-674.
- Yeh KY, Lai CY, Hsu CC, Lo CP, Her GM. ATF4 overexpression induces early onset of hyperlipidaemia and hepatic steatosis and enhances adipogenesis in zebrafish. *Scientific Reports*. 2017; 7: 16362.
- Yi Y, Li Y, Meng Q, Li Q, Li F, Lu B, Shen J, Fazli L, Zhao D, Li C, Jiang W, Wang R, Liu Q Szczepanski A, Li Q, Qin W, Weiner AB, Lotan TL, Ji Z, Kalantry S, Wang L, Schaeffer EM, Niu H, Dong X, Zhao W,

- Chen K, Cao Q. A PRC2-independent function for EZH2 in regulating rRNA 2'-O methylation and IRES-dependent translation. *Nature Cell Biology*. 2021; 23: 341-354.
- Yiew NKH, Greenway C, Zarzour A, Ahmadiéh S, Goo B, Kim D, Benson TW, Lu X-Y, Hui DY, Kim HW, Weintraub NL. Enhancer of zeste homolog 2 (EZH2) regulates adipocyte lipid metabolism independent of adipogenic differentiation: Role of apolipoprotein E. *Journal of Biological Chemistry*. 2019; 294(21): 8577-8591.
- Yoo JY, Choi KC, Kang H, Kim YJ, Lee J, Jun WJ, Kim MJ, Lee YH, Lee O, Yoon HG. Histone deacetylase 3 is selectively involved in L3MBTL2-mediated transcriptional repression. *FEBS Letters*. 2010; 584(11): 2225-30.
- Yuan W, Xu M, Huang C, Liu N, Chen S, Zhu B. H3K36 methylation antagonizes PRC2-mediated H3K27 methylation. *Journal of Biological Chemistry*. 2011; 286: 7983-7989.
- Zao Y, Garcia A. Comprehensive Catalog of Currently Documented Histone Modifications. *Cold Spring Harbor Perspectives in Biology*. 2015; 7(9): a025064.
- Zaret KS, Grompe M. Generation and Regeneration of Cells of the Liver and Pancreas. *Science*. 2008; 322(5907): 1490-1494.
- Zdziebło D, Li X, Lin Q, Zenke M, Illich DJ, Becker M, Müller AM. Pcgf6, a Polycomb Group Protein, Regulates Mesodermal Lineage Differentiation in Murine ESCs and Functions in iPS Reprogramming. *Stem Cells*. 2014; 32(12): 3112-
- Zee BM, Levin RS, Xu B, LeRoy G, Wingreen NS, Garcia BA. *In Vivo* Residue-specific Histone Methylation Dynamics. *Genomics, Proteomics, and Bioinformatics*. 2010; 285(5): 3341-3350.
- Zencak D, Lingbeek M, Kostic C, Tekaya M, Tanger E, Hornfeld D, Jaquet M, Munier FL, Schorderet DF, van Lohuizen M, Arsenijevic Y. Bmi1 Loss Produces an Increase in Astroglial Cells and a Decrease in Neural Stem Cell Population and Proliferation. *Journal of Neuroscience*. 2005; 25(24): 5774-5783.
- Zhan Y, Zhang Y, Shengbo Zhang, Coughlan H, Baldoni pL, Jacquelot N Cao WHJ, Preston S, Louis C, Rautela J, Pellegrini M, Wicks IP, Alexander WS, Harrison LC, Lew AM, Smyth GK, Nutt SL, Chopin M. Differential requirement for the Polycomb repressor complex 2 in dendritic cell and tissue-resident myeloid cell homeostasis. *Science of Immunology*. 2021 ; 6(63): eabf7268.gg.
- Zhang Q, Wang W, Gao Q.  $\beta$ -TRCP-mediated AEBP2 ubiquitination and destruction controls cisplatin resistance in ovarian cancer. *Biochemical Biophysical Research Communication*. 2020; 523(1): 274-279.
- Zhang X, Wang W, Zhu W, Dong J, Cheng Y, Yin Z, Shen F. Mechanisms and Functions of Long Non-Coding RNAs at Multiple Regulatory Levels. *International Journal of Molecular Sciences*. 2019; 20(22): 5573.
- Zhang Y, Endam LM, Filali-Mouhim A, Zhao L, Desrosiers M, Han D, Zhang L. Polymorphisms in RYBP and AOA genes are associated with chronic rhinosinusitis in a Chinese population: a replication study. *PLOS ONE*. 2012;7(6):e39247.
- Zhang Y, Zhang JJ, Liu XH, Wang L. CBX7 suppression prevents ischemia-reperfusion injury-induced endoplasmic reticulum stress through the Nrf-2/HO-1 pathway. *Am J Physiol Renal Physiol*. 2020 Jun 1;318(6):F1531-F1538. doi: 10.1152/ajprenal.00088.2020. Epub 2020 May 11. PMID: 32390514.



- Zhao J, Oshumi TK, Kung JT, Ogawa Y, Gruau DJ, Sarma K, Song JJ, Kingston RE, Borowsky M, Lee JT. Genome-wide Identification of Polycomb-Associated RNAs by RIP-seq. *Molecular Cell*. 2010; 40: 939-953.
- Zhao J, Sun BK, Erwin JA, Song JJ, Lee JT. Polycomb Proteins Targeted by a Short Repeat RNA to the Mouse X Chromosome. *Science*. 2008; 322(5902): 750-756.
- Zhao X, Wu X. Polycomb-group proteins in the initiation and progression of cancer. *Journal of Genetics and Genomics*. 2021; 48(6); 433-443.
- Zhao X, Wu X. Polycomb-group proteins in the initiation and progression of cancer. *Journal of Genetics and Genomics*. 2021; 48(6): 433-443.
- Zhong Y, Ye Q, Chen C, Wang M, Wang H. Ezh2 promotes clock function and hematopoiesis independent of histone methyltransferase activity in zebrafish. *Nucleic Acids Research*. 2018; 46(7): 3382-99.
- Zhou T, Wang L, Zhu KY, Dong M, Xu PF, Chen Y, Chen SJ, Chen Z, Deng M, Liu TX. Dominant-negative C/ebp $\alpha$  and Polycomb group protein Bmi1 extend short-lived hematopoietic stem/progenitor cell life span and induce lethal dyserythropoiesis. *Blood*. 2011; 118(14): 3842-3852.