



**HAL**  
open science

# Cutting edge genome editing technologies in zebrafish : from efficient and precise genetic modifications to disease modeling and genetic tool engineering

Marion Rosello

► **To cite this version:**

Marion Rosello. Cutting edge genome editing technologies in zebrafish : from efficient and precise genetic modifications to disease modeling and genetic tool engineering. Development Biology. Sorbonne Université, 2021. English. NNT : 2021SORUS521 . tel-03872925

**HAL Id: tel-03872925**

**<https://theses.hal.science/tel-03872925>**

Submitted on 26 Nov 2022

**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.

---

**THESE DE DOCTORAT DE SORBONNE UNIVERSITE  
ÉCOLE DOCTORALE 158  
« ED3C CERVEAU COGNITION COMPORTEMENT »**

*Institut Curie, Unité Génétique et Biologie du Développement  
Équipe Del Bene « Développement des circuits neuronaux »*

*Institut De la Vision, Département Biologie du Développement  
Équipe Del Bene « Développement et fonctionnement du système visuel des  
vertébrés »*

CUTTING EDGE GENOME EDITING TECHNOLOGIES  
IN ZEBRAFISH:  
FROM EFFICIENT AND PRECISE GENETIC  
MODIFICATIONS TO DISEASE MODELING AND  
GENETIC TOOL ENGINEERING.

Présentée et soutenue publiquement par

**Marion ROSELLO**

le 25 novembre 2021

Directeur de thèse :

**Dr. Filippo DEL BENE**

---

Devant le Jury composé de :

**Prof. Dr. Joachim WITTBRODT**, Heidelberg University

**Dr. Annarita MICCIO**, INSERM

**Dr. Carine GIOVANNANGELI**, CNRS

**Dr. Marie BREAU**, INSERM

**Dr. Filippo DEL BENE**, INSERM

Rapporteur

Rapporteur

Examineur

Examineur

Directeur de thèse







# REMERCIEMENTS

*Merci à toutes les personnes qui ont mis leur pierre à cet édifice, scientifiquement ou non. Cette aventure a commencé bien avant ma thèse et m'a permis de rencontrer des personnes extraordinaires sans qui ce travail n'aurait jamais abouti. Ce fut un long chemin qui a commencé à se dessiner dans un amphithéâtre de Biologie Animale, s'est défini à l'observatoire de Villefranche sur mer et s'est concrétisé dans le laboratoire de Filippo Del Bene.*

Mes premiers remerciements vont à Filippo, pour m'avoir permise d'intégrer son laboratoire en stage de Master 2 et de rester pour mon Doctorat. J'avais pris le risque de postuler tard et de ne déposer qu'une seule candidature souhaitant vivement intégrer son équipe. Avec du recul c'était complètement inconscient mais ça valait définitivement le coup ! Merci Filippo de m'avoir fait confiance à travers toutes ces opportunités, projets, collaborations, encadrements et liberté dans le développement de nos projets, tout en tentant de sauver nos articles scoop après scoop. Cela m'a fait grandir humainement et scientifiquement pendant ces 4 années passées dans ton équipe.

Je remercie très sincèrement Jean-Paul Concordet. Merci pour ton temps et ton expertise sur tous les aspects de ces projets mais aussi pour ton enthousiasme qui a toujours su me rebooster, notamment lors de nos brainstorming zoom pendant le confinement !

J'adresse aussi mes remerciements sincères à Marina Mione pour m'avoir aidée et conseillée avec une bienveillance inégalable.

Je tiens à remercier ma tutrice Jamilé Hazan et mes expertes scientifiques de comité de thèse Sylvie Schneider-Maunoury et Coralie Fassier pour vos précieux conseils et votre temps tout au long de ces quatre ans.

Coralie, nos déménagements de labo m'ont finalement permis de t'avoir pas loin de moi au quotidien, merci d'avoir été présente à chaque fois que j'en avais besoin.

Un grand merci à mon Jury de thèse, Prof. Dr. Joachim Wittbrodt, Dr. Annarita Miccio, Dr. Carine Giovannangeli et Dr. Marie Breau pour avoir accepté d'examiner ces projets et de participer à mon jury de thèse. J'en suis très sincèrement honorée et chanceuse.

Je remercie l'équipe FDB/S3 ainsi que les anciens membres de l'équipe. Je remercie plus particulièrement Céline Revenu pour m'avoir accueillie en Master 2 et pour son aide extrêmement précieuse lors des demandes de financement de thèse. Merci Elena pour tes conseils toujours très sages. Je remercie aussi Evelyne Coudrier et Marie Thérèse Prospero pour leur aide et expertise sur la Myosine. Merci à mes collaborateurs dont Emilie Dambroise, Antonella Lauri, Alice Brion, Panagiotis Antoniou et Grégoire Haouy. Merci Carine Giovannangeli pour toutes ces analyses NGS super utiles et intéressantes !

Je tiens ensuite à remercier Shahad Al-Badri, Juliette Vouigny, Malo Serafini et Nour Al-Badri pour avoir été si patients avec mes affaires dont moi seule comprend l'organisation, j'ai donné mon maximum pour ne pas envahir l'étage...

Même si je ne te remercierais jamais assez, un milliard de mercis à toi ma Shahad, tu as été mon Pilier pour absolument tous et à chaque seconde de cette thèse de l'Institut Curie à l'IDV, des USA à la Chine, du 75 au 94...what's next? Scientifiquement, sportivement, psychologiquement et personnellement. Quelques dents de lait sont tombées depuis notre première conversation sur *gf1ab*... Merci aussi pour tes précieuses corrections sur ce manuscrit.

Juliette, mon rayon de soleil et la reine des J-gel d'agarose, merci pour ton efficacité et ta bonne humeur. Ton rire, tes éternuements et ta positivité ont été mes bouffées d'oxygène. Ta propagande pour le surnom Marionnette est une réussite, je dois finalement l'avouer...j'attends toujours le single star de l'été !

Merci Malo pour ton aide immense et inespérée, ton implication et ta grande sagesse ! J'ai adoré te transmettre ce que j'ai appris, ça a été très enrichissant et on forme vraiment une super équipe ! Je dois ouvrir un bar à pain au chocolat pour te remercier comme il se doit. Brad Pitt restera toujours le plus bogoss des poissons !

Nour, ma petite padawane, je suis honorée d'avoir pu t'enseigner la classification ARN/enzyme.../ADN. Wooooow lololooo, on n'aura pas eu de poissons classe 3 mais Marie Antoinette était pas mal quand même ! On s'est bien amusées, ce qu'il s'est passé dans la pièce noire... restera dans la pièce noire. Merci !

Un grand merci à Audrey Michaud. Merci ma Audrey pour les clémentines, le thé, ta franchise, ton soutien et les western-blots ! Mais aussi de m'avoir donné une image de toi au labo mangeant du poulet qui restera à jamais gravée dans ma mémoire ! Merci d'avoir amené Haroun dans notre vie.

Merci Tarek Haddouche pour avoir pris soin entre autres de mes poissons !! Seynabou Diop, Aurélien Bore, Daniel Jr, Pierre Romero et Samuel Le Cam merci d'avoir continué à me faire tellement rire entre les aller-retours Curie/IDV.

Merci à Fiona Roche, Marion Lerat et Mickaël Le, mes anges gardiens de l'IDV, aussi ex-colocataires de bureau, sans qui ce déménagement de labo serait vite devenu un enfer. Merci pour votre accueil chaleureux à l'Institut de la Vision et pour votre approvisionnement assidu et massif en nourriture, en café et en bonne humeur, mais aussi en enzymes, bactéries et kits d'extraction. Ce tabouret toujours présent pour râler, rigoler, travailler, discuter, stresser... a été un immense soutien. Merci d'avoir été mes pompom girls à plusieurs conférences !

Fiona, merci pour ta bonne humeur et ta gentillesse inégalables, ton aide précieuse et ta grande disponibilité à chaque fois que j'avais besoin d'un Onyx ! Merci aussi pour ma rencontre avec le clonage oligo !

Mickaël, tu n'auras pas réussi à me convertir à ta secte Batman mais le gâteau était un très bon début, continues. Merci pour avoir transformé ma paillasse en œuvre d'art, je n'ose plus synthétiser de l'ARN... Merci de m'avoir sauvée tant de fois, notamment pour le CD du désespoir !

Marion, merci pour ton honnêteté, tes conseils toujours extrêmement précieux et pour ton rire qui m'a fait sourire seule en pleine manip de nombreuses fois.

Merci Franck Maurinot pour ton accueil et ton humour, finalement mon grigri de Chine n'aura pas mis feu à l'IDV ! Merci Morgane Belle pour ta gentillesse et d'avoir fait attention à ne pas m'assommer à chaque ouverture de porte. Merci Samya de ne pas toujours tout comprendre.

Merci à Eric Quéinnec pour avoir été un soutien depuis ce fameux amphi de Biologie animale, tes conseils et points de vue ont toujours été essentiels. Merci à Pascal Lapébie, mon premier mentor, pour m'avoir donné une chance et m'avoir appris tant de choses. Merci aussi à Evelyn Houliston, Carine Barreau, Sandra Chevalier, Tsuyoshi Momose et Gonzalo Quiroga pour m'avoir fait découvrir la recherche académique dans la passion et la bienveillance.

Je remercie Douglas W. Houston pour ma première rencontre avec CRISPR/Cas9 !

Mes remerciements vont aussi à Eric Quéinnec, Jean-Philippe Chambon, Karen Pottin, Gabriel Krasovec et Thomas Condamine pour leur soutien depuis mon Master 1.

Gabriel, les Cnidaires restent dans mon cœur même si j'ai viré dans le côté obscur des vertébrés. On sera peut-être amenés à retravailler ensemble qui sait... j'adorerais.

Merci à mes amis Mel, Zak, Tiff, Edma, Marion et Malina pour leurs encouragements et leurs compréhensions sur la surcharge de travail dont moi seule suis-je responsable, je suis chanceuse de vous avoir. Malina, malgré la distance on a finalement passé chaque étape de nos thèses en simultané, ces longs messages vocaux m'ont toujours fait un bien fou !

Merci à ma famille. Jean-Pierre, Mélanie, Pierre et Pascal pour ces week-ends ressourçant, Nahla et Hoda pour ces délicieux repas et gâteaux. Un immense MERCI à mes parents pour avoir toujours cru en moi, ce que je fais vous dépasse complètement mais sachez que je vous dois absolument tout. Merci Maman pour ton aide précieuse sur la mise en page de ce manuscrit.

Enfin, François, merci à l'infinie et au-delà ! Merci pour ta patience et ton soutien au quotidien. La liste serait longue, mais merci entre autres d'avoir fait « le taxi direction labo » à pas d'heure, le script python qui m'a fait gagner un temps énorme sur mes designs de sgRNA et merci de m'avoir écoutée t'expliquer mes manip pendant des heures. Merci d'être tout simplement toi !



# SUMMARY

Zebrafish is an ideal model organism to study developmental biology and has become a powerful system for disease modeling. Since the emergence of the CRISPR/Cas9 technology, several disease models have been generated to introduce loss-of-function mutations, putting this vertebrate model at the forefront of genetic studies. However, the generation of zebrafish mutant lines harboring precise pathological mutations is still challenging as knock-in approaches are still poorly efficient.

In this study we investigated the capacities of several new cytosine base editor variants engineered in introducing precise C-to-T or G-to-A mutations into the genome. Using these tools, we broadened the gene editing possibilities in zebrafish by obtaining mutagenesis up to 100% of efficiency in injected embryos and we developed new approaches to perform multiplex mutagenesis and co-selection. We further explored biological applications of this technology such as endogenous activation of signaling pathway or cancer and rare syndrome modeling. In addition, we were able to validate in zebrafish the pathogenicity of mutations found in patients. Finally, we aimed at developing and optimizing the versatile prime editing technology in this animal model, a recent technology that although highly promising remains to date poorly efficient.

To conclude, this study contributed in providing a large panel of genetic tools working in zebrafish without generating DNA double strand breaks and we assessed their limitations and complementarity for zebrafish gene editing. Through the accessibility and ease of genome manipulations in this animal model, this work further highlights the use of zebrafish as a good platform to engineer new genome editing approaches in animals and to study genetic disorders.

# LIST OF FIGURES AND TABLES

**Figure 1:** Examples of zebrafish disease models generated using CRISPR/Cas9 technology.

**Figure 2:** Schematic representation of CRISPR/Cas9.

**Figure 3:** Schematic representation of tissue-specific knock-out approaches.

**Figure 4:** Schematic representations of the Cas9 nickases.

**Figure 5:** Schematic representation of the base excision repair pathways.

**Figure 6:** Schematic representation of the cytosine base editor.

**Figure 7:** Schematic representation of the Prime Editor technology.

**Figure 8:** 3 mpf dwarf *MZcbl*<sup>-/-</sup> zebrafish display skeletal malformations.

**Figure 9:** BE4-gam shows high fidelity of base editing in zebrafish.

**Figure 10:** High mortality rate observed by expressing the human PIP5K1C (E146K), (Y205C) and (Y221C) variants.

**Figure 11:** Embryos expressing PIP5K1C (Y205C), (Y221C) and (E146K) present a spectrum of ocular abnormalities and microcephaly at 2 and 3 dpf.

**Figure 12:** Ocular abnormality and microcephaly phenotypes are found upon the expression of PIP5K1C (Y205C), (Y221C) and (E146K).

**Figure 13:** Larvae expressing PIP5K1C(205C) and PIP5K1C(Y221C) show head cartilage defects.

**Figure 14:** Amino acid alignment of the human PIP5K1C with the zebrafish *Pip5k1ca* and *Pip5k1cb* proteins.

**Figure 15:** *pip5k1ca* and *pip5k1cb* are expressed in the developing central nervous system in zebrafish.

**Figure 16:** Introduction of the human E146K mutation in *Pip5k1ca* and *Pip5k1cb* after C-to-T conversion using base editors.

**Figure 17:** Mutations in *pip5k1ca* using base editor induce a similar phenotype to the one obtained by the human *PIP5K1C(E146K)* mRNA overexpression.

**Figure 18:** Prime editing technology in zebrafish for small insertion and deletion.

**Figure 19:** Prime editing *versus* base editing for C-to-T conversion in zebrafish.

**Figure 20:** Pigmentation loss in zebrafish embryos by G-to-T conversions using prime editing technology.

**Figure 21:** pegRNA design strategies for in-frame insertions of the *V5-tag* coding sequence in *gfi1ab* gene.

**Figure 22:** Maternal expression of prime editor proteins.

**Figure 23:** Fish overexpressing the BE4-gam using the UAS/Gal4 system display toxicity effect.

**Figure 24:** CBE4max-SpRY(R33A) generates C-to-T conversions at a lower efficiency than CBE4max-SpRY in zebrafish.

**Figure 25:** Establishing the tissue specific base editing approach in zebrafish using the Gal4/UAS system.

**Figure 26:** Schematic representation of the two MS2-PE2 approaches.

**Table 1:** Cas9 variants recognizing other PAMs than the classical 5'-NGG-3' PAM.

**Table 2:** Design of the pegRNAs used in the study and respective gene editing frequencies obtained.

# ABBREVIATIONS

**ABE:** Adenine Base Editor  
**ALS:** amyotrophic lateral sclerosis  
**AP site:** apurinic or apyrimidinic site  
**APOBEC1:** apolipoprotein B mRNA editing enzyme catalytic subunit 1  
**BE:** base editor  
**BER:** base-excision repair  
**CBE:** Cytosine Base Editor  
**crRNA:** CRISPR RNA  
**CRISPR/Cas9:** Clustered Regularly Interspaced Short Palindromic Repeats/CRISPR-associated  
**CRISPRa:** CRISPR activation  
**CRISPRi:** CRISPR interference  
**dCas9:** dead Cas9  
**DSB:** double-strand break  
**Dpf:** day-post-fertilization  
**ENU:** N-ethyl-N-nitrosourea  
**FEN1:** flap structure-specific endonuclease 1  
**FGFR3:** fibroblast growth factor receptor 3  
**gDNA:** genomic DNA  
**GOF:** gain-of-function  
**HDR:** homology-directed repair  
**HR:** homologous recombination  
**Kb:** kilobase  
**KI:** knock-in  
**KO:** knock-out  
**LOF:** loss-of-function  
**Mpf:** month-post-fertilization  
**nCas9:** Cas9 nickase  
**NGS:** next-generation sequencing  
**NHEJ:** non-homologous end joining  
**nt:** nucleotides

**NUC:** nuclease lobe  
**PAM:** protospacer adjacent motif  
**PBS:** primer binding site  
**PE:** prime editing  
**pegRNA:** prime editing guide RNA  
**PIM:** PAM interacting motif  
**REC:** recognition lobe  
**RT:** reverse transcriptase  
***SaCas9:*** *Staphylococcus aureus* Cas9  
**SAM:** Synergistic Activation Mediator  
***SauriCas9:*** *Staphylococcus auricularis* Cas9  
**sgRNA:** single-guide RNA  
**SNP:** single nucleotide polymorphism  
***SpCas9:*** *Streptococcus pyogenes* Cas9  
**ssDNA:** single strand DNA  
**TALEN:** Transcription Activator-Like Effector Nucleases  
**tracrRNA:** trans-activating RNA  
**UAS:** upstream activation sequence  
**UGI:** inhibitors of uracil DNA glycosylase  
**Wt:** wild-type  
**ZFN:** Zinc Finger Nucleases

# TABLE OF CONTENTS

SUMMARY .....	9
LIST OF FIGURES AND TABLES .....	10
ABBREVIATIONS .....	12
CHAPTER I: INTRODUCTION .....	17
<b>I.1. Disease modeling in zebrafish</b> .....	17
I.1.1. Zebrafish, a powerful vertebrate model .....	17
I.1.1.1. The advantages of zebrafish .....	17
I.1.1.2. Forward genetic approaches .....	18
I.1.2. Ease of genetic and gene expression manipulations to model diseases .....	19
I.1.2.1. Overexpression of mutated gene assays .....	19
I.1.2.2. Reverse genetic approaches .....	20
<b>I.2. CRISPR/Cas9 in zebrafish</b> .....	22
I.2.1. Efficient generation of knock-out alleles using CRISPR/Cas9 technology .....	22
I.2.1.1. The technology .....	22
I.2.1.2. Conditional knock-out using CRISPR/Cas9 .....	24
I.2.1.3. The use of Cas9 with mutated catalytic domains .....	27
I.2.2. CRISPR/Cas9-based knock-in strategies .....	29
<b>I.3. Next generation genome editing technologies based on     CRISPR/Cas9</b> .....	31
I.3.1. Base editing .....	31
I.3.1.1. Cytosine base editor generates C:G to T:A mutations .....	31
I.3.1.2. Optimization of cytosine base editors and their applications .....	33
I.3.1.3. Additional base editors .....	34
I.3.1.4. Limitations of the BE technology .....	35
I.3.2. Prime editing .....	37

I.3.2.1. General principles of the prime editing technology .....	37
I.3.2.2. State of the art on the use of prime editing.....	39
I.3.2.3. Optimization of the technique .....	40
I.3.2.4. Prime editing <i>versus</i> Base editing.....	41
I.3.3. Cas9 variants recognizing other PAMs .....	42
<b>I.4. Aims of the study .....</b>	<b>45</b>
<b>CHAPTER II: RESULTS .....</b>	<b>47</b>
<b>II.1. Establishment and use of several cytosine base editors in zebrafish.....</b>	<b>47</b>
II.1.1. Summary article 1 .....	47
II.1.2. Article 1: Precise base editing for the <i>in vivo</i> study of developmental signaling and human pathologies in zebrafish.....	48
II.1.3. Skeletal phenotypic analysis of the dwarf <i>MZcbl<sup>-/-</sup></i> mutants .....	64
II.1.4. In-depth analysis of the BE4-gam specificity in zebrafish.....	65
<b>II.2. Establishment of the near PAM-less base editor strategy for multiplex gene editing and base editing co-selection approach.....</b>	<b>67</b>
II.2.1. Summary article 2 .....	67
II.2.2. Article 2: Disease modeling by efficient genome editing using a near-PAMless base editor <i>in vivo</i> .....	68
<b>II.3. The use of base editors in zebrafish to study mutations in <i>PIP5K1C</i> gene found in patients with a poly-malformative syndrome .....</b>	<b>96</b>
II.3.1. Context .....	96
II.3.2. Causality link between mutations found in <i>PIP5K1C</i> of patients and the disease .....	97
II.3.3. Endogenous mutation of <i>PIP5K1C</i> zebrafish orthologues using base editors: disease modeling and limitations .....	102

<b>II.4. The versatile prime editing technology in zebrafish</b> .....	108
II.4.1. Establishment of the technology .....	108
II.4.2. Maternal contribution of the prime editor .....	117
<b>CHAPTER III: DISCUSSION</b> .....	119
<b>III.1. Disease modeling in zebrafish using base editors</b> .....	119
<b>III.2. Base editors in zebrafish: a powerful approach with limitations.</b> .....	122
III.2.1. The power and limitations of base editors .....	122
III.2.2. Addressing RNA off-targets concern for conditional base editing.....	125
<b>III.3. Prime editing in zebrafish</b> .....	129
III.3.1. A need of optimization to routinely use prime editing in zebrafish .....	129
III.3.2. MS2-PE2 technology .....	130
<b>CHAPTER IV: CONCLUSION</b> .....	133
<b>MATERIALS AND METHODS</b> .....	135
<b>REFERENCES</b> .....	142
<b>ANNEXES</b> .....	160

# CHAPTER I: INTRODUCTION

## I.1. DISEASE MODELING IN ZEBRAFISH.

Zebrafish (*Danio rerio*) is nowadays a model of choice to study human diseases *in vivo*<sup>1, 2, 3, 4</sup>. This organism has numerous advantages due to its external and rapid embryonic development, high reproduction rate and optical accessibility during the first few days of its development. Together with the rapid emergence and increasing efficiency of genetic tools, these advantages make this vertebrate a relevant model to study the causal links between genetic mutations and developmental and/or physiological defects but also to elucidate the mechanisms responsible for the human diseases associated with the same genetic alterations.

### I.1.1. ZEBRAFISH, A POWERFUL VERTEBRATE MODEL.

#### I.1.1.1. THE ADVANTAGES OF ZEBRAFISH.

For several decades, rodents have been the main experimental animal model used in biomedical research. Over time other laboratory models have emerged, including zebrafish. Although zebrafish is a non-mammalian animal that does not have certain organs like mammary glands and lungs, this vertebrate model presents many advantages over rodents for the study of embryonic development and diseases. Indeed, a single zebrafish couple can generate hundreds of embryos per week. Their rapid and external development, small size and optical clarity allow to perform live-imaging *in vivo*<sup>5, 6, 7</sup>. As the embryonic development is external, the eggs are easily collected and micro-injected in a non-invasive manner. Motor neuron axonogenesis starts then during the first day of development<sup>8, 9</sup>, the heart beats already the second day-post-fertilization (dpf)<sup>10</sup>, and by day 5 the retina is entirely formed and connected to the brain<sup>11</sup>. The larvae start to eat at 5/6 dpf<sup>12</sup> and adults become sexually mature at 3-

month-post-fertilization (mpf)<sup>13, 14</sup>. This fast-embryonic development and generation time account for some of the so many advantages conferred by this animal model.

The genome of *Danio rerio* is completely sequenced and at least one ortholog for 70% of human genes is present in the genome<sup>15</sup>. In particular, it has been reported that 82% of the genes causing human diseases have one orthologue in zebrafish so far<sup>15</sup>. This high degree of conservation provides support for the use of *Danio rerio* as a laboratory model to investigate human diseases.

Zebrafish embryos are also well suited to pharmacological treatments by the addition of the drug directly delivered in the medium, allowing treatments of intact and living embryos or larvae. These high-throughput approaches have been used for drug discovery by screening several small molecules and they will have an important future as the generation of precise human pathogenic mutations will be routinely performed in animal models, allowing drug screening for personalized medicine<sup>16, 17, 18, 19, 20</sup>.

#### I.1.1.3. FORWARD GENETIC APPROACHES.

In addition to the previously listed advantages, zebrafish has emerged as an important model with the conduction of a large-scale N-ethyl-N-nitrosourea (ENU) mutagenesis screens which introduced random mutations into the genome, thus providing an extensive collection of phenotypes and hundreds of mutant fish lines<sup>13, 21, 22, 23</sup>. The role of a large number of factors during early embryonic development and organogenesis were studied with this approach. Some cancers were modeled<sup>24</sup> such as epidermal papilloma<sup>25</sup> or renal cancer<sup>24, 26</sup>. However, although this approach generated a high number of zebrafish mutant lines and allowed successfully functional genome studies, the identification of mutated genes and genetic background clearing to isolate the mutation of interest can be complex and time-consuming. Moreover, forward genetic approaches remain poorly useable for genetic disease modeling when the aim is to target a specific gene supposedly implicated in a disease.

Since then, researches developed genetic tools and opened the possibility to manipulate the genome of zebrafish such as transgenesis in order to overexpress proteins of interest. These advances have placed zebrafish as a relevant vertebrate model for genetic disease modeling<sup>1</sup>.

## I.1.2. EASE OF GENETIC AND GENE EXPRESSION MANIPULATIONS TO MODEL DISEASES.

### I.1.2.2. OVEREXPRESSION OF MUTATED GENE ASSAYS.

The injection of an *in vitro* synthesized and capped mRNA encoding a protein of interest, such as the mRNA of a pathological mutated gene, into one-cell stage embryos is a standard approach employed by the zebrafish community<sup>27, 28</sup>. Upon injection, the mRNA diffuses in almost every cells of the embryos during the early stages of embryogenesis. Although being a powerful strategy, the limited stability of the mRNA after 2 dpf or the strong lethality phenotype that may occur due to this overexpression approach that is not cell type specific are limitations that does not allow its systematic use and may false the interpretation of the results.

Next, the generation of transgenic lines stably expressing mutated genes and/or fluorescent marker genes became also central to biomedical studies in zebrafish. This method allows to model diseases and/or to perform live imaging of cells and to track cell dynamics *in vivo*. Transgenesis in zebrafish was improved and is nowadays widely used in combination with transposons like Tol2 identified in the genome of the medaka fish, allowing the insertion into the genome of large DNA sequences kilobases (kb)<sup>29, 30, 31, 32, 33</sup>.

Several transgenic line with tissue-specific promoter driving the human mutated gene fused to fluorescent marker were then generated. This approach allowed to study cancer formation and monitor its progression in zebrafish<sup>34</sup> and also to model neurodegenerative diseases such as tauopathies, polyglutamine diseases or motor neuron degenerations<sup>35</sup>. Moreover, in order to circumvent lethality issues and to be able to express a mutated gene of interest such as oncogene in one or several tissues of choice, researchers developed the conditional Gal4/UAS system<sup>36</sup>. The transcriptional activator Gal4 controls gene expression by binding the upstream activation sequences (UAS). The Gal4 is placed under the control of a tissue-specific promoter and activates the *UAS* sequences driving the expression of the gene of interest. Thus, the *UAS* is activated exclusively in the cells expressing the Gal4. Several transgenic lines expressing a Gal4 in different and specific tissues are now available to achieve tissue-specific gene expression.

Using this approach or the CRE-Lox recombination system, a large number of transgenic lines overexpressing mutated genes such as cancer genes in a spatiotemporal manner have been generated, using promoters driving the gene expression in specific cells with direct expression control<sup>24, 37</sup>.

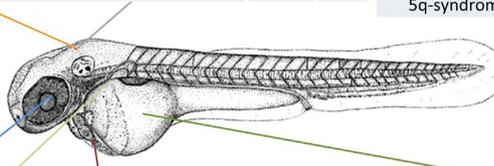
All these versatile strategies developed over the years and their feasibility rendered zebrafish an ideal model system to study many cancer types and other diseases. However under these conditions, endogenous gene regulations, often critical in these diseases, are not maintained, representing a major limitation by employing such approaches.

### I.1.2.2. REVERSE GENETIC APPROACHES.

Reverse genetic approaches in zebrafish have notably begun with the development of the Zinc Finger Nucleases (ZFNs)<sup>38, 39, 40, 41, 42, 43</sup>. Loss-of-function (LOF) mutations using ZFNs have been introduced in several embryonic developmental genes, including *gata2*, a gene associated with familial coronary artery disease<sup>42</sup>. Soon after another strategy, easier to design than ZFNs, has increased the use of reverse genetic approaches and thus disease modeling in zebrafish, the Transcription Activator-Like Effector Nucleases (TALENs)<sup>44, 45</sup>. For instance, models for acute myelogenous leukemia (AML)<sup>46</sup>, collagen VI-related diseases<sup>47</sup> or Rett syndrome<sup>48</sup> were generated using TALEN. Rapidly, the extremely powerful Clustered Regularly Interspaced Short Palindromic Repeats/CRISPR-associated (CRISPR/Cas9) technology, which will be the focus of the next section (*c.f.* section I.2), revolutionized genome editing and the generation of knock-out (KO) zebrafish mutant lines<sup>49</sup>.

Many disorders has been modeled and studied in zebrafish by the generation of LOF mutations using CRISPR/Cas9, generating stable lines or by directly working with Fo injected embryos (crispants)<sup>50</sup>. These include, among others, ciliopathies<sup>51</sup>, nervous system<sup>52, 53, 54, 55, 56, 57, 58</sup>, retinal<sup>59, 60, 61</sup>, kidney<sup>62, 63, 64, 65</sup>, cardiovascular<sup>66, 67, 68, 69, 70, 71, 72</sup>, liver<sup>73</sup>, hematological and immunological<sup>74, 75, 76</sup> diseases, craniofacial and skeletal abnormalities<sup>77, 78, 79, 80, 81, 82, 83</sup>, heterotaxy<sup>84, 85</sup> and other syndromes<sup>58, 86, 87</sup> (Fig. 1). Some tumor suppressor genes were also targeted<sup>24, 37</sup>.

Nervous system disease				Craniofacial and skeletal abnormalities				L-R assymetry and heterotaxy			
Hirschsprung disease	<i>dennd3, ncl1, nup98, tbata</i>	crispant	(52)	Craniofacial abnormalities	<i>psmd12, brf1b, smchd1, bptf</i>	crispant	(77-80)	Heterotaxy	<i>mmp21</i>	crispant and line	(84)
Occipital encephalocele	<i>wdr63</i>	crispant	(53)	Craniofacial abnormalities	<i>itga8</i>	line	(81)	Heterotaxy	<i>klf8</i>	line	(85)
Hypermanganesemia	<i>slc39a14</i>	line	(54)	Osteoperosis	<i>atp6vh1</i>	line	(82)	<b>Heamotological and immunological disease</b>			
Epileptic encephalopathy	<i>stxbp1</i>	line	(55)	Osteoperosis	<i>lrp5</i>	crispant and line	(83)	Congenital neutropenia	<i>csf3, cdk9</i>	crispant	(74, 75)
Pyridoxine dependent epilepsy	<i>aldh7a1</i>	line	(56)					5q-syndrome	<i>rsp14</i>	crispant	(76)
Autism spectrum disorder	<i>shank3b</i>	line	(57)								
<b>Retinal disease</b>				<b>Cardiovascular disease</b>				<b>Liver disease</b>			
Nanophthalmia	<i>mrfp</i>	crispant	(59)	Decreased # cardiomyocytes	<i>dab2</i>	crispant	(66)	Familial intrahepatic cholestasis	<i>abcb11</i>	line	(73)
Leber Congenital Ameurosis	<i>cct2</i>	line	(60)	Hypertrophic cardiomyopathy	<i>gstk1</i>	crispant	(67)	<b>Ciliopathies</b>			
Microphthalmia	<i>il7r</i>	line	(61)	Pulmonary arterial hypertension	<i>ccdc30</i>	crispant	(68)	Joubert syndrome	<i>armc9</i>	crispant	(51)
<b>Kidney Disease</b>				Dilated cardiomyopathy	<i>titin</i>	crispant	(69)	<b>Other syndroms</b>			
Renal agenesis	<i>greb1</i>	crispant	(62, 63)	Decreased # cardiomyocytes	<i>pr130</i>	line	(70)	Marphan	<i>fbn1</i>	line	(86)
Tubulo-interstitial, glomerulocystic	<i>sec61a1</i>	crispant	(64)	Hypoplastic left heart syndrome	<i>sap130a</i>	line	(71)	Severe X-linked intellectual disability	<i>EIF2S3</i>	line	(58)
Glomerular abnormalities	<i>apol1</i>	crispant	(65)	Atrial septal defects	<i>lhx2</i>	line	(72)	Connective tissue disorder	<i>lpo8</i>	line	(87)



**Figure 1: Examples of zebrafish disease models generated using CRISPR/Cas9 technology.** For each line of the table, the disorder modeled in zebrafish is in the first column and the targeted genes in the second column. Crispant: F0 injected embryos, line: stable mutant zebrafish line. Adapted from <sup>50</sup>.

CRISPR/Cas9 has placed zebrafish at the forefront of human disease modeling by generating KO alleles efficiently. However, although successful studies reported its use for the generation of knock-in (KI), efficient and base specific mutagenesis in this animal model is still challenging and further technological advances are needed. Indeed, all diseases cannot be modeled through the generation of LOF mutations and being able to precisely and efficiently insert any type of mutations found in human patients would have a big impact and remains to this date one of the biggest challenges in this field.

## I.2. CRISPR/CAS9 IN ZEBRAFISH.

The CRISPR/Cas9 technology has had a major impact in life sciences, enabling directed genome mutagenesis with unprecedented efficiency in plants, animals, microorganisms and cultured cells. The breakthrough importance of this technological advancement is highlighted by the fact that two of the main players in its initial discovery and technical adaptation in eukaryotic cells<sup>88</sup>, Dr. Emmanuelle Charpentier and Dr. Jennifer A. Doudna, were awarded by the Nobel prize in chemistry in 2020. In zebrafish, this technology is now extremely targeted and efficient to generate knock-out alleles and is currently being optimized for knock-in approaches.

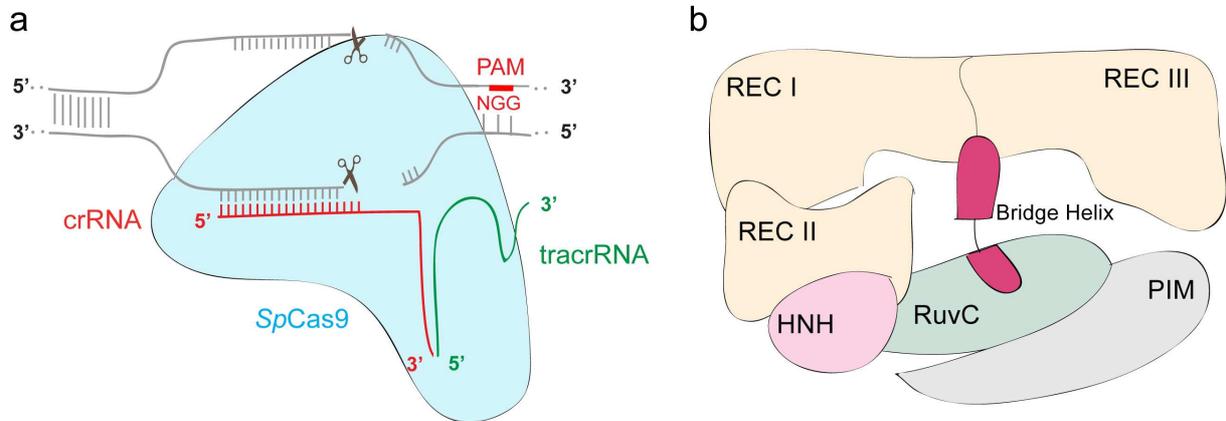
### I.2.1. EFFICIENT GENERATION OF KNOCK-OUT ALLELES USING CRISPR/CAS9.

#### I.2.1.1. THE TECHNOLOGY.

The CRISPR/Cas9 system was described in bacteria where it acts as an adaptive immune system protecting against invading viruses<sup>89</sup>. This system is composed by the endonuclease Cas9 and a single-guide RNA (sgRNA). The sgRNA is an RNA duplex structure: the CRISPR RNA (crRNA) at the 5' end has 20 nucleotides (nt) complementary to the targeted genomic DNA (gDNA) locus and the trans-activating RNA (tracrRNA) at the 3' end, forming loops on which the Cas9 protein binds. The sgRNA thus guides the Cas9 enzyme to the target locus (Fig. 2.a)<sup>88, 90</sup>.

The most widely used Cas9 for genome editing is the *Streptococcus pyogenes* Cas9 (*SpCas9*) which site-specifically recognizes and generates double-strand break (DSB) on any DNA sequences upstream of the 5'-NGG-3' protospacer adjacent motif (PAM). The PAM is thus a key recognition element which allowed to target specific loci using CRISPR/Cas9 technology<sup>88</sup>. The Cas9 presents seven domains: REC I, REC II, REC III, HNH, RuvC, PAM interacting motif (PIM) and the bridge helix (Fig. 2.b)<sup>91, 92, 93, 94</sup>. The recognition (REC) lobe is constituted by the 3 REC domains, sensing nucleic acids and regulating the conformational change of the Cas9<sup>93</sup>. The nuclease lobe (NUC) is composed by the PIM and the HNH and RuvC domains which catalyze the cleavage

of the targeted DNA strand and the non-targeted DNA strand respectively. The bridge helix connects the REC and the NUC lobes (Fig. 2.b).



**Figure 2: Schematic representation of CRISPR/Cas9.** (a) The *SpCas9* in blue binds the sgRNA composed by the crRNA in red which binds to the locus of interest and the tracrRNA in green. The *SpCas9* generates a double strand break if the 5'-NGG-3' PAM is present on the non-targeted strand downstream of the sgRNA. (b) Structure of the *SpCas9* endonuclease (adapted from Cavanagh & Garrity, "CRISPR Mechanism", CRISPR/Cas9, Tufts University, 2014. <https://sites.tufts.edu/crispr/> (Access in 2021)). The 3 REC domains sense nucleic acids and regulate the conformational change of the Cas9, HNH and RuvC domains catalyze the DNA cleavage, PIM domain interacts and recognizes the PAM and the bridge helix connects the REC lobe and the other domains<sup>91, 92, 93, 94</sup>.

The *SpCas9* forms a complex with the sgRNA that directs the endonuclease to the targeted locus. Upon stable annealing of the sgRNA immediately upstream and on the other strand of the 5'NGG-3' PAM, the *SpCas9* generates DSB (Fig. 2.a). The DSB is repaired by the non-homologous end joining (NHEJ) repair pathway. As consequence of the NHEJ repair which works without a DNA template, randomly sized insertions or deletions (INDELS) are produced, sometimes leading to a shift in the open reading frame and the generation of premature stop codons. Thus, the generated by-products cannot be precisely predicted and this approach is mainly used to introduce LOF mutations. To circumvent this limitation, another pathway, the homology-directed repair (HDR), can be stimulated by providing a repair DNA template. This will be the focus of the next section (*c.f.* section I.2.2).

Consequently, the design of this technology to generate LOF relies on the used sgRNA. The *SpCas9* protein is now widely commercially available from several companies and the production of the short oligonucleotide is simple and fast. Moreover, sgRNAs are now made available from commercial companies such as IDT

and Synthego as synthetic RNAs with chemical modifications increasing the stability of the sgRNA. The ease of the design and the availability of the web tools have had a major impact in the zebrafish field, allowing to generate zebrafish KO mutant lines in a specific and rapid manner.

However, in some cases lethality, compensation mechanisms, and pleiotropic phenotypes could be induced by constitutive gene LOF mutations<sup>95, 96</sup>. For these reasons, the ability to control precisely the generation of mutations in a spatiotemporal-regulated manner has been and still is the focus of several studies.

#### 1.2.1.2. CONDITIONAL KNOCK-OUT USING CRISPR/CAS9.

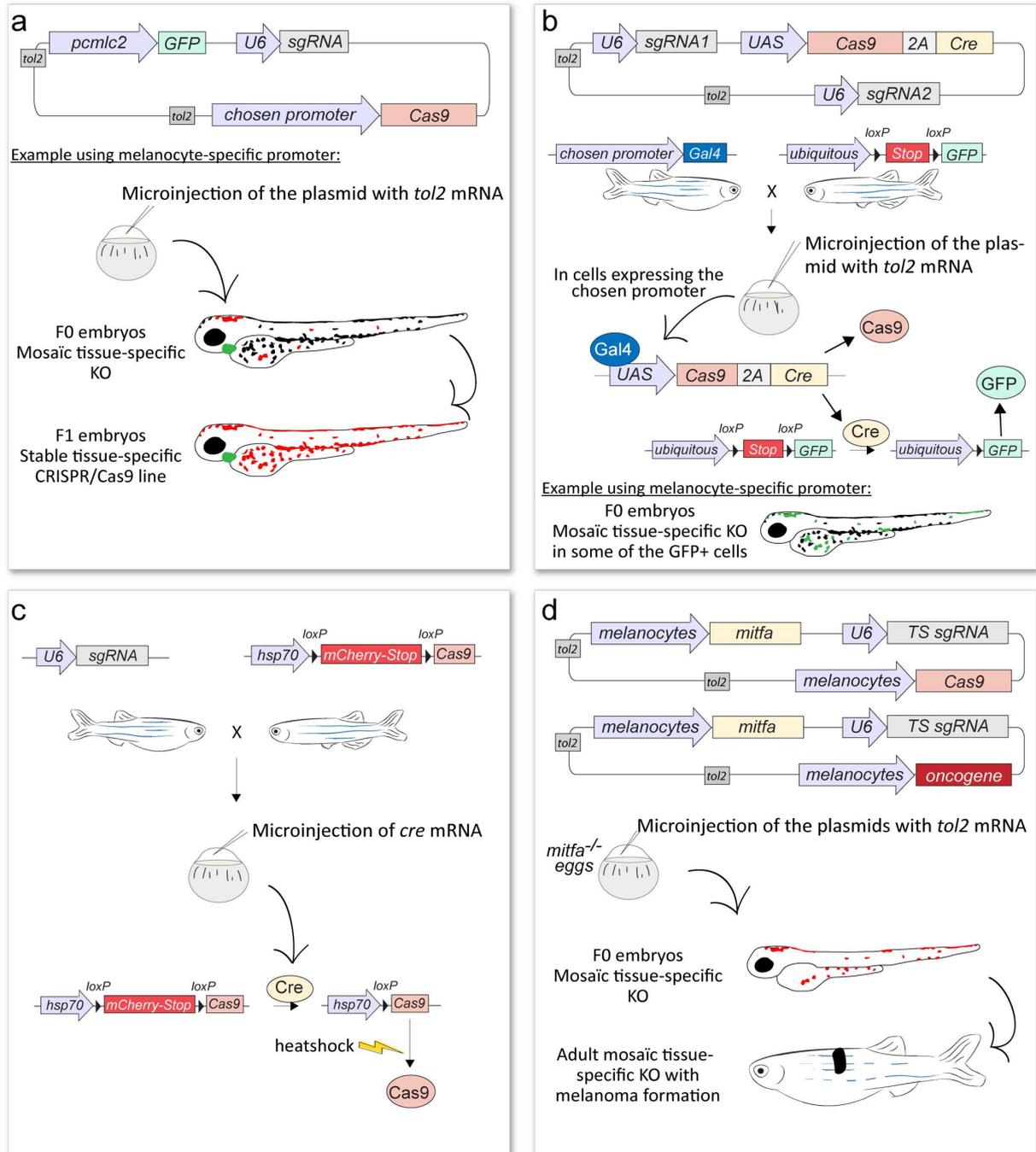
The use of CRISPR/Cas9 in zebrafish makes it possible to efficiently generate constitutive KO alleles and to explore gene functions *in vivo*. However, as previously mentioned, through a stable gene disruption, limitations such as compensatory mechanisms or embryonic lethality can arise.

For these reasons, conditional knock-out approaches were developed in zebrafish. In order to spatially regulate the generation of KO, the expression of the Cas9 was driven by a tissue-specific promoter and the expression of the sgRNA by the constitutive *U6* promoter recognized by the RNA polymerase III<sup>97</sup>. In these conditions, the sgRNA is expressed in all the cells that have integrated the plasmid and gene editing can occur only in the cell expressing the Cas9<sup>98</sup> (Fig. 3.a). The success of this approach relies on the transgenesis and gene editing efficiencies and the LOF is thus generated in an unlabeled mosaic manner, making the phenotype characterization challenging. In this context, the 2C-Cas9 method was developed, combining the Gal4/UAS system with the CRISPR/Cas9 technologies (Fig. 3.b), enabling the generation of targeted mutagenesis in the tissue of choice. Furthermore, the Cre-lox system was added to this approach which permitted to label the cell populations that have at some point expressed the Cas9 and that are therefore potentially knocked out, allowing the study of cell-autonomous defects and cell lineages<sup>99</sup>.

In parallel, another strategy was developed to temporally regulate the KO alleles generation. Indeed, to this end, the *Tg(hsp70:loxP-mCherry-STOP-loxP-cas9)* and *Tg(U6:sgRNA)* transgenic lines were made. Here, the expression of the Cre recombinase is ubiquitously expressed by mRNA injection in *Tg(hsp70:loxP-mCherry-STOP-loxP-cas9;U6:sgRNA)* embryos where the *mCherry-STOP* sequence

is floxed upon Cre activity. The expression of the Cas9 in these cells is then itself controlled by heat-shock induction. This approach allows a spatiotemporal control of the KO<sup>100</sup> (Fig. 3.c). Also, the injection of a plasmid containing the *cre* cassette controlled by a tissue-specific promoter could direct the expression of the Cas9 in the tissue of choice in addition to the temporally control.

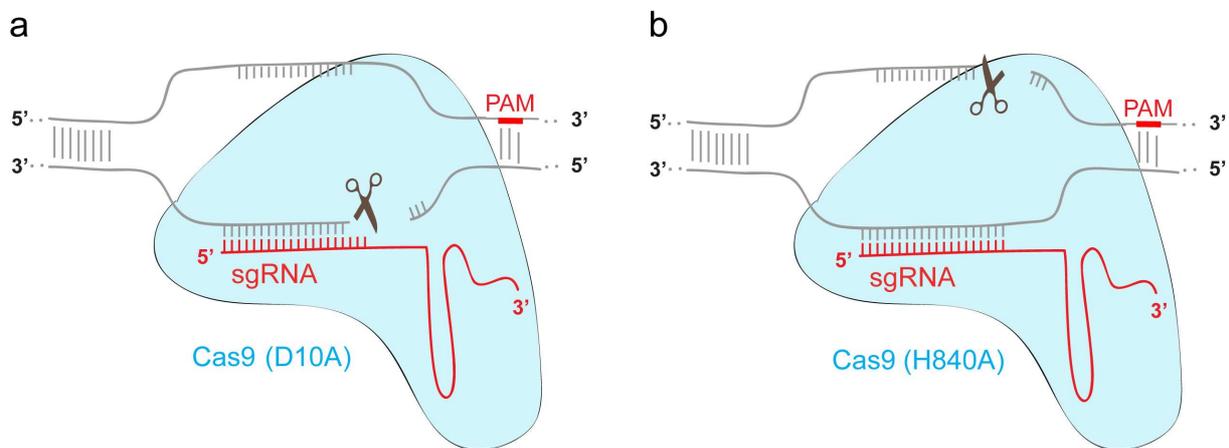
More recently, the MAZERATI, Modeling Approach in Zebrafish for Rapid Tumor Initiation, platform was engineered by combining overexpression of oncogene and tumor suppressor genes KO in the tissue of choice. This method allows to rapidly model potential cancer drivers *in vivo*<sup>101</sup> (Fig. 3.d).



**Figure 3: Schematic representation of tissue-specific knock-out approaches. (a)** Cas9 cell-specific expression is dependent on the time of the chosen promoter activity (red cells). The U6 promoter allows the ubiquitous expression of the sgRNA (the green heart marker serves as a transgenesis reporter for embryos expressing the plasmid). Putative Cas9+/sgRNA+ edited cells cannot here be visualized. Adapted from<sup>98</sup>. **(b)** GFP+ cells have once expressed the Cas9 and sgRNA and may be edited. Adapted from<sup>99</sup>. **(c)** Temporal control of the Cas9 expression. No visualization of the cas9-expressing cells. Adapted from<sup>100</sup>. **(d)** Cas9, sgRNA and/or oncogene expression specifically in melanocytes dependent on the time of the chosen promoter activity (red cells). These cells overexpressing an oncogene and/or knocked out for a tumor suppressor gene are visualized upon the pigmentation rescue through *mitfa* expression. Adapted from<sup>101</sup>.

## I.2.1.3. THE USE OF CAS9 WITH MUTATED CATALYTIC DOMAINS.

The Cas9 protein carries two defined catalytic domains which have their specific target strand, the RuvC and HNH domains. By generating precise mutations on each domain, researchers were able to create Cas9 nickases (nCas9) that only induce a single-strand DNA break named a nick. By introducing the D10A mutation in the Cas9, the RuvC domain was inactivated, leading to the generation of a nick on the targeted DNA strand 3 bases upstream of the PAM<sup>88</sup> (Fig. 4.a). Another mutation, the H840A, led instead to the inactivation of the HNH domain and generated a nick 3 bases upstream of the PAM on the non-targeted DNA strand<sup>92</sup> (Fig. 4.b). These mutations gave rise to two nCas9 variants: Cas9(D10A) and Cas9(H840A) (Fig. 4).

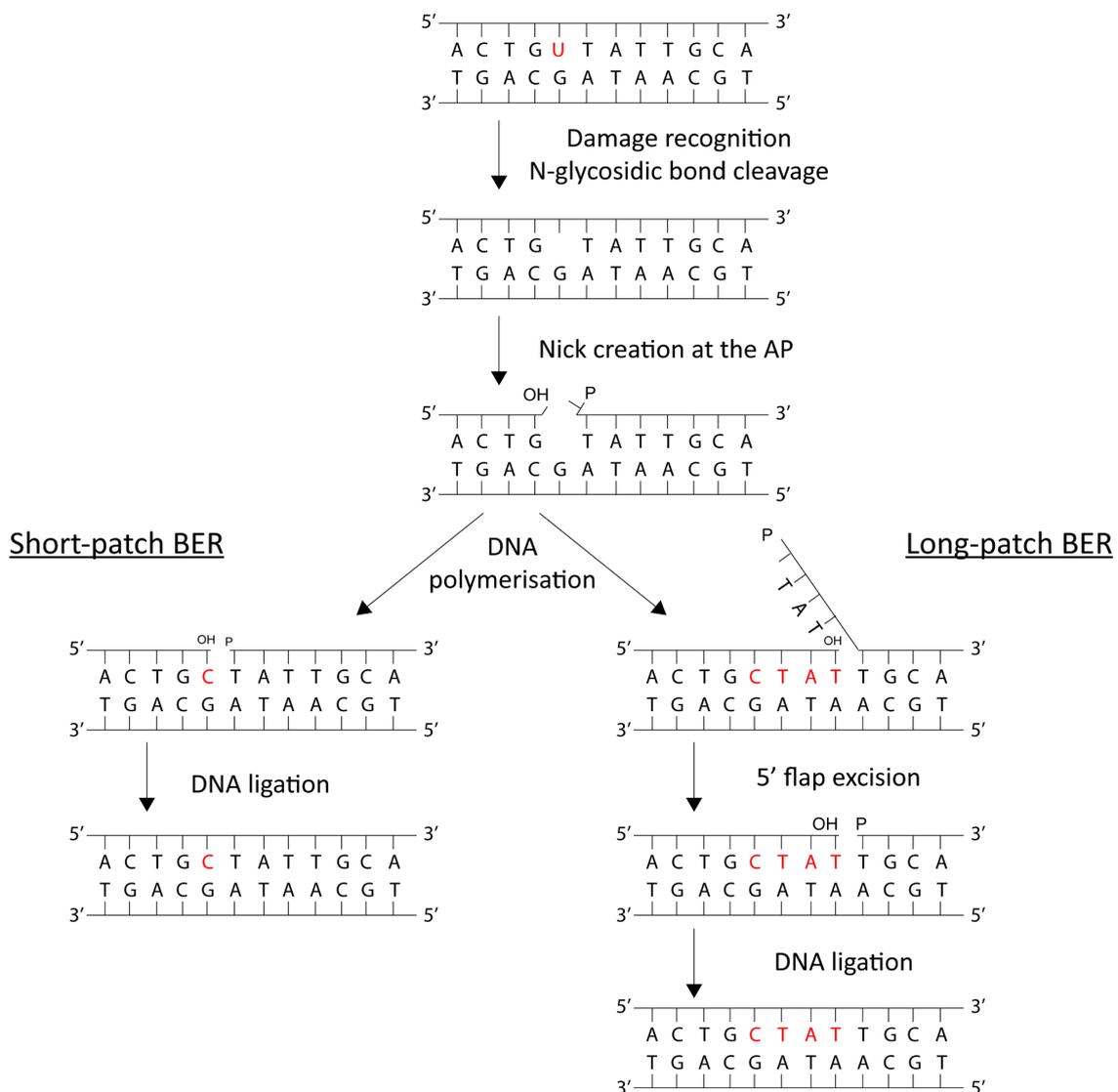


**Figure 4: Schematic representations of the Cas9 nickases. (a)** The Cas9(D10A) in blue has the RuvC domain inactivated and generated a nick 3 bp upstream of the PAM on the targeted DNA strand<sup>92</sup>. **(b)** The Cas9(H840A) in blue has the HNH domain inactivated and generated a nick 3 bp upstream of the PAM on the non-targeted DNA strand<sup>92</sup>.

It was previously reported that the Cas9 activity is not affected when few mismatches are present in the crRNA sequence, leading to potential off-targets<sup>102, 103, 104, 105, 106</sup>. To overcome this issue, many online tools have been developed and are able to generate prediction scores for gene editing efficiency and off-target risks in order to choose the suitable sgRNA<sup>107, 108</sup>. Another solution that was reported to be efficient against this problem is the use of nCas9 with two sgRNAs targeting each DNA strand of the same locus in order to perform precise deletion<sup>109</sup>. Indeed, using this approach, the gene editing efficiency on the on-target was similar to the classical CRISPR/Cas9 assays and no edits were observed in predicted DNA off-target sites<sup>110, 111</sup>.

Single-strand nicks are generally repaired via the highly conserved base-excision repair (BER) pathway<sup>109, 112</sup>. BER follows two principal pathways: the short-patch BER

and the long-patch BER<sup>113</sup>. The short-patch BER is implicated in the correction of single bases by first removing the damaged base via the cleavage of the N-glycosidic bond, creating an apurinic or apyrimidinic site (AP site). The AP endonuclease or AP lyase then induces a nick in the DNA at the AP site. The induced gap is then corrected by the DNA polymerase and the helix integrity is restored by the DNA ligase<sup>113</sup> (Fig. 5, short-patch BER). When several bases are damaged, the long-patch BER pathway is favored. During the long-patch repair process, the DNA polymerase synthesizes more than one base, leading to the formation of the 3' and 5' flaps. The flap synthesized by the DNA polymerase (3' flap) is then favorably kept by the flap structure-specific endonuclease 1 (FEN1) which preferably removes the 5' flap<sup>113</sup> (Fig. 5, long-patch BER).



**Figure 5: Schematic representation of the base excision repair pathways.** In red is the mutation that will be repaired using the base excision repair pathway. For single nucleotide mutation, the short-patch BER is usually favored instead for several mutations the long-patch BER is usually used. Adapted from<sup>113</sup>.

Integrating both the D10A and H840A mutations into the Cas9, researchers were able to create the dead Cas9 (dCas9) without any endonuclease activity and they used it as a cargo to bring proteins at the locus of interest into the genome. For instance, through this ability, CRISPR activation (CRISPRa) was developed by fusing the dCas9 to transcriptional activator factors and sgRNAs were designed to target promoter region. Using this approach, gene expression was increased specifically. CRISPR interference (CRISPRi) was also engineered based on the same principle but fusing to the dCas9 transcriptional repressor factors<sup>114</sup>. Furthermore, the Synergistic Activation Mediator (SAM) system was developed by combining the MS2 system with the CRISPRa. The bacteriophage coat protein MS2 binds with high affinity and specificity the MS2 RNA hairpins by its RNA-binding site. These MS2 RNA stem loop structures are found only in the bacteriophage RNA, conferring the high specificity of this system. In SAM, the dCas9 was fused to the transcription activator factor such as VP64 on the one hand and on the other hand the MS2 protein to additional transcription activation domains. The MS2 RNA loops were added to the 3' end of the sgRNA which targeted the promoter regions. The expression of these components allowed to boost the expression of the target gene. For instance, the authors reported that PrP expression was increased up to 10-fold in mRNA and protein using SAM<sup>115</sup>. Among the CRISPRa systems, SAM showed the highest levels of gene expression activation<sup>116</sup>.

### 1.2.2. CRISPR/CAS9-BASED KNOCK-IN STRATEGIES.

The generation of precise DSBs using programmable nucleases, such as Cas9, has further enabled to develop knock-in approaches by stimulating the homology-directed repair (HDR) with the addition of an exogenous DNA template.

In order to insert a precise genetic modification in frame with the targeted coding sequence into the zebrafish genome, different strategies varying the design of the DNA templates have been tested. Researchers have used DNA templates with two arm regions homologous to the DSB site of the genome of different lengths, either short or long arms, to respectively stimulate the microhomology-mediated end joining (MMEJ) or the homologous recombination (HR) pathways<sup>117, 118</sup>. For instance, HA tags<sup>119</sup>, point mutations<sup>120, 121, 122, 123, 124, 125</sup>, *cre/loxP* sequences<sup>126, 127</sup> or fluorescent protein coding sequences<sup>128, 129</sup> were introduced into the zebrafish genome using these strategies. The use of TALEN to generate the DSB and knock-ins was also successful<sup>130, 131, 132</sup>.

In order to improve the HDR-based KI approaches in zebrafish several optimizations were performed, such as the use of morpholino to enhance MMEJ<sup>133</sup>, *in vivo* or *in vitro* linearization of the donor plasmid to stimulate HDR<sup>117, 130, 131</sup> or variation of the homology arm length<sup>129</sup>. Recently, phosphorothioate modifications on the oligonucleotide also improved knock-in efficiency in cells and in fish<sup>124, 134, 135</sup>.

Although successful, these approaches remain challenging in zebrafish and poorly efficient. In addition, their efficiency appears to be highly locus and DNA template dependent as the frequencies of correct gene editing are still low. Furthermore, the designs of the donor template are not straightforward and can be time-consuming. Insertions performed using these strategies are usually restricted to dividing cells and they are thus difficult to apply for somatic gene editing. Moreover, although important optimizations were successfully done, some concerns such as DNA off-targets<sup>136</sup> and the high rate of unwanted mutation generation are still limitations of these techniques. Recently, exciting advances in the field have been reported allowing the generation of precise point mutations, the base editing technology, which avoids the use of DSBs and DNA templates and will be the focus of the next section.

## I.3. NEXT GENERATION GENOME EDITING TECHNOLOGIES BASED ON CRISPR/CAS9.

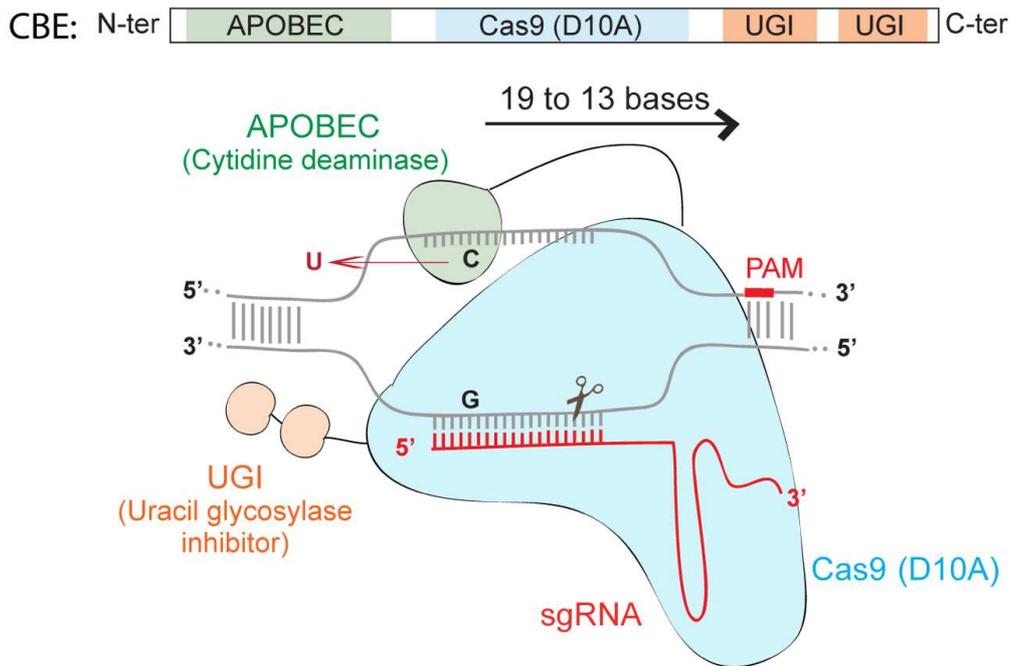
The discovery of CRISPR/Cas9 has revolutionized the field of life sciences as well as gene and cell therapy by enabling the manipulation of DNA with unprecedented efficiency in a wide range of organisms and cell types<sup>137, 138</sup>. Recently and based on CRISPR/Cas9, the laboratory of Dr. David R. Liu engineered cutting edge precise genome editing tools by fusing enzymes to the Cas9-nickase in which one of the two catalytic domains is inactivated. These approaches allow the generation of specific and precise genetic modifications without double-strand DNA breaks or donor DNA template and without requiring cell division, meaning that base editing can also be induced in the genome of post mitotic cells. Since 2016, two principal classes of base editors (BEs) have been developed and optimized to perform precise base pair transition mutations<sup>139</sup>. In 2019, the CRISPR/Cas9 toolkit was further expanded by developing a more versatile approach, the prime editing (PE)<sup>140</sup>.

### I.3.1. BASE EDITING.

#### I.3.1.1. CYTOSINE BASE EDITOR GENERATES C:G TO T:A MUTATIONS.

The “Cytosine Base Editor” (CBE) was developed in cultured cells to precisely generate C:G to T:A conversions using a *SpCas9*(D10A) nickase fused to the APOBEC1 (apolipoprotein B mRNA editing enzyme catalytic subunit 1) cytidine deaminase<sup>139</sup>. The APOBEC1 catalyzes the hydrolytic deamination of a **C** base leading to its conversion into a **U** base which will be later read as a T base. Operating on single strand DNA (ssDNA), the sgRNA used in this system binds the G strand of **C:G**. The *SpCas9*(D10A) nickase has one of the two catalytic domains inactivated and cuts only the G strand to direct DNA repair on the non-edited strand in addition to bringing the APOBEC1 to its target site to act locally (Fig. 6). The 5'-NGG-3' PAM position is thus important to obtain an efficient base conversion. C-to-T conversion can occur in an editing window located at a distance of 13 bp to 19 bp away from the PAM, meaning that all C bases present within this frame can potentially be edited. Thus, in this system the *SpCas9*(D10A) nickase is used firstly to bring the APOBEC1 to its target site to

deaminate the cytidine to an uracil resulting in U:G and secondly to direct DNA repair through the break of the non-edited strand. To prevent the change of the U back to a C mutation by BER (Fig. 5), two inhibitors of Uracil DNA Glycosylase (UGI) were fused to the C terminal end of the *SpCas9*(D10A) (Fig. 6).



**Figure 6: Schematic representation of the cytosine base editor.** The Cas9(D10A) in blue brings the APOBEC in green to the site and cuts only the DNA strand complementary to the sgRNA and to direct the repair on the non-edited strand. The APOBEC converts all C bases to U in the editing window found [19 to 13 bp] away from the PAM. The UGIs in orange prevent the change of the U back to a C by base excision repair. Adapted from<sup>141</sup>.

This ability to introduce precise and efficient C:G to T:A conversions is of high interest with considerable applications such as the introduction of precise premature stop codons in any gene of interest, as it does not rely on the error-prone NHEJ repair mechanism for which the generation of stop codons following the repair cannot be predicted. Also, it is especially important for medical research to model pathologies by introducing point mutations found in human patients or to rescue a pathogenic mutation. Indeed, so far the predominant class of human pathogenic mutations is estimated to be single nucleotide polymorphism (SNP) representing 58% of the variants sequenced from patients and among them 47% are C:G to T:A transitions<sup>142, 143, 144</sup>.

One of the first developed CBE, the BE3, has been used in several animal models such as mice<sup>145, 146, 147, 148, 149</sup>, zebrafish<sup>150</sup>, rabbits<sup>151</sup>, pig<sup>152</sup> and xenopus<sup>153, 154</sup>, allowing the creation of a large panel of disease models or rescue applications but INDELs mutations or unwanted conversions were frequently observed as well in these studies at the target loci. This has indeed been the case in zebrafish, even if the use of the BE3 has been demonstrated to be a promising tool in this animal model for which knock-ins are still challenging. BE3 in this model showed rather low base editing efficiency (less than 29%) and more importantly generated many INDELs or unwanted SNPs (more than 5%)<sup>150</sup>. Therefore, although novel and promising, the BE3 variant has not been a genome editing tool widely used by the zebrafish community.

**For these reasons, during my PhD we explored the capacities of several new CBE variants engineered from improved versions of BE3 in zebrafish. I will develop this work in the Chapter II. 1 section of this manuscript.**

### I.3.1.2. OPTIMIZATION OF CYTOSINE BASE EDITORS AND THEIR APPLICATIONS.

Since the development of the first base editors has been reported in 2016, intensive work has been done in cultured cells in order to increase the base conversion efficiency and decrease the production of undesired by-products. Different linkers added between the enzymes were tested to increase flexibility, improving the gene editing efficiency by 50% and leading to the generation of the BE4 editor<sup>155</sup>. In addition, the bacteriophage Mu Gam protein was fused to the new version of the base editor to reduce INDELs frequencies, creating the BE4-gam variant. Here, the Gam protein binds and protects the DSBs free ends that could result from a nick generated on the edited strand (the U strand of U:G) by the AP lyase (DNA-apurinic or apyrimidinic site lyase) after the cleavage of the N-glycosidic bond during BER (Fig. 5)<sup>155, 156</sup>. In parallel, a bis-bpNLS BE4 variant was generated to improve nuclear localization signals and was codon optimized using GenScript. This new variant, containing all these optimized features, was called BE4max<sup>157</sup>. Following that, work was performed to do an ancestral sequence reconstruction on APOBEC1 where several APOBEC1 sequences were tested leading to the generation of the AncBE4max variant, which showed an increased efficiency for several mutations in cultures cells in comparison to the use of BE4 to

introduce the same mutations. In addition, the AncBE4max was also slightly more efficient than the BE4max for some of the targeted loci<sup>157</sup>.

Using base editing, numerous disease models have been generated. For instance, in mice, mutations found in Alzheimer's disease could be modeled<sup>158</sup> and a model of human ablepharon macrostomia syndrome has been reported in zebrafish<sup>159</sup>. In monkey, the BE4max has been used to create a model of Hutchinson–Gilford progeria syndrome<sup>160</sup>. Additional studies have shown rescues of pathogenic mutations in mice models through injections of adenoviral vectors containing BE and sgRNAs. Indeed using this approach, a reduction of plasma cholesterol was measured upon base editing in liver cells of a hypercholesterolemic mouse model<sup>161</sup>; the rescue *Fah* loss of function and its lethal phenotype causing hereditary tyrosinemia type 1 could be performed<sup>162</sup>; astrocytes were edited resulting in delayed progression of the amyotrophic lateral sclerosis (ALS) disease<sup>163</sup> and it has been possible to rescue features of a hearing deficit<sup>164</sup>. However, only few studies have reported the generation of multiplex base editing (namely in pig<sup>165</sup> and in monkey<sup>166</sup>) which is necessary to model multigenic disorders in animal models.

**To address this challenge, we decided to explore the capacities of CBE in zebrafish to perform multiplex base editing as well as a co-selection strategy. I will develop this work in the Chapter II. 2 section of this manuscript.**

### I.3.1.3. ADDITIONAL BASE EDITORS.

In parallel of the development of CBEs, another base editor has been engineered in order to perform A:T to G:C transitions, the Adenine Base Editor (ABE)<sup>139, 157</sup>. It is based on the same principle of CBEs, but instead of fusing the *SpCas9*(D10A) to the APOBEC1, an adenine deaminase which converts an adenine into an inosine was used in this system. Using this tool, researchers were able to perform several disease rescues in mice. For instance, the altered visual function of adult mice was repaired by correcting a pathogenic mutation by subretinal injections of lentivirus expressing ABE and sgRNA as shown in a recent study<sup>167</sup>. Another example is the restored dystrophin expression in Duchenne muscular dystrophy mouse model by intramuscular injections of dual-AAV9 to deliver ABE and sgRNA<sup>168</sup>. In sheep, researchers were able to generate the Booroola fecundity mutation<sup>169</sup>. In zebrafish, only one study in 2018 has reported

the use of a codon optimized zABE7.10 that can induce base editing<sup>170</sup>. A germline transmission of the mutation for several genes was furthermore obtained<sup>170</sup>. Nevertheless, they also got a very high INDELS frequency, often at the same frequency or higher than the wanted A-to-G conversions. Although this tool expands the conversion abilities of base editing, ABE appears so far to be less reliable than CBE in zebrafish.

Recently, one study has further reported the fusion of CBE and ABE to create a new base editor, the ACBE<sup>171</sup>. Using this genetic tool, they could introduce simultaneously the C-to-T and A-to-G transitions in immortalized cells up to 19.65% of efficiency and in mouse and porcine fetal fibroblasts up to 1.69% of efficiency<sup>171</sup>. Interestingly, different laboratories then decided to extend the range of genetic manipulation by developing approaches in order to perform precise transversions. Indeed, in 2021, three studies have described base editing systems allowing C-to-G mutations in cultured cells<sup>172, 173, 174</sup>. The final products obtained in these studies were less pure than the use of ABE or CBE. Nevertheless, while they have not been established yet in animal models, these reports further enlarge the range of base conversion possibilities using base editors and they are at a starting point of future optimizations.

#### I.3.1.4. LIMITATIONS OF THE BE TECHNOLOGY.

Although highly efficient, there are still some intrinsic limitations to the base editor techniques such as the restricted editing window in which base modifications can occur and the limited conversion possibilities. As a result, these tools cannot be applied to any mutation of interest, nor to generate deletions or insertions.

Indeed, the editing window is an important aspect of the system as the presence of the PAM at the right distance of the targeted base is necessary and therefore a first limitation to take into consideration. Another significant parameter is the presence of several cytosines in the editing window that can interfere with the final desired products. To circumvent these issues, many Cas9 recognizing other PAMs than the classical 5'-NGG-3' one have been developed and integrated to the base editing technology. It is now possible to take advantage of them to target bases that were not targetable before using the 5'-NGG-3' PAM as well as exclude some of the Cs present

in the editing window by the design of other sgRNAs on the same locus. This part will be developed in the Chapter I.3.c of this manuscript.

In addition, although the predominant class of human pathogenic mutations is estimated to be SNP, which account for 58% of the variants sequenced in patients, other human genetic variants associated with disorders are deletions, duplications, or insertions, as in the case of triplet-repeat diseases<sup>142, 143, 144</sup>. Therefore, in order to be able to model or to repair these pathogenic mutations, an efficient method to perform precise deletion and insertion is still needed.

More recently, the novel prime editing approach was engineered in cell culture and has a remarkable potential to drastically broaden gene editing possibilities which will be introduced in this next chapter<sup>140</sup>. Although it appears to be much less efficient than base editors, prime editing will and already has allowed important advancements in the field of genome editing engineering, disease modeling and gene therapies.

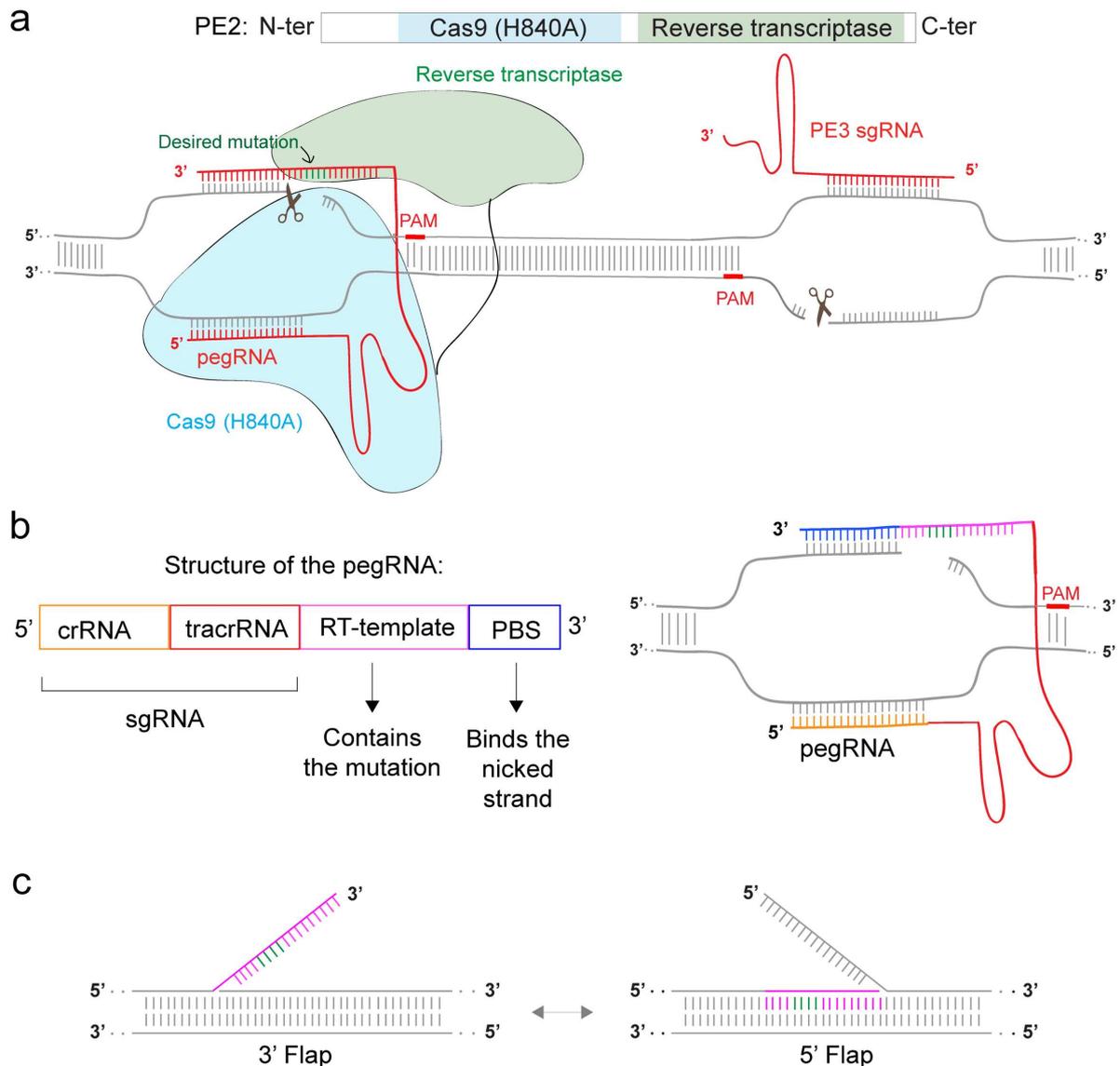
## I.3.2. PRIME EDITING.

### I.3.2.1. GENERAL PRINCIPLES OF THE PRIME EDITING TECHNOLOGY.

As for the BE system, the Prime Editing technology (also called PE2) relies on a modified CRISPR/Cas9 system. It is based on using a fusion of the *SpCas9*(H840A) nickase to an optimized moloney murine leukemia virus (M-MLV) reverse transcriptase (RT) which copies the desired mutation directly from a prime editing guide RNA (pegRNA)<sup>140</sup> (Fig. 7.a).

In this system, the pegRNA contains 3 major parts: the first one, at the 5' end, is a classical sgRNA a sequence targeting 20 bases of the locus upstream of the PAM, the second named "RT-template" carries the desired mutation and acts as the template for the reverse transcriptase and the last one is the primer binding site (PBS) which binds the other gDNA strand of the locus (Fig. 7.a, b). The binding of the PBS sequence upstream of the nick creates a DNA:RNA hybrid with a 3' hydroxyl group where the reverse transcriptase transcribes and inserts the information encoded within the RT-template of the pegRNA (Fig. 7.b). The pegRNA guides the Cas9(H840A), and thus the fused RT, to the correct target locus while simultaneously harboring the sequence that will be used to modify the genome. As for the BE, the Cas9 of PE2 has one catalytic domain inactivated, conferring it a nickase activity. Here, compared to the BE, it is the other domain of the Cas9 which is inactivated, which leads to the generation of a nick 3 bases upstream of the PAM on the DNA strand that is not targeted by the sgRNA contained in the pegRNA (Fig. 7.a).

After the DNA synthesis by the RT and from the RT-template has occurred, two redundant single-stranded DNA flaps are generated: a 3' flap containing the mutation copied from the pegRNA and a 5' flap composed by the unedited DNA sequence (Fig. 7.c). Among the two flaps, the 5' overhanging flap is preferably removed by specific flap endonucleases such as FEN1 during the long-patch BER, the DNA repair process initiated to repair at least two nucleotides (Fig. 5). Thus, the unedited sequence is eliminated and the mutation is incorporated to the genome after the ligation step (Fig. 7.c).



**Figure 7: Schematic representation of the Prime Editor technology. (a)** The Cas9(D10A) in blue brings the Reverse Transcriptase (RT) in green to the site and cuts only one strand. The RT retrotranscribes the desired mutation from the pegRNA. **(b)** Schematic representation of the different parts of the pegRNA. The crRNA in orange and tracrRNA in red (the sgRNA) bind the locus of interest and the Cas9, the RT-template in magenta containing the desired mutation in green is the template of the RT. The PBS sequence in blue binds the nicked DNA strand and serves as a primer for the RT as it forms a DNA:RNA complex. **(c)** The two redundant single-stranded DNA flaps generated after the DNA synthesis by the RT in magenta, a 3' flap containing the mutation copied from the pegRNA in green and a 5' flap composed by the unedited DNA sequence. The 5' overhang flap is preferably removed by specific flap endonucleases such as FEN1 during the long-patch BER (Fig. 5). Adapted from<sup>140</sup>.

The prime editor thus generates a heteroduplex of wild-type and mutant DNA sequences. As for the strategy developed in BE to force DNA repair on the non-edited DNA strand, PE2 efficiency is significantly improved by adding a sgRNA that drive the same nCas9-RT fusion protein  $\pm 14$ -116 bp away from the nick induced using the

pegRNA to induce another nick on the non-edited DNA strand in 3' of the edited site (Fig. 7.a). This strategy thus favors repair of the heteroduplex into the homoduplex mutant sequence, a system called prime editor 3 (PE3) as it is composed by a tripartite system: the pegRNA, the nCas9-RT protein and a second sgRNA (Fig. 7.a). Using PE3, an increase of efficiency by 1.5-4.2 fold was obtained in cultured cells where all 12 possible base conversions as well as insertions up to 44 bp and deletions up to 80 bp with an efficiency average of  $\pm 41\%$  were generated<sup>140</sup>. Nevertheless, in this study, the PE3 approach showed a slight increase of INDEL frequencies as both strands are nicked, potentially generating a double strand break. To partially circumvent this problem, Anzalone *et al.* have shown a reduced INDEL rates for some loci using the so called PE3b strategy. In this approach, the second sgRNA binds the edited DNA sequence in order to create the nick on the non-edited strand at the edited locus. The second nick will thus be generated only after editing has occurred as the second sgRNA only recognizes the edited sequence<sup>140</sup>. The design of the pegRNAs is less straightforward than the classical sgRNAs. Anzalone *et al.* suggest different parameter to consider when designing the pegRNAs. They indeed recommend to start with a PBS of  $\pm 13$  nt and  $\pm 40-60\%$  GC, RT-template of  $\pm 10-16$  nt and the second nick at  $\pm 40-90$  bp away from the first nick for the PE3 system<sup>140</sup>.

**As part of my PhD work, I undertook the challenge to set up this promising strategy in zebrafish as soon as the first report of prime editing was published in cultured cells. The results that we so far obtained and their future optimizations will be detailed in the Chapter II.4 and III.3 of this manuscript.**

### I.3.2.2. STATE OF THE ART ON THE USE OF PRIME EDITING.

PE2 was rapidly applied in plants, animal organisms and cultured cells. Several proofs of concept and applications have been reported in plants<sup>175, 176, 177, 178, 179, 180, 181, 182</sup> such as tomato, rice, potato, also in wheat<sup>175</sup>, *Drosophila*<sup>183</sup>, mouse<sup>184</sup> and in iPS Cells and organoids<sup>185, 186</sup>. In primary patient-derived ductal liver and intestinal organoids, for two analyzed targets (*HEK3* and *CTNNB1*), it has been reported that the desired edit can be obtained with 30%-50% of efficiency while only 1-4% of unwanted mutation were generated using the PE3 system. It should be noted that in this system, the direct comparison of PE3 with HDR for small deletion generation results in higher INDEL

formation (80% vs 8%) and lower wanted mutation frequency (6% vs 21%) using the HDR approach<sup>185</sup>. Similar results were obtained in the first report of prime editing<sup>140</sup>.

The first prime editing use reported in animals applied micro-injections into mouse embryos of the PE3 system components and PE2 as mRNA. Prior to the injection, many pegRNAs were screened in cultured mouse cells to select the most efficient one by modulating the RT-template and PBS lengths. The highest efficiency observed for different tested lengths were locus-dependent. Furthermore, in this study, the authors observed that the PE3b system was overall less efficient than the PE3 strategy. The selected pegRNA was then injected into mouse embryos and 44%-75% of the blastocysts were edited, with mutation frequencies ranging from 1.1% to 18.5%, with low unwanted INDEL frequency. However, a high proportion of other mutations was also obtained. As an example, one injected embryo for a G-to-C conversion presented in 95.81 % of the mutated sequences a G-to-T modification and only 3.55% the G-to-C conversion. As a conclusion, even though promising, the PE3 system presented in this study a low conversion fidelity in murine embryos and needs to be improved<sup>184</sup>.

In *Drosophila* prior to *in vivo* gene editing, researchers first tested PE in *Drosophila*-derived cultured S2R+ cells to validate the technique and select pegRNAs working in these cells<sup>183</sup>. Three different genes were tested for which up to 8% of correct edits with a very low frequency of INDELs (less than 1%) using PE2 system were obtained in the transfected cells. Performing this approach *in vivo* using a stable line in order to express the PE2 and the selected pegRNA from their previous analysis, they obtained precise gene editing with an efficiency reaching up to 35.2%. Furthermore, they were able to demonstrate that *in vivo*, the PE3 strategy was much more efficient than the PE2 system but a high INDEL frequency was also obtained using the PE3. For instance, for the *forked* target, they obtained at least twice as much of wanted edits using the PE3 strategy leading to 11.6% of efficiency but, 67.9% of INDELs was also obtained in these gene edited animals. In contrast, using PE2 this INDELs rate was less than 1%<sup>183</sup>.

### I.3.2.3. OPTIMIZATION OF THE TECHNIQUE.

As it has been the case for all the genetic modification tools so far, these first studies on prime editing emphasize its feasibility. These findings will inspire future

optimizations and applications in animal models and gene therapy<sup>138</sup>. Indeed, this technology presents a high potential and versatility for precise mutagenesis. In most cases, it is however still poorly efficient with a high rate of unwanted mutation generation.

So far, one of the first optimization attempts consists in the development of many software and web tools made to improve the design of the pegRNA, a task that is less straightforward than for sgRNAs. Numerous pegRNA design tools are nowadays available, among which for instance PE-designer<sup>187</sup>, PE-analyzer<sup>187</sup>, PrimeDesign<sup>188</sup>, PINE-CONE<sup>189</sup>, pegIT<sup>190</sup>, pegFinder<sup>191</sup>, multicrispr<sup>192</sup>, PnB Designer<sup>193</sup>, Easy-Prime<sup>194</sup>. Parameters such as the length and base composition of PBS and the RT template sequences are still not well-defined as they change depending on the locus and they can massively impact on the efficacy. The tools therefore aim to help designing pegRNA taking these parameters into account.

#### I.3.2.4. PRIME EDITING *VERSUS* BASE EDITING.

These reports on the use of prime editing in several different models highlights that this technique is less efficient than BE and not yet optimized to avoid INDELS creation during the editing process. So far it does not replace the capacity of BEs when BEs can be used to perform the desired genetic modification. However, this very promising genetic tool has the potential to increase drastically the editing possibilities on the type of genetic modifications and the presence of the PAM is less limiting than in Base Editing.

First, the length of the RT template of the pegRNA can be adjusted to get a nick upstream of the mutation integration site and PE3 was indeed shown to generate base changes up to 34 bp downstream of the nick<sup>140</sup>. Second, there is no concern about multiple editing as it is the case for base editing when there are several Cs in the editing window. Although the PAM requirements in Prime Editing is much less stringent than in Base Editing, for some loci no PAM can be found. Therefore, being flexible with the PAM sequence to design different pegRNAs for the same mutation would be highly valuable as a slight difference in the pegRNA design can make a significant difference in the gene editing efficiency. For these reasons, intensive work has been done in order to engineer many Cas9 recognizing other PAMs than the classical 5'-NGG-3' sequence, some of which have been already implemented in BE and PE systems.

### I.3.3. CAS9 VARIANTS RECOGNIZING OTHER PAMS.

Although CRISPR/Cas9 is highly efficient and flexible to generate knock-out alleles in numerous laboratory animal models, base editing and prime editing present some important limitations including the design of the guide RNAs often constrained to the presence of a PAM at the right distance where the modification has to be introduced.

Intensive work has been done and is still on-going in order to get more flexibility in the design of the guide RNAs. The current biggest goal is to engineer an efficient PAM-less Cas9 nuclease variant. With the emergence of the new base editing and prime editing technologies, these new engineered flexible Cas9 are particularly necessary to develop these new strategies as genetic tools of choice for the genetic engineering of laboratory model organisms and for gene therapy in humans. Indeed, the classical and most efficient *SpCas9* requires the 5'-NGG-3' sequence downstream the sequence targeted by the sgRNA to break the DNA strands (Fig. 2.a). With the aim to broaden the genome editing possibilities, several Cas9 showing different PAM specificities, usually a 2-4 bp motif, have been reported (Table 1).

Target sites are defined by the presence of a PAM, essential for the activity of the Cas9, including mutated Cas9 nickases. Some of these new Cas9 have thus been integrated in the base editors and prime editor systems to increase the target scope. Indeed, to overcome the limitations of the restricted window and the bystander editing of base editing, researchers have developed base editors incorporating different Cas9 nuclease enzymes. They implemented in CBEs these diverse nucleases: xCas9 (NGN PAMs)<sup>195, 196</sup>, *SpCas9-VQR/SpCas9-VRQR* (NGA PAMs)<sup>197, 198, 199</sup>, *SpCas9-VRER* (NGCG PAMs)<sup>197, 198</sup>, *SpCas9-NG* (NG PAMs)<sup>196, 200, 201</sup>, EQR-Cas9 (NGAG PAMs)<sup>197, 198, 199</sup>, *SaCas9* (NNGRRT PAMs)<sup>155, 197, 198, 202</sup>, KKH-*SaCas9* (NNNRRT PAMs)<sup>198, 203</sup>, *SpymacCas9* (NAA PAMs)<sup>141, 202, 204</sup>, *SauriCas9* (NNGG PAMs)<sup>205</sup> and ScCas9 (NNG PAMs)<sup>206</sup>. Moreover, ABE systems using some of these different enzymes have also been created: VQR-ABE<sup>207</sup>, VRGR-ABE<sup>208</sup>, VRER-ABE<sup>207, 208</sup>, *SaCas9*-ABE<sup>208, 209</sup> and *SaKKH*-ABE<sup>208, 210, 211</sup>, xABE<sup>195, 208</sup> and NG-ABE<sup>207, 208</sup>.

More recently, two Cas9 variants and their implementations in the last versions of CBE and ABE (CBE4max and ABEmax<sup>157</sup>) have been reported: the SpG-Cas9 recognizing NG PAMs and the SpRY-Cas9 as a near-Pamless Cas9<sup>212</sup>. The CBE4max-SpRY is particularly promising as it can edit nearly all PAMs, with a higher efficiency

with the NRN PAMs (R=A or G). No studies have reported the CBE4max-SpRY working in animal models so far.

**In this context, we developed the *SpymacAncBE4max* (c.f. Chapter II.1) and the highly flexible CBE4max-SpRY in zebrafish (c.f. Chapter II.2). We furthermore aimed at integrating the SpRY domain in the prime editor (c.f. Chapter II.5).**

In the meantime, a study has reported many PE2 variants engineered and tested with several PAMs in cultured cells: PE2-VQR, PE2-VRQR, PE2-VRER, PE2-NG, PE2-SpG, and PE2-SpRY<sup>213</sup>. Using the PE2 approach at non-NGG NGN sites, gene editing was obtained for all of these variants with a highest activity reached using the PE2-SpG. Moreover, the design of pegRNAs with 13-nt PBS and 14-nt RT template was the one with which the highest efficiency for PE2-NG, PE2-SpG and PE2-SpRY was achieved, with respectively 12.8%, 26.9% and 21.4% of efficiency for C-to-G conversion using the 5'-AGC-3' PAM. The flexibility of the PE2-SpRY as the SpRY-Cas9 was tested and has been reported as a near-PAMless nuclease, and gene editing for NNN PAMs was obtained. In addition, a low off-target frequency and an increase of prime editing efficiency up to 2.8-fold using PE3 and PE3b approaches combined with the PE2-SpRY and without an increase of INDELS was observed. Finally, in this study the authors could show a higher editing efficiency using the PE2-SpG than the “classical” PE2 for the same targets. Overall, this demonstrates the important benefit to have new variants for these technologies and to develop two highly versatile tools, PE2-SpG and PE2-SpRY.

Cas9 variant	PAM (5'-3')	ABE variant	CBE variant	PE2 variant
xCas9 (195)	NGN	xCas9(3.7)-ABE (195) xABE <sub>max</sub> (208)	xCas9(3.7)-BE3 (195, 196)	
SpCas9-VQR/ SpCas9-VRQR (197)	NGAN	VQR-ABE (207) VRQR-ABE <sub>max</sub> (208)	VQR-BE3 (198) VQR-BE4-gam (199)	PE2-VQR (213) PE2-VRQR (213)
SpCas9-VRER (197)	NGCG	VRER-ABE (207) VRER-ABE <sub>max</sub> (208)	VRER-BE3 (198)	PE2-VRER (213)
SpCas9-NG (200)	NG	NG-ABE (207) NG-ABE <sub>max</sub> (208)	hA3A-BE3-NG (201) CBE-NG (196)	PE2-NG (213)
EQR-Cas9 (197)	NGAG		EQR-BE3 (198) EQR-BE4-gam (199)	
SaCas9 (197, 202)	NNGRRT	SaCas9-ABE (209) SaABE <sub>max</sub> (208)	SaBE3 (198) SaBE4 (155) SaBE4-gam (155)	
KKH-SaCas9 (203)	NNNRRT	SaKKH-ABE (210, 211) SaKKH-ABE <sub>max</sub> (208)	SaKKH-BE3 (198)	
SauriCas9 (205)	NNGG	SauriABE <sub>max</sub> (205)	SauriBE4 <sub>max</sub> (205)	
SpymacCas9 (204)	NAA	SpymacABE <sub>max</sub> (202)	Spymac-BE3 (204) SpymacBE4 <sub>max</sub> (202) SpymacAncBE4 <sub>max</sub> (141)	
SpG-Cas9 (212)	NG	ABE <sub>max</sub> -SpG (212)	CBE4 <sub>max</sub> -SpG (212)	PE2-SpG (213)
SpRY-Cas9 (212)	NRN (NYN)	ABE <sub>max</sub> -SpRY (212)	CBE4 <sub>max</sub> -SpRY (212)	PE2-SpRY (213)
ScCas9 (206)	NNG	ScCas9-ABE(7.10) (206)	ScCas9-BE3 (206)	

**Table 1: Cas9 variants recognizing other PAMs than the classical 5'-NGG-3' PAM.**

#### I.4. AIMS OF THE STUDY.

Zebrafish is a model organism of choice to study developmental biology and has become an important animal model for disease modeling, including genetic disorders. However, the generation of mutant lines harboring precise pathological mutations is still a challenge in the field of gene editing for this vertebrate model.

In this study we investigate the capacities of several new cytosine base editor variants engineered from the BE3, including several CBE recognizing other PAMs than the classical 5'-NGG-3' in zebrafish in order to broaden the gene editing possibilities and to perform multiplex mutagenesis. To go further, we also explore biological applications of this technology to complement the current studies and possibilities of disease modeling such as cancer for which the main strategy so far is the overexpression of mutated oncogenes in the tissue of choice. In addition to cancer, we intend to introduce point mutations found in patients which developed a rare polymalformative syndrome to validate their pathogenicity in zebrafish.

Finally, we aim at developing and optimizing the versatile prime editing technology in this animal model, a recent technology which is still poorly efficient although highly promising.



# CHAPTER II: RESULTS

## II.1. ESTABLISHMENT AND USE OF SEVERAL CYTOSINE BASE EDITORS IN ZEBRAFISH.

### II.1.1. SUMMARY ARTICLE 1.

Dissecting the genetic alterations responsible for the cellular changes leading to disease is a central goal of medical research. It is thus necessary to be able to mimic mutations found in human patient in animal models, especially to further develop personalized medicine approaches. We therefore decided to develop the cytidine base editor technology in zebrafish for which the generation of knock-ins is still poorly efficient in order to introduce specific C-to-T conversions. This ability is of high interest to generate stop codons at defined locations and to mimic pathologies by reproducing point mutations found in humans. In the article 1, we have established 3 different CBE variants all working and never published in zebrafish, providing new targeting possibilities in this animal model. Using these CBE variants, we were able to:

- Efficiently generate C-to-T conversions up to 83%, without any unwanted mutations using the BE4-gam CBE variant.
- Introduce a C-to-T conversion up to 91% of efficiency using the ancBE4max variant, and for C bases for which we did not get any editing using the BE4-gam variant.
- Perform base editing using the *Spymac*-ancBE4max, a new base editor recognizing NAA PAM. This new variant never established for zebrafish before has increased the gene editing possibilities in this animal model.
- Endogenously activate the major developmental Wnt signaling pathway by introducing specifically the S33L mutation in  $\beta$ -catenin leading to its constitutive activation.
- Develop a python script in order to facilitate the design of sgRNAs to introduce precise STOP codon and made it available to the community through this work.

- Target 10 cancer genes, performing loss-of-function for 7 tumor suppressor genes and gain-of-function (GOF) for 3 oncogenes.
- Create a new model for dwarfism with the loss-of-function of *cbl* tumor suppressor gene.

## II.1.2. ARTICLE I: PRECISE BASE EDITING FOR THE *IN VIVO* STUDY OF DEVELOPMENTAL SIGNALING AND HUMAN PATHOLOGIES IN ZEBRAFISH.

Rosello M, *et al.* Precise base editing for the *in vivo* study of developmental signaling and human pathologies in zebrafish. *Elife* **10**, (2021).

# Precise base editing for the *in vivo* study of developmental signaling and human pathologies in zebrafish

Marion Rosello<sup>1,2</sup>, Juliette Vouigny<sup>2</sup>, François Czarny<sup>1</sup>, Marina C Mione<sup>3</sup>, Jean-Paul Concordet<sup>4</sup>, Shahad Albadri<sup>1\*</sup>, Filippo Del Bene<sup>1,2\*</sup>

<sup>1</sup>Sorbonne Université, INSERM, CNRS, Institut de la Vision, Paris, France; <sup>2</sup>Institut Curie, PSL Research University, Inserm U934, CNRS UMR3215, Paris, France; <sup>3</sup>Department of Cellular, Computational and Integrative Biology – CIBIO, University of Trento, Trento, Italy; <sup>4</sup>Muséum National d'Histoire Naturelle, INSERM U1154, CNRS UMR 7196, Paris, France

**Abstract** While zebrafish is emerging as a new model system to study human diseases, an efficient methodology to generate precise point mutations at high efficiency is still lacking. Here we show that base editors can generate C-to-T point mutations with high efficiencies without other unwanted on-target mutations. In addition, we established a new editor variant recognizing an NAA protospacer adjacent motif, expanding the base editing possibilities in zebrafish. Using these approaches, we first generated a base change in the *ctnnb1* gene, mimicking oncogenic an mutation of the human gene known to result in constitutive activation of endogenous Wnt signaling. Additionally, we precisely targeted several cancer-associated genes including *cbl*. With this last target, we created a new zebrafish dwarfism model. Together our findings expand the potential of zebrafish as a model system allowing new approaches for the endogenous modulation of cell signaling pathways and the generation of precise models of human genetic disease-associated mutations.

**\*For correspondence:**

shahad.albadri@inserm.fr (SA);  
filippo.del-bene@inserm.fr (FDB)

**Competing interests:** The authors declare that no competing interests exist.

**Funding:** See page 12

**Received:** 07 December 2020

**Accepted:** 10 February 2021

**Published:** 12 February 2021

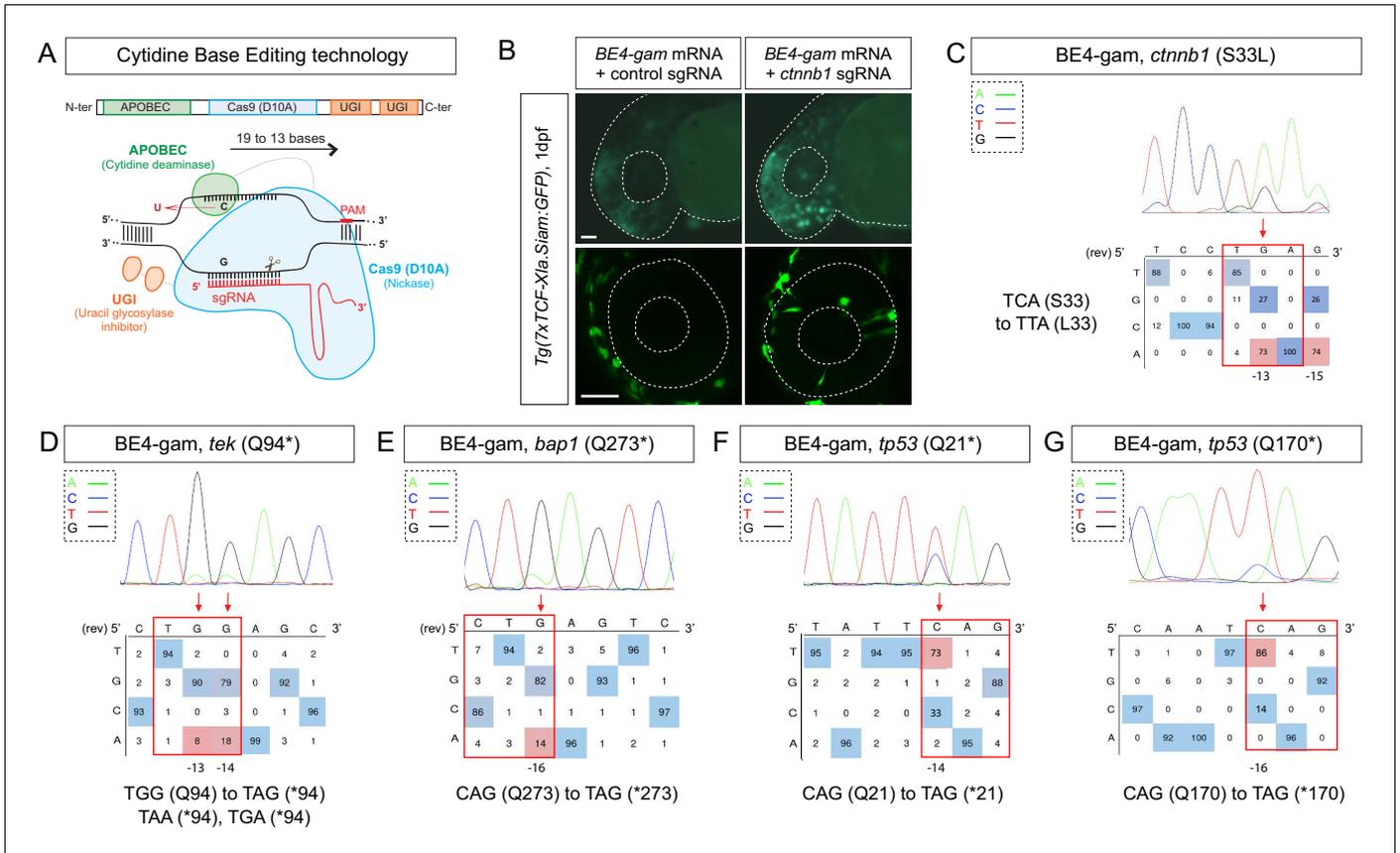
**Reviewing editor:** Tanya T Whitfield, University of Sheffield, United Kingdom

© Copyright Rosello et al. This article is distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use and redistribution provided that the original author and source are credited.

## Introduction

With the recent technological advances in precise gene editing, the use of zebrafish in genetic engineering studies has drastically increased in the last 5 years (*Patton and Tobin, 2019; Santoriello and Zon, 2012*). The CRISPR (clustered regularly interspaced short palindromic repeats)/Cas9 system is indeed a remarkably powerful gene-editing tool (*Sander and Joung, 2014*) that enables the rapid and efficient generation of loss-of-function mutations in this animal model. This system relies on the specific binding of a sgRNA-Cas9 complex that initially interacts with DNA 20 base pair (bp) upstream of a NGG protospacer adjacent motif (PAM) sequence that triggers the Cas9 protein to introduce a double-strand break (DSB). This technique is nowadays widely used in zebrafish notably to produce knock-out alleles (*Hwang et al., 2013*), and more recently, it has also been demonstrated that CRISPR/Cas9-mediated homology-directed repair (HDR) can be used to introduce exogenous DNA and single-nucleotide polymorphisms (*Prykhozhij et al., 2018; Tessadori et al., 2018; Wierson et al., 2020*).

Recently, a CRISPR/Cas9-based technology has been developed to precisely edit single bases of DNA without introducing DSBs in human cells (*Koblan et al., 2018; Komor et al., 2016; Komor et al., 2017*). The method is based on the fusion of a Cas9-D10A nickase with a cytidine deaminase giving rise to a cytidine base editor (CBE). CBE converts C-to-T bases in a restricted window of 13–19 nucleotides (nt) upstream of the PAM sequence (*Figure 1A*). In zebrafish, a CBE was shown to work but with limited efficiencies, inducing less than 29% of gene editing and, in most



**Figure 1.** Efficient endogenous activation of Wnt signaling pathway and tumor suppressor genes targeting using BE4-gam in zebrafish. (A) Schematic representation of the cytidine base editor technology. (B) Activation of Wnt signaling via S33L mutation in  $\beta$ -catenin. 1 dpf *Tg(7xTCF-Xla.Siam:GFP)* representative embryos injected with BE4-gam mRNA and *ctnnb1* (S33L) sgRNA or control scrambled sequence. The upper panel shows an overall increase of GFP-positive cells in the head/anterior region upon the injection of the BE4-gam mRNA and *ctnnb1* (S33L) sgRNA compared to the control situation. The lower panel shows maximal z-projection of lateral view of the injected embryos where ectopic GFP signal in retinal progenitor cells (white stars) can be detected, whereas control embryos do not show any fluorescence in the retina at this stage. (C–G) DNA sequencing chromatogram of targeted loci with the BE4-gam and obtained C-to-T conversion efficiencies. The chromatograms correspond to the highest efficiency reported for the single embryos analyzed as detailed in **Table 2**. (C) S33L mutation in  $\beta$ -catenin upon C-to-T conversion in *ctnnb1* reached 73% of gene-editing efficiency. The other edited C led to a silent mutation GAC (D) to GAT (D). (D) Q94\* mutation in Tek upon C-to-T conversion in *tek* reached 18% of gene-editing efficiency. (E) Q273\* mutation in Bap1 upon C-to-T conversion in *bap1* reached 14% of gene-editing efficiency. (F) Q21\* mutation in p53 upon C-to-T conversion in *tp53* reached 73% of gene-editing efficiency. (G) Q170\* mutation in p53 upon C-to-T conversion in *tp53* reached 86% of gene-editing efficiency. For (C) and (E), the reverse complement of the sgRNA sequence is shown. Scale bars: (B) 50  $\mu$ m. (D–G) Numbers in the boxes represent the percentage of each base at that sequence position. In red are highlighted the base substitutions introduced by base editing, while the original bases are in blue. The color code of the chromatogram is indicated in the upper left corner (Adenine green, Cytosine blue, Thymine red, Guanine black). The distance from the PAM sequence of the targeted C base is indicated below each chromatogram. It is considered that the quantifications under 5% are due to the background signal from Sanger sequencing and are thus non-significant (Kluesner et al., 2018). The online version of this article includes the following figure supplement(s) for figure 1:

**Figure supplement 1.** List of targeted loci.

cases, at least 5% of unwanted INDEL (insertion or deletion) mutations were also detected (Carrington et al., 2020; Zhang et al., 2017). For these reasons, this editing strategy has not been so far favored by the zebrafish community. However, since this first generation of CBEs, several studies in cell culture have optimized and engineered new base editor variants with increased gene-editing efficiency reaching up to 90% without creating undesired INDEL mutations (Koblan et al., 2018). A recent study reported the use of a second-generation CBE to generate a zebrafish model of human ablepharon macrostomia syndrome (Zhao et al., 2020).

Recent progress has also been made on the generation of CBEs able to recognize other PAM sequences, allowing to broaden gene-editing possibilities (Jakimo *et al.*, 2018; Koblan *et al.*, 2018). Thus, base editing offers a complementary and powerful approach in zebrafish to introduce specific single-nucleotide variants into the zebrafish genome. Here based on these technological advances, we optimized these second-generation gene-editing tools in zebrafish. As reported in *ex vivo* studies (Koblan *et al.*, 2018), we tested different CBE variants and obtained highly efficient C-to-T conversion, reaching up to 91% efficiency without unwanted mutations and expanded base editing possibilities using a CBE variant recognizing the NAA PAM. Furthermore, compared to previous studies, here we used these tools to target Wnt signaling, thus proving that endogenous pathways can be modulated in their natural context. Finally, we demonstrated the power of this technology for introducing precise mutations in human cancer-associated genes with high efficiency in zebrafish and created a new fish model for dwarfism.

## Results and discussion

### BE4-gam base editing for the endogenous activation of Wnt signaling pathway

To date, the main strategies used in zebrafish to study the constitutive activation of signaling pathways and to dissect their role during embryonic development or tumorigenesis were based on over-expressing mutated genes. To gain further insights and to complement these studies, an important requirement is to have the ability to maintain the endogenous genetic and regulatory contexts by generating mutations of endogenous genes *in vivo*.

To address this challenge, we decided to introduce an activating mutation in the *ctnnb1* gene coding for the key effector  $\beta$ -catenin of canonical Wnt signaling, a major signaling pathway during embryonic development which is activated in many cancers (Steinhart and Angers, 2018). It was previously shown that the mutation of the Serine33 of the human  $\beta$ -catenin protein into a Leucine prevents its degradation by the ubiquitin-proteasome system, leading to its stabilization and thereby to the constitutive activation of Wnt signaling pathway (Hart *et al.*, 1999; Liu *et al.*, 1999).

We first aimed at introducing this mutation in the genome of the zebrafish by using the Base Editor 4 fused to the gam domain (BE4-gam) (Figure 1A). This CBE was indeed one of the first variants of CBEs to show high efficiency of gene editing and fewer INDELS formation in cultured cells (Komor *et al.*, 2017). We injected the BE4-gam mRNA and synthetic *ctnnb1* S33L sgRNA into one-cell stage *Tg(7xTCF-Xla.Siam:GFP)* zebrafish embryos to directly monitor the effect of the introduced mutation on the activity of the canonical Wnt signaling (Moro *et al.*, 2012). Upon *ctnnb1* S33L sgRNA injection, we observed an increase of GFP-positive cells at 1 dpf ( $n = 39/50$  embryos) compared to the control embryos ( $n = 27$  embryos) resulting from three independent experiments. By confocal imaging we quantified ectopic activation of the pathway in retinal progenitor cells and observed an average of 12 GFP-positive clones per retina, while GFP-positive cells were never detected in the retina of control injected *Tg(7xTCF-Xla.Siam:GFP)* embryos ( $n = 4$  each) (Figure 1B). Using this strategy, we observed base editing in five of eight randomly chosen embryos and were able to reach up to 73% of editing efficiency in single embryo analysis (Table 1, Table 2). In addition, we also observed the conversion of another cytidine within the PAM [−19, −13 bp] window leading to a silent mutation (GAC-to-GAT (D)) in four of the eight analyzed embryos with up to 74% efficiency (Figure 1C, Table 1, Table 2, Figure 1—figure supplement 1).

With these results, we demonstrated that it is now possible to constitutively and efficiently activate important developmental signaling pathways in their endogenous context, as we show here for Wnt signaling. Furthermore, several studies have implicated the S33L  $\beta$ -catenin mutation in tumorigenesis, making it possible to study the role of this oncogenic mutation in cancer development in zebrafish. In order to test the potential of CBE targeting in cancer modeling, we next decided to use it to target a series of tumor suppressor genes and oncogenes using the same editing strategy applied to endogenous  $\beta$ -catenin.

### Base-editing strategies for the generation of human cancer mutations

Zebrafish is a powerful model system to study cancer genetics *in vivo* (Cagan *et al.*, 2019; Cayuela *et al.*, 2018). However, a robust method for modeling cancer-associated mutations in

**Table 1.** Base-editing efficiency using different CBE variants.

Number of edited embryos randomly chosen after injection of CBE mRNA and sgRNA. The efficiency varies between non-detected (n.d.) and 91% depending on the targeted locus, the sgRNA, and the CBE used. Editing efficiency was quantified by editR analysis (Kluesner et al., 2018), which does not detect editing efficiency below 5%.

Targeted gene CBE used induced mutation	<i>ctnnb1</i> (S33L) BE4-gam	<i>tp53</i> (Q170*) BE4-gam	<i>cbl</i> (W577*) BE4-gam	<i>kras</i> (E62K) BE4-gam	<i>Kras</i> (E62K) ancBE4max	<i>dmd</i> (Q8*) BE4-gam	<i>dmd</i> (Q8*) ancBE4max	<i>rb1</i> (W63*) ancBE4max	<i>nras</i> (G13S) spymac- ancBE4max	<i>tp53</i> (Q170*) spymac- ancBE4max
Number of edited embryos	5/8	7/8	8/10	0/8	4/7	0/8	2/4	8/8	2/4	1/4
Highest obtained efficiency	73%	86%	C16 35% C15 50%	n.d.	19%	n.d.	14%	C17 91% C16 65%	19%	16%

**Table 2.** Editing efficiency quantification.

Editing quantification of up to 10 single embryos randomly chosen after injection of indicated CBE mRNA and sgRNA. The efficiency varies between non-detected (n.d.) to 91% in a single embryo depending on the targeted locus, the sgRNA, and the CBE used. Editing efficiency was quantified by editR analysis (Kluesner et al., 2018), which does not detect editing efficiency below 5%.

Targeted gene CBE used	Number of edited embryos	Emb. 1	Emb. 2	Emb. 3	Emb. 4	Emb. 5	Emb. 6	Emb. 7	Emb. 8	Emb. 9	Emb. 10
<i>ctnnb1</i> (S33L) BE4-gam	5/8	C15 74% C13 73%	C15 n.d. C13 40%	C15 44% C13 25%	C15 7% C13 16%	C15 n.d. C13 11%	n.d.	n.d.	n.d.	–	–
<i>tek</i> (Q94*) BE4-gam	5/8	C14 18% C13 8%	C14 10% C13 n.d.	C14 8% C13 n.d.	C14 6% C13 n.d.	C14 8% C13 9%	n.d.	n.d.	n.d.	–	–
<i>Bap1</i> (Q273*) BE4-gam	4/8	14%	12%	9%	8%	n.d.	n.d.	n.d.	n.d.	–	–
<i>tp53</i> (Q21*) BE4-gam	6/8	63%	33%	37%	58%	8%	50%	n.d.	n.d.	–	–
<i>tp53</i> (Q170*) BE4-gam	7/8	86%	46%	51%	62%	45%	20%	33%	n.d.	–	–
<i>cbl</i> (W577*) BE4-gam	8/10	C16 35% C15 50%	C16 19% C15 31%	C16 22% C15 38%	C16 25% C15 41%	C16 20% C15 35%	C16 7% C15 9%	C16 7% C15 12%	C16 10% C15 17%	n.d.	n.d.
<i>kras</i> (E62K) BE4-gam	0/8	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	–	–
<i>kras</i> (E62K) ancBE4max	4/7	C17 19% C16 21%	C17 8% C16 11%	C17 6% C16 8%	C17 9% C16 10%	n.d.	n.d.	n.d.	–	–	–
<i>dmd</i> (Q8*) BE4-gam	0/8	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	–	–
<i>dmd</i> (Q8*) ancBE4max	2/4	14%	6%	n.d.	n.d.	–	–	–	–	–	–
<i>sod2</i> (Q145*) ancBE4max	8/8	64%	45%	21%	54%	52%	24%	26%	33%	–	–
<i>rb1</i> (W63*) ancBE4max	8/8	C19 n.d. C17 91% C16 65%	C19 21% C17 79% C16 75%	C19 n.d. C17 27% C16 18%	C19 13% C17 81% C16 60%	C19 8% C17 48% C16 33%	C19 13% C17 76% C16 64%	C19 13% C17 78% C16 69%	C19 21% C17 77% C16 63%	–	–
<i>nras</i> (G13S) spymac-ancBE4max	2/4	19%	18%	n.d.	n.d.	–	–	–	–	–	–
<i>tp53</i> (Q170*) spymac-ancBE4max	1/4	16%	n.d.	n.d.	n.d.	–	–	–	–	–	–

zebrafish is still lacking to date. We decided to create predictable premature stop codons in tumor suppressor genes and to generate activating mutations in oncogenes of the RAS family (Li *et al.*, 2018) in order to test the ability of CBEs to induce cancer-related mutations in zebrafish.

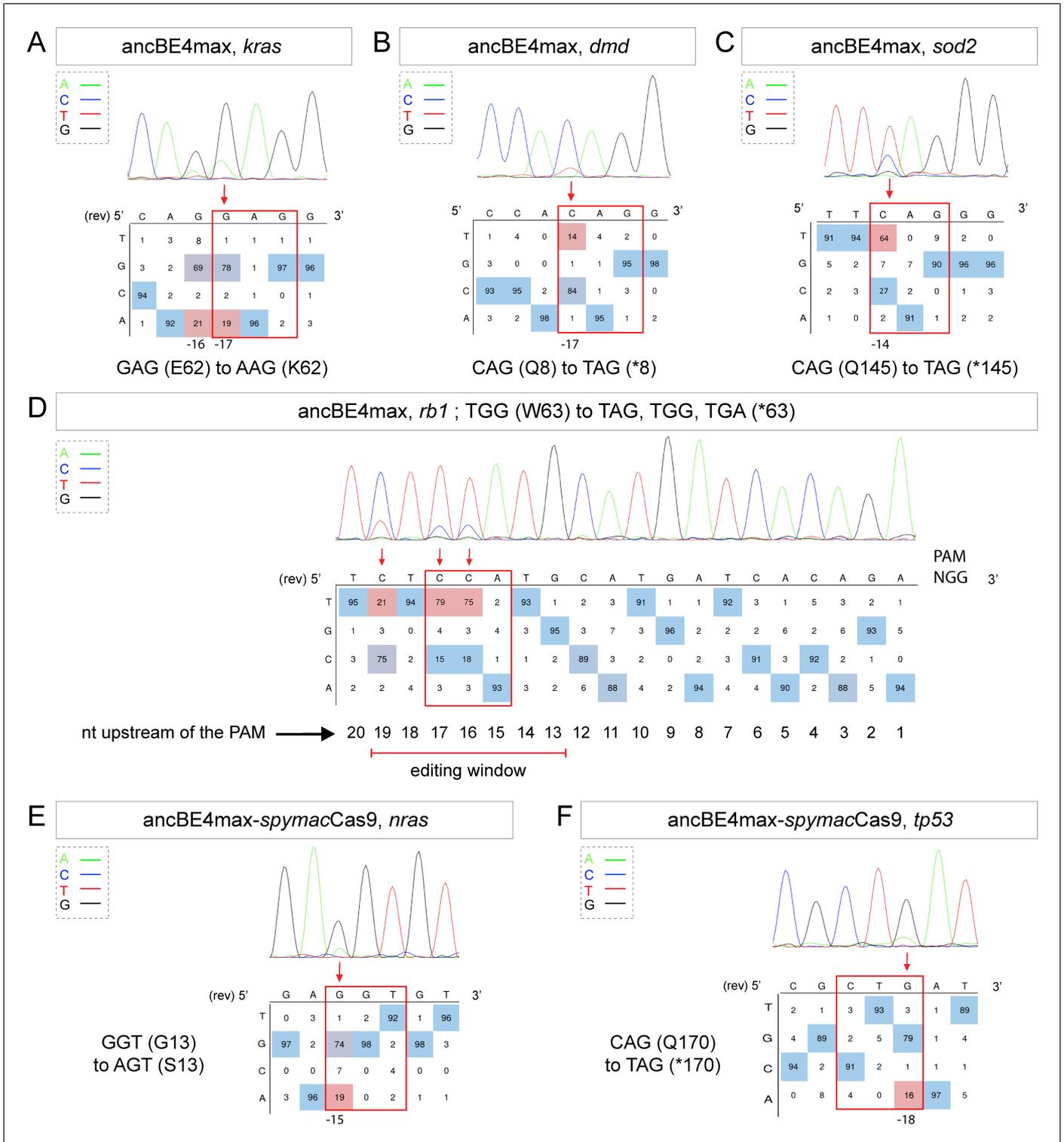
We first developed an automated script to rapidly detect codons allowing to generate nonsense mutations after a C-to-T conversion within the restricted PAM [−19, −13] bp editing window (Source code 1). Using this script, we designed a series of sgRNAs targeting in a selection of tumor suppressor genes. We induced the Q94\* mutation in Tek in five of eight randomly chosen embryos, Q273\* mutation in Bap1 in four of eight randomly chosen embryos and Q21\* mutation in p53 in six of eight randomly chosen embryos as well as Q170\* in p53 in seven of eight randomly chosen embryos by C-to-T conversions (Figure 1D–G, Table 1, Table 2). Among the different targeted mutations, the highest efficiency was achieved with the *tp53* tumor suppressor gene, for which we reached up to 86% of C-to-T conversion for the introduction of the Q170\* mutation (Figure 1G, Table 1, Table 2). To assess the presence of INDELS or unwanted mutations upon BE4-gam injections in our targets, we amplified, cloned, and sequenced all targeted loci. For *tek* 6 of 20, *bap1* 2 of 12, *tp53* (Q21\*) 12 of 24, and lastly *tp53* (Q170\*) 21 of 24 colonies showed precise C-to-T conversions, whereas all the other analyzed sequences were wild type, without any error or INDEL formation. Together these results show that, using BE4-gam, we efficiently targeted several genes implicated in tumorigenesis in zebrafish without generating any unwanted INDELS, unlike what was previously reported with the BE3 variant.

More recently, a new CBE variant, the ancBE4max, has been engineered and optimized in cell culture with increased efficiency compared to the classical BE4-gam, reaching up to 90% efficiency and very low rates of INDELS (Koblan *et al.*, 2018). We therefore decided to use this new CBE variant to target the oncogenic mutation E62K in *Kras* and induce the creation of a Q8\* stop codon in the *Dmd* tumor suppressor for which we did not obtain any C-to-T conversions or other unwanted changes using the BE4-gam in randomly chosen eight embryos for each condition (Table 1, Table 2). By co-injection of *ancBE4max* mRNA with the *kras* E62K sgRNA, we were able to introduce the E62K mutation in four of seven randomly chosen embryos and we were able to reach up to 19% of editing efficiency in single embryo analysis. Another cytidine in the editing window was also converted and led to the generation of a silent mutation (CAG-to-CAA (Q)) (Figure 2A, Table 1, Table 2). Similar to what we observed in the case of *kras* editing, we were able to obtain a Q8\* mutation in the *Dmd* tumor suppressor in two of four randomly chosen embryos with up to 14% of editing efficiency (Figure 2B, Table 1, Table 2). Thus, with this new ancBE4max variant, we are able to introduce mutations that could not be achieved with BE4-gam using the same sgRNAs. Remarkable editing efficiency was also observed using this CBE for two additional targets: the tumor suppressor genes *sod2* and *rb1*, for which, respectively, up to 64% and 91% of editing were reached and 100% of the sequenced embryos were precisely mutated (from single embryo analysis of  $n = 8$  randomly analyzed embryos) (Figure 2C,D, Table 1, Table 2; Bravard *et al.*, 1992; Dyson, 2016).

It is interesting to note that in general all the cytidine bases present in the PAM [−19, −13] bp window can be edited by the CBE, with a higher efficiency for the cytidine bases located in the middle of this window while editing was below detection levels for cytidines located only 12 bp upstream (Figure 2D). With the use of ancBE4max CBE, these results highlight the importance of the cytidine distance from the PAM for efficient editing in zebrafish as shown previously in cell culture assays (Gaudelli *et al.*, 2017).

## Expanding gene-editing possibilities in zebrafish using a CBE recognizing NAA PAM

Due to the PAM-dependent restriction of the editing window, many mutations could not be generated so far. We therefore decided to expand the editing possibilities in zebrafish by associating *SpymacCas9* recognizing NAA PAMs with the efficient conversion capacity of the ancBE4max. To this end, we replaced the PAM-interacting motif (PIM) domain of the *SpCas9* with the one of the *SpymacCas9* in the ancBE4max (Jakimo *et al.*, 2018). The inserted PIM domain was codon optimized for zebrafish. Using this newly generated ancBE4max-*SpymacCas9*, we were able to reproduce the human G13S mutation in *Nras* oncogene in zebrafish in 2 out of 4 randomly analyzed embryo with up to 19% of efficiency reached in single embryo analysis (Figure 2E, Table 1, Table 2). We also introduced a stop codon by a C-to-T conversion in the *tp53* gene in 1 out of 4 randomly analyzed embryo with 16% efficiency (Figure 2F, Table 1, Table 2). These results demonstrate that in addition



**Figure 2.** Tumor suppressor genes and oncogenes targeting by the highly efficient ancBE4max and the ancBE4max-SpymacCas9 recognizing NAA PAM. (A–F) DNA sequencing chromatogram of targeted loci with the ancBE4max (in A–D) or ancBE4max-SpymacCas9 (in E, F) and obtained C-to-T conversion efficiencies. (A) E62K mutation in *Kras* upon C-to-T conversion in *kras* reached 19% gene-editing efficiency. The other edited C led to a silent mutation CAG (Q) to CAA (Q). (B) Q8\* mutation in *Dmd* upon C-to-T conversion in *dmd* reached 14% of gene-editing efficiency. (C) Q145\* mutation in *Sod2* upon C-to-T conversion in *sod2* reached 64% of gene-editing efficiency. (D) W63\* mutation in *Rb1* upon C-to-T conversion in *rb1*

Figure 2 continued on next page

Figure 2 continued

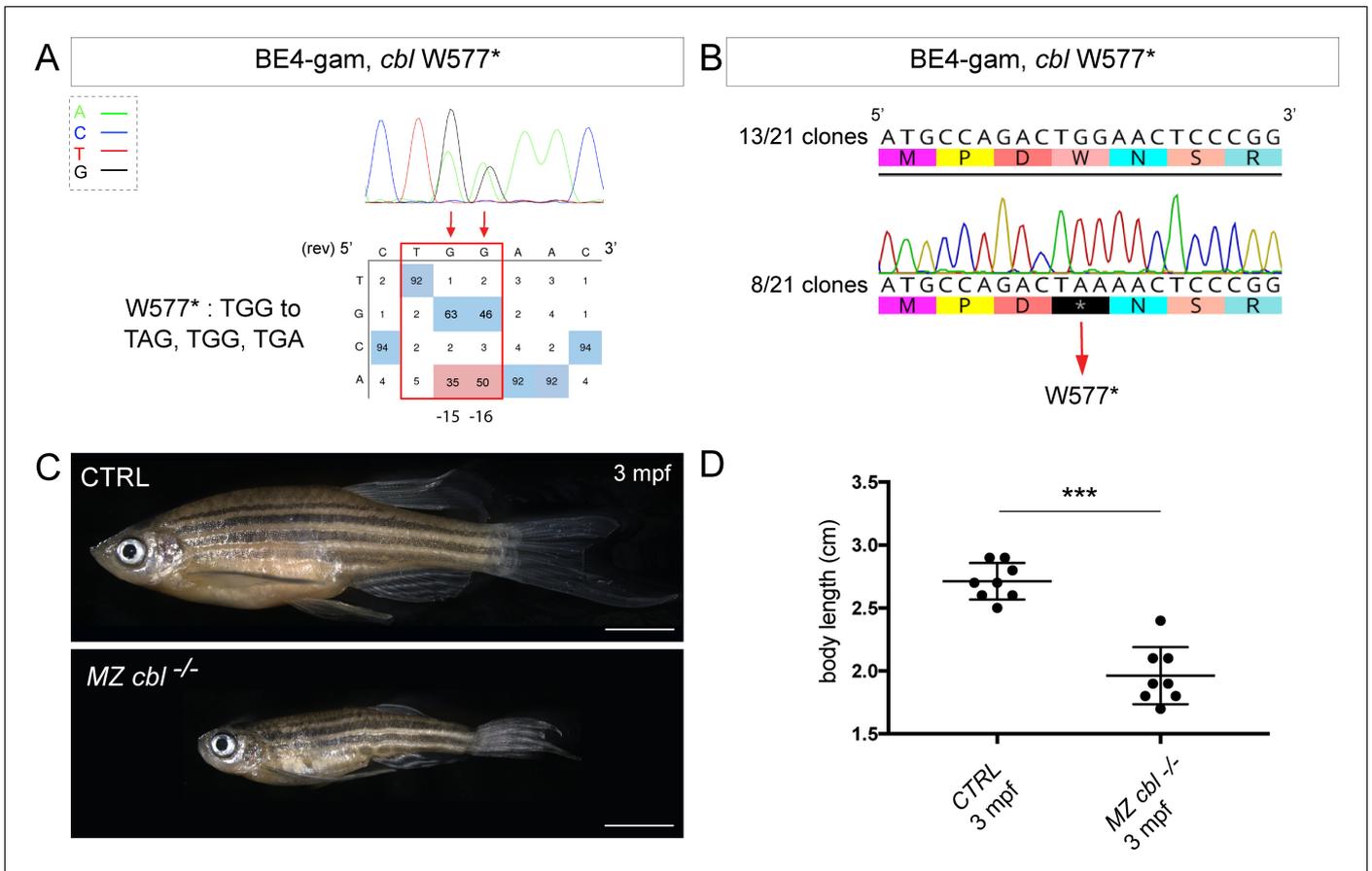
reached 21% for the C19 base, 79% for C17, and 75% for the C16 of gene-editing efficiency. (E) G13S mutation in *Nras* upon C-to-T conversion in *nras* reached 19% of gene-editing efficiency. (F) Q170\* mutation in p53 upon C-to-T conversion in *tp53* reached 16% of gene-editing efficiency. For (A, D–F), the reverse complement of the sgRNA sequence is shown. (A–F) The chromatograms correspond to the efficiency reported for the single embryos provided in the first column of **Table 2**. The numbers in the boxes represent the percentage of each base at that sequence position. In red are highlighted the base substitutions introduced by base editing, while the original sequence is in blue. The color code of the chromatogram is indicated in the upper left corner (Adenine green, Cytosine blue, Thymine red, and Guanine black). The distance from the PAM sequence of the targeted C base is indicated below each chromatogram. It is considered that the quantifications under 5% are due to the background signal from Sanger sequencing and are thus non-significant (Kluesner et al., 2018).

to the classical NGG PAM it is now also possible to target NAA PAMs in zebrafish, thereby significantly expanding the range of cytidine bases that can be converted. For these new CBEs, we added a function in our script to choose the PAM recognized by the Cas9-D10A of the chosen CBE to generate the desired base editing (**Source code 1**). Together with the use of the ancBE4max and ancBE4max-SpymacCas9 CBE variants, we were now able to target mutations that could not be generated with the BE4-gam base editor and reproduce a wider range of human cancer mutations in zebrafish.

Genetic alterations that lead to oncogene activation and/or tumor suppressor inactivation are responsible for tumorigenesis. It is indeed well-established that in cancer patients, a series of genetic mutations in tumor suppressor genes and/or oncogenes are combined to all together lead to the appearance of the disease (Dash et al., 2019). With these efficient genetic tools that are now established in zebrafish, we have the possibility to rapidly test precise combinations of mutations identified in cancer patients.

### Precise gene editing in the *cbl* tumor suppressor gene for the generation of human disease phenotypes in zebrafish

With the technological advances in CRISPR/Cas9 gene editing, zebrafish has become an even more attractive system for modeling human genetic diseases. Among the chosen loci to test the efficiency of the BE4-gam, we targeted the tumor suppressor gene encoding for Cbl, an E3 ubiquitin ligase, that is found mutated in Noonan syndrome patients presenting short stature and other bone malformations among several other phenotypes (Martinelli et al., 2010). In human, activating mutations in the *fibroblast growth factor receptor 3* (*FGFR3*) gene are a leading cause of dwarfism achondroplasia and related dwarf conditions. Indeed, *FGFR3* hyperactivation triggers intracellular signaling within the chondrocytes of the growth plate which terminates its proliferation and bone growth (Harada et al., 2009). Interestingly, another study based on *in vitro* systems reported that some of these activating mutations in *FGFR3* disrupt c-Cbl-mediated ubiquitination that serves as a targeting signal for lysosomal degradation and termination of receptor signaling (Cho et al., 2004). Using the CBE BE4-gam as previously described, we obtained up to 50% of gene-editing efficiency (**Figure 3A**, **Table 1**, **Table 2**), with 80% of the analyzed embryos showing the expected editing ( $n = 10$  randomly analyzed embryos). Four of 15 adults carried the Cbl W577\* mutation in germ cells and one of these carriers transmitted it to 28% of its F1 offspring (44 of 153 analyzed fish carried the mutation). The target sequence was analyzed in the F1 embryos and no INDELS were found (**Figure 3B**). The zygotic homozygous mutant fish (*cbl*<sup>-/-</sup>) deriving from the incross of two heterozygous parents (*cbl*<sup>+/-</sup>) did not develop any obvious phenotype and could be grown to adulthood. This could be due to the fact that in zebrafish, maternal factors stored as mRNAs and proteins in the egg can compensate for zygotic loss of function during embryonic stages. In order to obtain maternal-zygotic mutants (MZ *cbl*<sup>-/-</sup>) that lacked wild-type *cbl* mRNAs and proteins provided by the mother, *cbl*<sup>-/-</sup> mutant parents were incrossed. As controls, the *cbl*<sup>+/+</sup> siblings of the *cbl*<sup>-/-</sup> mutant fish were incrossed in parallel. Interestingly, 24% of the MZ *cbl*<sup>-/-</sup> mutants displayed a significantly reduced overall growth and size by 3 months post-fertilization, while 100% of the progeny of the *cbl*<sup>+/+</sup> sibling fish showed a normal body size (means: 2.7 cm for the wild-type controls and 1.96 cm for the dwarf MZ *cbl*<sup>-/-</sup>, **Figure 3C,D**). Furthermore, this dwarf phenotype was never observed in any of the fish derived from the incrosses of the wild-type stocks used to generate this mutant line, while it was observed in the progeny of two other crosses of the *cbl*<sup>-/-</sup> line. Although we cannot formally exclude the presence of a distinct maternal zygotic mutation linked to the *cbl*<sup>W577\*</sup> allele, our



**Figure 3.** BE4-gam generated *cbl* maternal zygotic mutant fish show a reduced growth phenotype. (A) DNA sequencing chromatogram of targeted *cbl* gene with the BE4-gam. W577\* mutation in *Cbl* upon C-to-T conversion in *cbl* reached 50% for the C16 base and 35% for the C15 base of gene-editing efficiency. The chromatogram refers to the efficiency reported for the embryo provided in the first column of **Table 2**. The numbers in the boxes represent the percentage of each base at that sequence position. In red are highlighted the base substitutions introduced by base editing, while the original sequence is in blue. The color code of the chromatogram is indicated in the upper left corner (Adenine green, Cytosine blue, Thymine red, and Guanine black). The distance from the PAM sequence of the targeted C base is indicated below the chromatogram. It is considered that the quantifications under 5% are due to the background signal from Sanger sequencing and are thus non-significant (Kluesner et al., 2018). (B) Sequencing of individual clones of a pool of F1 embryos from a founder carrying the W577\* mutation in *Cbl*. TGG-to-TAA precise mutation was found in 8 of 21 clones. No editing or INDELS were detected in all other clones. (C) Three months post-fertilization (mpf) *cbl* wild type derived from the incross of wild-type siblings (upper panel) and dwarf maternal zygotic (MZ) mutant fish found in 24% of the progeny (lower panel). (D) Quantification of the body length of the *cbl*<sup>+/+</sup> controls and of the dwarf MZ *cbl*<sup>-/-</sup>. The dwarf fish show a significant reduced size at three mpf compared to the wild-type controls. n = 8 for each group. Mann-Whitney test, p=0,0002. Scale bars: (C) 5 mm.

data strongly support the role of *Cbl* W577\* in the observed phenotype. Four germline mutations located in the RING domain of *Cbl* (Q367P, K382E, D390Y, and R420Q) have been previously identified and associated to Noonan syndrome and related phenotypes (Martinelli et al., 2010). Our results are in line with the growth defect phenotypes observed in these patients and directly implicate *Cbl* loss-of-function as a cause of bone malformations in an animal model. In addition, a point mutation in zebrafish *Cbl* (H382Y) has been implicated in myeloproliferative disorders. Unlike our mutant, *cbl*<sup>H382Y</sup> mutant fish do not survive to adulthood, suggesting that the *Cbl*<sup>W577\*</sup> premature stop reported here may have different consequences on the multiple functions of *Cbl* (Peng et al., 2015). Although not lethal, it would be of interest to assess whether any hematopoietic defects are present in our MZ *cbl*<sup>-/-</sup> mutants or whether this phenotype is only linked to the *Cbl*<sup>H382Y</sup>

substitution found in the *LDD731* zebrafish mutant (Peng et al., 2015). Our model represents a powerful *in vivo* system to dissect the role of Cbl in bone morphogenesis and to explain the human phenotypes related to bone malformations.

## Conclusions

In our work, we took advantage of base editors to generate C-to-T point mutations at unprecedented high efficiencies (up to 91%) without detecting any unwanted mutations that were often problematic when using CBEs in zebrafish. In comparison, previous work has reported an efficiency reaching a maximum of 29% using the BE3 (Zhang et al., 2017). Another more recent study employed the ancBE4max variant in zebrafish with a slight difference of efficiency that might be due to the choice of the specific locus targeted, the synthesis of the sgRNA (homemade vs commercially synthesized) and the injection mode (yolk vs cell) (Carrington et al., 2020). More recently, Zhao et al., 2020 have shown similar efficiencies as we obtained in our study. To expand the gene-editing possibilities in this animal model, we established in addition a new editor variant recognizing the NAA PAM. Using these approaches, we first performed the endogenous and constitutive activation of Wnt signaling by introducing the S33L mutation in  $\beta$ -catenin. In addition, we demonstrated using these strategies that we were able to precisely target several cancer-associated genes for which so far only transgenic over-expressions or imprecise deletions were used to elucidate their functions. Among our targets, the introduced mutation in the *cbl* gene allowed us to generate a new zebrafish model for dwarfism.

Together our work provides a panel of examples whereby, using gene-editing approaches, some of which we established here in zebrafish for the first time, it is now possible to manipulate endogenous signaling pathways, to generate models for human genetic disorders and to mimic precise cancer-associated mutations in zebrafish. While a recent study reported the use of ancBE4max in zebrafish (Zhao et al., 2020), in our work we provide a direct comparison of BE4-gam, ancBE4max, and Spymac-ancBE4max. Our study highlights the power and the need for these approaches to increase the efficiency and the targeting flexibility in order to model pathological human mutations in zebrafish.

Finally, the high efficiencies of CBEs obtained in this study should encourage future applications where they could be implemented with mosaic mutation induction technologies such as the MAZER-ATI (Modeling Approach in Zebrafish for Rapid Tumor Initiation) system (Ablain et al., 2018). This will allow to rapidly model and study *in vivo* combinations of endogenous mutations occurring in specific cancer patients or in genetic disorders caused by somatic mosaicism. Our approach could thus be applied in zebrafish for the precise modeling of complex combinations of cancer-causing mutations in adult animal models as currently possible by transgenic overexpression or somatic gene inactivation (Callahan et al., 2018).

## Materials and methods

### Key resources table

Reagent type (species) or resource	Designation	Source or reference	Identifiers	Additional information
Genetic reagent ( <i>Danio rerio</i> )	<i>Tg(7xTCF-Xla.Siam:GFP)</i>	ZIRC	ZFIN ID: ZBD-ALT-110113-1	
Recombinant DNA reagent	pCMV_BE4-gam (plasmid)	Addgene	Addgene:#100806 RRID:Addgene_100806	
Recombinant DNA reagent	pCMV_ancBE4max (plasmid)	Addgene	Addgene:#112094 RRID:Addgene_112094	
Recombinant DNA reagent	pCS2+_ancBE4max-SpymacCas9 (plasmid)	This paper		See Materials and methods
Commercial assay or kit	NEBuilder HiFi DNA Assembly Cloning Kit	New England Biolabs	Catalog# E5520S	

Continued on next page

Continued

Reagent type (species) or resource	Designation	Source or reference	Identifiers	Additional information
Commercial assay or kit	mMESSAGE mMACHINE T7 Ultra kit	Ambion	Catalog# AM1345	
Commercial assay or kit	mMESSAGE mMACHINE Sp6 kit	Ambion	Catalog# AM1340	
Commercial assay or kit	PCR clean-up gel extraction kit	Macherey-Nagel	Catalog# 740609.50	
Peptide, recombinant protein	Phusion high-fidelity DNA polymerase	ThermoFisher	Catalog# F-530XL	
Software, algorithm	SequenceParser.py	This paper		See <b>Source code 1</b>

### Fish lines and husbandry

Zebrafish (*Danio rerio*) were maintained at 28°C on a 14 hr light/10 hr dark cycle. Fish were housed in the animal facility of our laboratory which was built according to the respective local animal welfare standards. All animal procedures were performed in accordance with French and European Union animal welfare guidelines. Animal handling and experimental procedures were approved by the Committee on ethics of animal experimentation. The *Tg(7xTCF-Xla.Siam:GFP)* line was kindly provided by Sophie Vriz (**Moro et al., 2012**).

### Molecular cloning

To generate the pCS2+\_ancBE4max-SpymacCas9 plasmid, the SpymacCas9 PIM domain sequence has been codon optimized for expression in zebrafish using online software from IDT and synthesized with the first UGI sequence as G-block from IDT. Then, three fragments have been inserted into pCS2+ plasmid linearized with Xho1 using the NEBuilder HiFi DNA Assembly Cloning Kit (New England Biolabs # E5520S): a first fragment of 4161 bp of the ancBE4max to the PIM domain (amplified using the primers F-5'-CGATTCGAATTCAAGGCCCTCATGAAACGGACAGCCGAC-3' and R-5'-CGGTCTGGATCTCGGTCTTTTTCACGATATTC-3'), the Gblock fragment of 803 bp (amplified using the primers F-5'-AAAGACCGAGATCCAGACCGTGGGACAG-3' and R-5'-TCCC GCCGCTATCC TCGCCGATCTTGGAC-3'), and a third fragment of 654 bp of the rest of the ancBE4max from the PIM domain (amplified using the primers F-5'-CGGCGAGGATAGCGGCGGGAGCGGCGGG-3' and R-5'-CTCACTATAGTTCTAGAGGCTTAGACTTTCTCTTCTTGGGCTCGAATTCGCTGCCGTCG-3'). pCMV\_BE4-gam (a gift from David Liu, Addgene plasmid # 100806; **Anzalone et al., 2019**) has been used to generate BE4-gam mRNA *in vitro*. This plasmid was linearized with Pme1 restriction enzyme and mRNAs were synthesized by *in vitro* transcription with 1 µl of GTP from the kit added to the mix, followed by Poly(A) tailing procedure and lithium chloride precipitation (using the mMESSAGE mMACHINE T7 Ultra kit #AM1345, Ambion). pCMV\_ancBE4max (pCMV\_AncBE4max was a gift from David Liu [Addgene plasmid # 112094]) has been linearized using AvrII restriction enzyme; mRNAs were synthesized by *in vitro* transcription with 1 µl of GTP from the kit added to the mix and lithium chloride precipitation (using the mMESSAGE mMACHINE T7 Ultra kit #AM1345, Ambion). The pCS2+\_ancBE4max-SpymacCas9 has been linearized using KpnI restriction enzyme; mRNAs were synthesized by *in vitro* transcription with 1 µl of GTP added to the mix and lithium chloride precipitation (using the mMESSAGE mMACHINE Sp6 kit #AM1340, Ambion).

### sgRNA design

A sequenceParser.py python script was developed and used to design sgRNAs for the creation of a stop codon. The first function of the script is to ask which PAM will be used to then execute the rapid detection of codons that are in the right editing windows from this predefined PAM to generate a STOP in frame after C-to-T conversion. The ORF sequence file extension is .txt and the letters in lower cases. The script can be executed from the command line interface (such as the terminal or PowerShell console) using Python version 3.

Efficiencies of sgRNAs were validated using CRISPOR online tool (*Haussler et al., 2016*). All the synthetic sgRNAs were synthesized by IDT as Alt-R CRISPR-Cas9 crRNA and Alt-R CRISPR-Cas9 tracrRNA.

List of the crRNAs used in this study and the targeted C bases for each targeted locus. Sequences are oriented from 5' to 3':

	crRNA sequence used for base editing (5'–3')
<i>ctnnb1</i> (S33L)	CTGGACTCAGGAATACACTC
<i>tek</i> (Q94*)	GGAGCTCCAGGTGACGGTAG
<i>bap1</i> (Q273*)	GACTCAGCAAGAATCAGGCC
<i>tp53</i> (Q21*)	AGTATTCAGCCCCAGGTGG
<i>tp53</i> (Q170*)	CAATCAGCGAGCAAATTACA
<i>kras</i> (E62K)	CCTCCTGACCTGCAGTGTCC
<i>dmd</i> (Q8*)	CCACAGGACCAATGGGAGGA
<i>sod2</i> (Q145*)	GCTGTTTCAGGGCTCAGGCTG
<i>rb1</i> (W63*)	TCTCCATGCATGATCACAGA
<i>nras</i> NAA (G13S)	AACACCTCCTGCTCCCACAA
<i>tp53</i> NAA (Q170*)	ATCAGCGAGCAAATTACAGG
<i>cbl</i> (W577*)	AGTTCCAGTCTGGCATGTTG

### Micro-injection

Prior injections, a mix of 2  $\mu$ L of the Alt-R CRISPR-Cas9 crRNA (100 pmol/ $\mu$ L) and 2  $\mu$ L of Alt-R CRISPR-Cas9 tracrRNA (100 pmol/ $\mu$ L) from IDT was incubated at 95°C for 5 min, cooled down at room temperature, and then kept on ice to form the synthetic sgRNA complex. One nanoliter of another mix containing CBE mRNA (600 ng/ $\mu$ L) and the synthetic sgRNA complex (43 pmol/ $\mu$ L) was then injected into the cell at one-cell stage zebrafish embryos.

### Genotyping

To genotype the *cbl* mutant line, a PCR was performed with primers Fwd-5'-GTACGCCTGGA-GACCCATCTC-3' and Rev-5'-CTTTTGGACTGTCATAATCCGATGC-3'. The PCR product was digested with the restriction enzyme BsrI, which cut only on the WT allele. The WT allele resulted in two fragments (300 bp and 69 bp) and the mutant allele only one fragment (369 bp).

### Whole-embryo DNA sequencing

A series between 4 and 10 single embryos randomly chosen was analyzed for each target sequence, and the embryo with the highest efficiency is shown. Generally, between 25% and 100% embryos were positive for gene editing, that is showed >16% expected sequence modification. For genomic DNA extraction, each single embryo was digested for 1 hr at 55°C in 0.5 mL lysis buffer (10 mM Tris, pH 8.0, 10 mM NaCl, 10 mM EDTA, and 2% SDS) with proteinase K (0.17 mg/mL, Roche Diagnostics) and inactivated 10 min at 95°C. To sequence and check for frequency of mutations, each target genomic locus was PCR-amplified using Phusion High-Fidelity DNA polymerase (ThermoFisher Scientific, # F-530XL). PCR products have been extracted from an agarose gel and purified (using the PCR clean-up gel extraction kit #740609.50, Macherey-Nagel), and Sanger sequencing was performed by Eurofins. Sequence analyses were achieved using ApE software and quantifications of the mutation rate done using editR online software (*Kluesner et al., 2018*). For the verification of *cbl* mutant F1 embryos, *tek*, *bap1* and *tp53* mutations, PCR fragments were subsequently cloned into the pCR-bluntII-TOPO vector (Invitrogen). Plasmid DNA was isolated from single colonies and sent for sequencing. Mutant alleles were identified by comparison with the wild-type sequence using ApE and Geneious softwares.

Primer sequences used to amplify the targeted loci:

<b>tek (Q94*)</b>	<b>ATCTCAGACGTGACTCTGGTGAAC</b>	<b>TTCCTGTAGCATCTTGTAGGTGTAG</b>
<i>bap1</i> (Q273*)	TTGTTTATTTTTTCAGGACCATGGGG	CACCTGAAGGTATTGGTGTTCCTTG
<i>tp53</i> (Q21*)	CTTTGCATAAGAAACAACATCCCGA	GTTCAACACTGAAAACCAAAGAGG
<i>tp53</i> (Q170*)	ATATCTTGCTGTTTTCTCCCTGCT	GTCCTACAAAAAGGCTGTGACATAC
<i>kras</i> (E62K)	CGTTCCACTATGTCCACACATTTAG	AACAGTACATTTTCTGCATACTCGC
<i>dmd</i> (Q8*)	AGGGCTCCTTCCTTTTTCTGTTTAT	TGATCGAGTTTTGATGATTTCTCCG
<i>sod2</i> (Q145*)	GCATATGGCTGGAAATGATGAACC	GCACTTTATGTGCATTCACTGAGG
<i>rb1</i> (W63*)	TCTGTCAACTGTTGTTTTCCAGAC	TTCAATATCTGCCACACATACCTCA
<i>nras</i> (G13S)	CCTTTTCTCTTTTTGTCTGGGTG	CGCAATCTCACGTTAATTGTAGTGT
<i>cbl</i> (W577*)	GTACGCCTGGAGACCCATCTC	CTTTTGGACTGTCATAATCCGATGC

## Imaging

Embryos were oriented in low-melting agarose 0.6% with an anesthetic (Tricaine 0.013%) diluted in egg solution. The inverted laser scanning confocal microscope Zeiss CLSM-LSM780 was used for high-resolution microscopy, employing a 40× water immersion objective. Z-stacks were acquired every 1–2 μm. Leica MZ10F was used to image the whole embryos the *cbl* mutant adult fish. Image analyses were performed with ImageJ software.

## Body size quantifications

Eight control wild-type siblings and eight dwarf *MZcbl<sup>-/-</sup>* in total were used to measure the body size using a millimetric ruler. The length measured was from mouth to trunk. A non-parametric t-test with the Mann–Whitney correction was applied to determine significance in growth. The software used was Prism 7 (GraphPad).

## Acknowledgements

We thank Sophie Vríz for sharing the *Tg(7xTCF-Xla.Siam:GFP)* transgenic line and the members of the fish-facility in Institut Curie. We also thank Céline Revenu and Viviana Anelli for early contribution. MR was supported by the Fondation pour la Recherche Médicale (FRM grant number ECO20170637481) and la Ligue Nationale Contre le Cancer. Work in the Del Bene laboratory was supported by ANR-18-CE16 'iReelAx', UNADEV in partnership with ITMO NNP/AVIESAN (National Alliance for Life Sciences and Health) in the framework of research on vision and IHU FOReSIGHT [ANR-18-IAHU-0001] supported by French state funds managed by the Agence Nationale de la Recherche within the Investissements d'Avenir program. MCM was supported by World Wide Cancer Research, grant no. 0624, and by LILT –Trento, Program five per mille (year 2014).

## Additional information

### Funding

<b>Funder</b>	<b>Grant reference number</b>	<b>Author</b>
Agence Nationale de la Recherche	ANR-18-CE16 "iReelAx"	Filippo Del Bene
Agence Nationale de la Recherche	[ANR-18-IAHU-0001	Filippo Del Bene

Fondation pour la Recherche Médicale	ECO20170637481	Marion Rosello
Ligue Contre le Cancer		Marion Rosello
UNADEV/AVIESAN		Filippo Del Bene
Worldwide Cancer Research	grant no. 0624	Marina C Mione
LILT -Trento	Program 5 per mille	Marina C Mione

The funders had no role in study design, data collection and interpretation, or the decision to submit the work for publication.

#### Author contributions

Marion Rosello, Conceptualization, Formal analysis, Validation, Investigation, Visualization, Methodology, Writing - original draft, Writing - review and editing; Juliette Vouigny, Validation, Investigation, Methodology; François Czarny, Software; Marina C Mione, Conceptualization, Writing - review and editing; Jean-Paul Concordet, Conceptualization, Supervision, Methodology, Writing - original draft, Writing - review and editing; Shahad Albadri, Conceptualization, Formal analysis, Validation, Visualization, Writing - original draft, Writing - review and editing; Filippo Del Bene, Conceptualization, Supervision, Funding acquisition, Methodology, Writing - original draft, Writing - review and editing

#### Author ORCIDs

Marion Rosello  <https://orcid.org/0000-0003-3935-6971>  
Juliette Vouigny  <http://orcid.org/0000-0002-7361-8405>  
Marina C Mione  <http://orcid.org/0000-0002-9040-3705>  
Shahad Albadri  <https://orcid.org/0000-0002-3243-7018>  
Filippo Del Bene  <https://orcid.org/0000-0001-8551-2846>

#### Ethics

Animal experimentation: All procedures were performed on zebrafish embryos in accordance with the European Communities Council Directive (2010/63/EU) and French law (87/848) and approved by the Sorbonne Université ethic committee (Charles Darwin) and the French Ministry for research (APAFIS agreement #21323 2019062416186982) and by the Institut Curie ethic committee and the French Ministry for research (APAFIS agreement #6031 2016070822342309).

#### Decision letter and Author response

Decision letter <https://doi.org/10.7554/eLife.65552.sa1>  
Author response <https://doi.org/10.7554/eLife.65552.sa2>

## Additional files

#### Supplementary files

- Source code 1. SequenceParser.py STOP codon design source code. This python code highlights in capital the codons that can converted as STOP codon by C-to-T conversion with the chosen PAM sequence at the correct distance (PAM [−19, −13] bp window).
- Transparent reporting form

#### Data availability

All data generated or analysed during this study are included in the manuscript and supporting files.

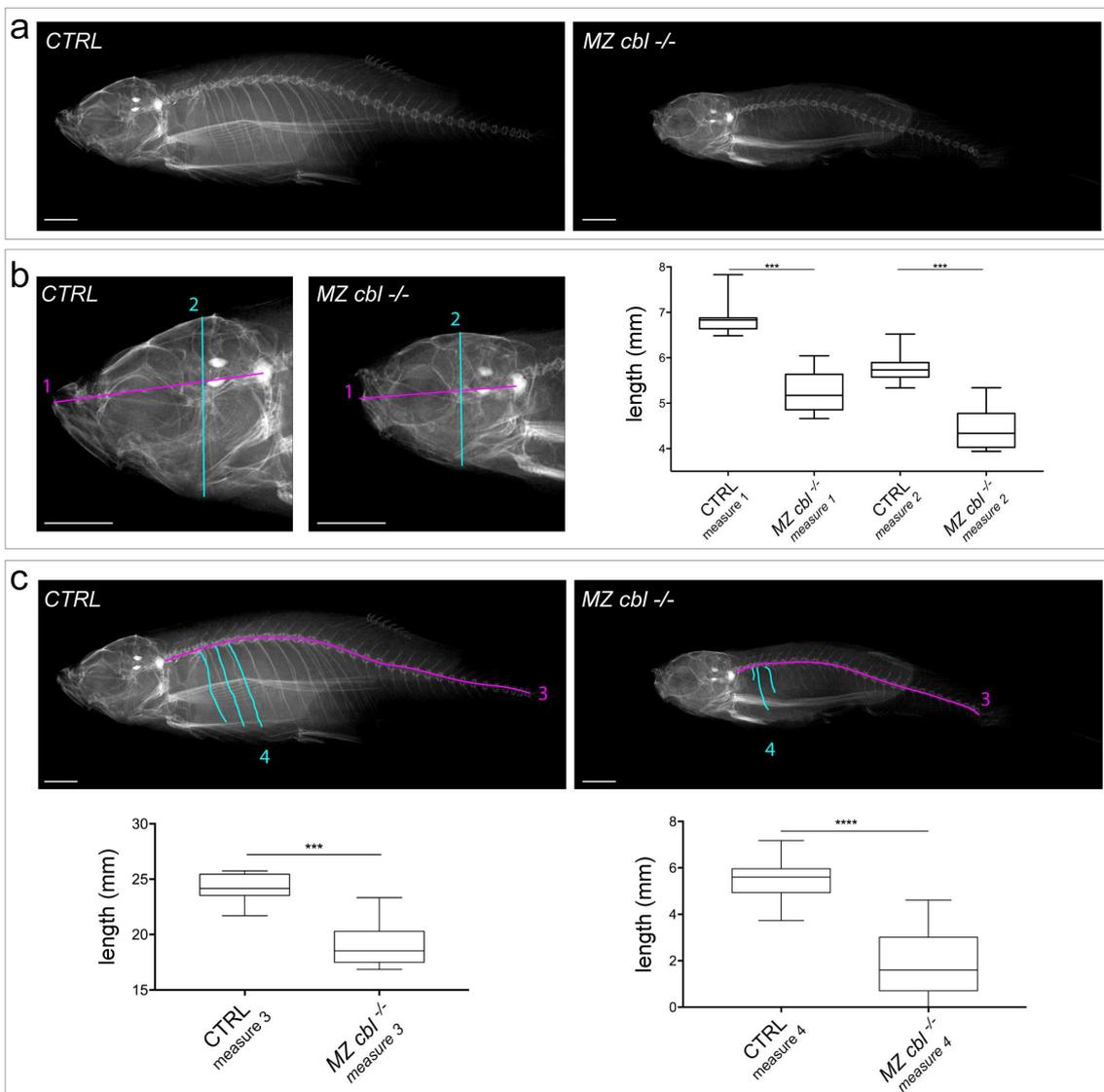
## References

- Ablain J**, Xu M, Rothschild H, Jordan RC, Mito JK, Daniels BH, Bell CF, Joseph NM, Wu H, Bastian BC, Zon LI, Yeh I. 2018. Human tumor genomics and zebrafish modeling identify *SPRED1* loss as a driver of mucosal melanoma. *Science* **362**:1055–1060. DOI: <https://doi.org/10.1126/science.aau6509>, PMID: 30385465
- Anzalone AV**, Randolph PB, Davis JR, Sousa AA, Koblan LW, Levy JM, Chen PJ, Wilson C, Newby GA, Raguram A, Liu DR. 2019. Search-and-replace genome editing without double-strand breaks or donor DNA. *Nature* **576**:149–157. DOI: <https://doi.org/10.1038/s41586-019-1711-4>, PMID: 31634902
- Bravard A**, Sabatier L, Hoffschir F, Ricoul M, Luccioni C, Dutrillaux B. 1992. SOD2: a new type of tumor-suppressor gene? *International Journal of Cancer* **51**:476–480. DOI: <https://doi.org/10.1002/ijc.2910510323>, PMID: 1592538
- Cagan RL**, Zon LI, White RM. 2019. Modeling Cancer with flies and fish. *Developmental Cell* **49**:317–324. DOI: <https://doi.org/10.1016/j.devcel.2019.04.013>, PMID: 31063751
- Callahan SJ**, Tepan S, Zhang YM, Lindsay H, Burger A, Campbell NR, Kim IS, Hollmann TJ, Studer L, Mosimann C, White RM. 2018. Cancer modeling by transgene electroporation in adult zebrafish (TEAZ). *Disease Models & Mechanisms* **11**:dmm034561. DOI: <https://doi.org/10.1242/dmm.034561>, PMID: 30061297
- Carrington B**, Weinstein RN, Sood R. 2020. BE4max and AncBE4max are efficient in germline conversion of C:G to T:A base pairs in zebrafish. *Cells* **9**:1690. DOI: <https://doi.org/10.3390/cells9071690>
- Cayuela ML**, Claes KBM, Ferreira MG, Henriques CM, van Eeden F, Varga M, Vierstraete J, Mione MC. 2018. The zebrafish as an emerging model to study DNA damage in aging, Cancer and other diseases. *Frontiers in Cell and Developmental Biology* **6**:178. DOI: <https://doi.org/10.3389/fcell.2018.00178>, PMID: 30687705
- Cho JY**, Guo C, Torello M, Lunstrum GP, Iwata T, Deng C, Horton WA. 2004. Defective lysosomal targeting of activated fibroblast growth factor receptor 3 in achondroplasia. *PNAS* **101**:609–614. DOI: <https://doi.org/10.1073/pnas.2237184100>, PMID: 14699054
- Dash S**, Kinney NA, Varghese RT, Garner HR, Feng WC, Anandkrishnan R. 2019. Differentiating between Cancer and normal tissue samples using multi-hit combinations of genetic mutations. *Scientific Reports* **9**:1005. DOI: <https://doi.org/10.1038/s41598-018-37835-6>, PMID: 30700767
- Dyson NJ**. 2016. RB1: a prototype tumor suppressor and an enigma. *Genes & Development* **30**:1492–1502. DOI: <https://doi.org/10.1101/gad.282145.116>, PMID: 27401552
- Gaudelli NM**, Komor AC, Rees HA, Packer MS, Badran AH, Bryson DI, Liu DR. 2017. Programmable base editing of A•T to G•C in genomic DNA without DNA cleavage. *Nature* **551**:464–471. DOI: <https://doi.org/10.1038/nature24644>, PMID: 29160308
- Haeussler M**, Schöning K, Eckert H, Eschstruth A, Mianné J, Renaud JB, Schneider-Maunoury S, Shkumatava A, Teboul L, Kent J, Joly JS, Concordet JP. 2016. Evaluation of off-target and on-target scoring algorithms and integration into the guide RNA selection tool CRISPOR. *Genome Biology* **17**:148. DOI: <https://doi.org/10.1186/s13059-016-1012-2>, PMID: 27380939
- Harada D**, Yamanaka Y, Ueda K, Tanaka H, Seino Y. 2009. FGFR3-related dwarfism and cell signaling. *Journal of Bone and Mineral Metabolism* **27**:9–15. DOI: <https://doi.org/10.1007/s00774-008-0009-7>, PMID: 19066716
- Hart M**, Concordet JP, Lassot I, Albert I, del los Santos R, Durand H, Perret C, Rubinfeld B, Margottin F, Benarous R, Polakis P. 1999. The F-box protein beta-TrCP associates with phosphorylated beta-catenin and regulates its activity in the cell. *Current Biology* **9**:207–211. DOI: [https://doi.org/10.1016/S0960-9822\(99\)80091-8](https://doi.org/10.1016/S0960-9822(99)80091-8), PMID: 10074433
- Hwang WY**, Fu Y, Reyon D, Maeder ML, Tsai SQ, Sander JD, Peterson RT, Yeh JR, Joung JK. 2013. Efficient genome editing in zebrafish using a CRISPR-Cas system. *Nature Biotechnology* **31**:227–229. DOI: <https://doi.org/10.1038/nbt.2501>, PMID: 23360964
- Jakimo N**, Chatterjee P, Nip L, Jacobson JM. 2018. A Cas9 with complete PAM recognition for Adenine dinucleotides. *bioRxiv*. DOI: <https://doi.org/10.1101/429654>
- Kluesner MG**, Nedveck DA, Lahr WS, Garbe JR, Abrahamte JE, Webber BR, Moriarity BS. 2018. EditR: a method to quantify base editing from Sanger sequencing. *The CRISPR Journal* **1**:239–250. DOI: <https://doi.org/10.1089/crispr.2018.0014>, PMID: 31021262
- Koblan LW**, Doman JL, Wilson C, Levy JM, Tay T, Newby GA, Maianti JP, Raguram A, Liu DR. 2018. Improving cytidine and Adenine base editors by expression optimization and ancestral reconstruction. *Nature Biotechnology* **36**:843–846. DOI: <https://doi.org/10.1038/nbt.4172>, PMID: 29813047
- Komor AC**, Kim YB, Packer MS, Zuris JA, Liu DR. 2016. Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage. *Nature* **533**:420–424. DOI: <https://doi.org/10.1038/nature17946>, PMID: 27096365
- Komor AC**, Zhao KT, Packer MS, Gaudelli NM, Waterbury AL, Koblan LW, Kim YB, Badran AH, Liu DR. 2017. Improved base excision repair inhibition and bacteriophage mu gam protein yields C:g-to-t:a base editors with higher efficiency and product purity. *Science Advances* **3**:eaao4774. DOI: <https://doi.org/10.1126/sciadv.aao4774>, PMID: 28875174
- Li S**, Balmain A, Counter CM. 2018. A model for RAS mutation patterns in cancers: finding the sweet spot. *Nature Reviews Cancer* **18**:767–777. DOI: <https://doi.org/10.1038/s41568-018-0076-6>, PMID: 30420765
- Liu C**, Kato Y, Zhang Z, Do VM, Yankner BA, He X. 1999. beta-Trcp couples beta-catenin phosphorylation-degradation and regulates *Xenopus* Axis formation. *PNAS* **96**:6273–6278. DOI: <https://doi.org/10.1073/pnas.96.11.6273>, PMID: 10339577
- Martinelli S**, De Luca A, Stellacci E, Rossi C, Checquolo S, Lepri F, Caputo V, Silvano M, Buscherini F, Consoli F, Ferrara G, Digilio MC, Cavaliere ML, van Hagen JM, Zampino G, van der Burgt I, Ferrero GB, Mazzanti L,

- Screpanti I, Yntema HG, et al. 2010. Heterozygous germline mutations in the CBL tumor-suppressor gene cause a noonan syndrome-like phenotype. *The American Journal of Human Genetics* **87**:250–257. DOI: <https://doi.org/10.1016/j.ajhg.2010.06.015>, PMID: 20619386
- Moro E, Ozhan-Kizil G, Mongera A, Beis D, Wierzbicki C, Young RM, Bournele D, Domenichini A, Valdivia LE, Lum L, Chen C, Amatruda JF, Tiso N, Weidinger G, Argenton F. 2012. In vivo wnt signaling tracing through a transgenic biosensor fish reveals novel activity domains. *Developmental Biology* **366**:327–340. DOI: <https://doi.org/10.1016/j.ydbio.2012.03.023>, PMID: 22546689
- Patton EE, Tobin DM. 2019. Spotlight on zebrafish: the next wave of translational research. *Disease Models & Mechanisms* **12**:dmm039370. DOI: <https://doi.org/10.1242/dmm.039370>, PMID: 30858282
- Peng X, Dong M, Ma L, Jia XE, Mao J, Jin C, Chen Y, Gao L, Liu X, Ma K, Wang L, Du T, Jin Y, Huang Q, Li K, Zou Li, Liu T, Deng M, Zhou Y, Xi X, et al. 2015. A point mutation of zebrafish c-cbl gene in the ring finger domain produces a phenotype mimicking human myeloproliferative disease. *Leukemia* **29**:2355–2365. DOI: <https://doi.org/10.1038/leu.2015.154>, PMID: 26104663
- Prykhodzhiy SV, Fuller C, Steele SL, Veinotte CJ, Razaghi B, Robitaille JM, McMaster CR, Shlien A, Malkin D, Berman JN. 2018. Optimized knock-in of point mutations in zebrafish using CRISPR/Cas9. *Nucleic Acids Research* **46**:e102. DOI: <https://doi.org/10.1093/nar/gky512>, PMID: 29905858
- Sander JD, Joung JK. 2014. CRISPR-Cas systems for editing, regulating and targeting genomes. *Nature Biotechnology* **32**:347–355. DOI: <https://doi.org/10.1038/nbt.2842>, PMID: 24584096
- Santoriello C, Zou LI. 2012. Hooked! modeling human disease in zebrafish. *Journal of Clinical Investigation* **122**:2337–2343. DOI: <https://doi.org/10.1172/JCI60434>
- Steinhart Z, Angers S. 2018. Wnt signaling in development and tissue homeostasis. *Development* **145**:dev146589. DOI: <https://doi.org/10.1242/dev.146589>, PMID: 29884654
- Tessadori F, Roessler HI, Savelberg SMC, Chocron S, Kamel SM, Duran KJ, van Haelst MM, van Haften G, Bakkers J. 2018. Effective CRISPR/Cas9-based nucleotide editing in zebrafish to model human genetic cardiovascular disorders. *Disease Models & Mechanisms* **11**:dmm035469. DOI: <https://doi.org/10.1242/dmm.035469>, PMID: 30355756
- Wiersen WA, Welker JM, Almeida MP, Mann CM, Webster DA, Torrie ME, Weiss TJ, Kambakam S, Vollbrecht MK, Lan M, McKeighan KC, Levey J, Ming Z, Wehmeier A, Mikelson CS, Haltom JA, Kwan KM, Chien CB, Balciunas D, Ekker SC, et al. 2020. Efficient targeted integration directed by short homology in zebrafish and mammalian cells. *eLife* **9**:e53968. DOI: <https://doi.org/10.7554/eLife.53968>, PMID: 32412410
- Zhang Y, Qin W, Lu X, Xu J, Huang H, Bai H, Li S, Lin S. 2017. Programmable base editing of zebrafish genome using a modified CRISPR-Cas9 system. *Nature Communications* **8**:118. DOI: <https://doi.org/10.1038/s41467-017-00175-6>, PMID: 28740134
- Zhao Y, Shang D, Ying R, Cheng H, Zhou R. 2020. An optimized base editor with efficient C-to-T base editing in zebrafish. *BMC Biology* **18**:190. DOI: <https://doi.org/10.1186/s12915-020-00923-z>, PMID: 33272268

### II.1.3. SKELETAL PHENOTYPIC ANALYSIS OF THE DWARF $MZCBL^{-/-}$ MUTANTS.

In collaboration with Dr. Dambroise (*Imagine Institute, Paris*), phase-contrast X-ray imaging was performed on 8 dwarf  $MZcbl^{-/-}$  mutants and 8 control fish to analyze skeletal malformations. We can observe a shortened length and height of the head compared to the control fish heads, suggesting a microcephalia phenotype in the mutant fish (Fig. 8.b, measures 3 and 4). Moreover, although the number of vertebrae were the same in the control and the  $cbl$  mutant fish, the spine and the ribs were shorter in the dwarf fish than in the control fish (Fig. 8.c, measures 3 and 4).



**Figure 8: 3 mpf dwarf  $MZcbl^{-/-}$  zebrafish display skeletal malformations. (a)** Phase-contrast X-ray imaging on 3 mpf control and dwarf  $MZcbl^{-/-}$  fish. Scale bar=2 mm. **(b)** Quantifications of the length (measure 1, n=8 per condition) and the height (measure 2, n=8 per condition) of the head. **(c)** Quantifications of the spine length (measure 3, n=8 per condition) and the ribs length (measure 4, n=80 ribs per condition) (10

ribs per fish)). **(c, d)** Wilcoxon-Mann-Whitney test was used for the measures 1, 2 and 3 and T-test was performed for the measure 4 were performed. \*\*\*:  $p < 0.001$ , \*\*\*\*:  $p < 0.0001$ .

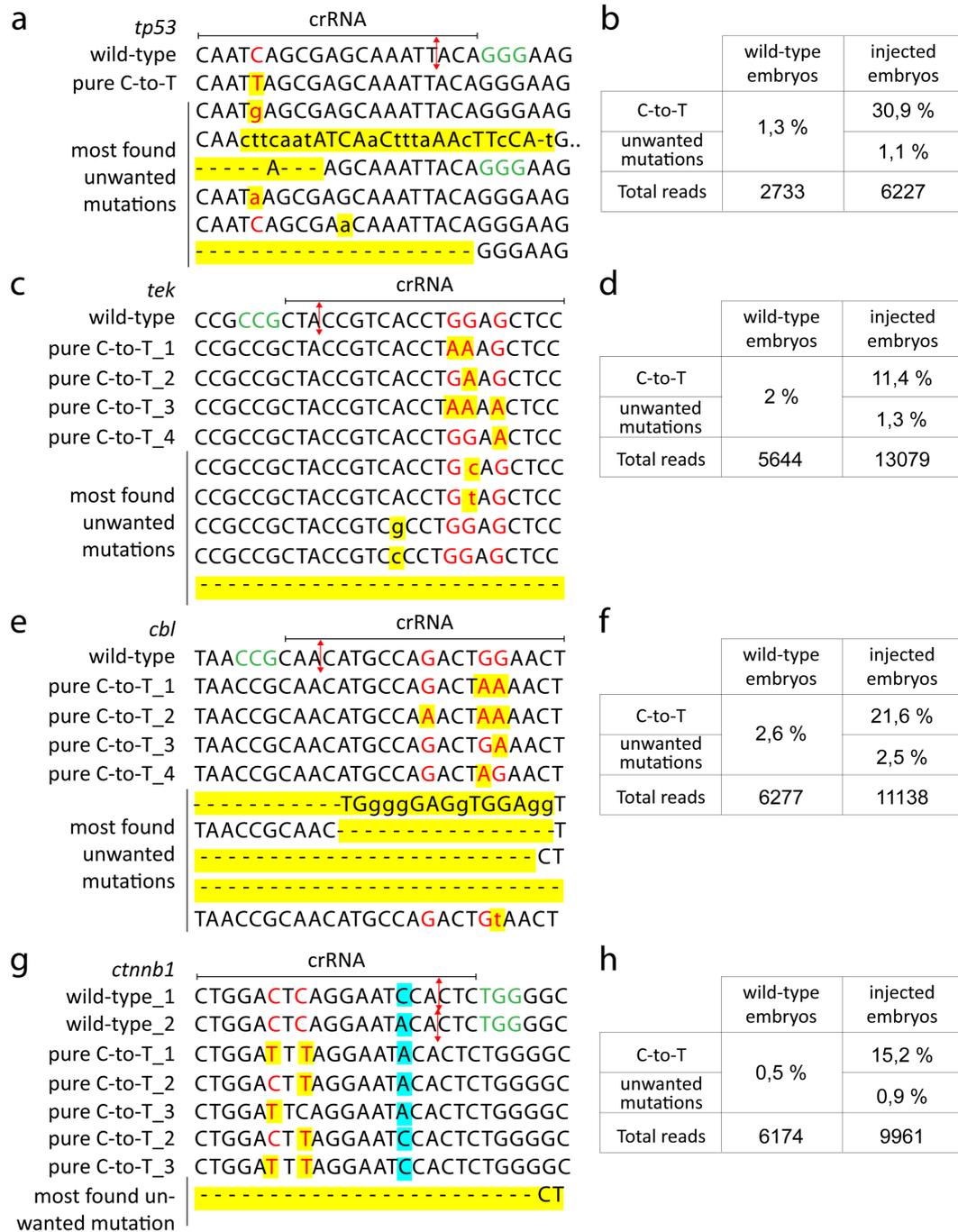
#### II.1.4. IN-DEPTH ANALYSIS OF THE BE4-GAM SPECIFICITY IN ZEBRAFISH.

Using the BE4-gam variant we were able to edit at high efficiency several genes in zebrafish. These base conversions were precise as unwanted mutations were not detectable after single clones sequencing. Indeed, for *tek* 6 out of 20, *bap1* 2 out of 12, *tp53 (Q21\*)* 12 out of 24, and *tp53 (Q170\*)* 21 out of 24 colonies showed precise C-to-T transitions, whereas all the others were wild-type and no unwanted mutations were found<sup>142</sup>. However, although the reads were enough to evaluate the efficiency and the precision of the technique, the number of reads was poor using this approach.

In order to go deeper in the analysis of the by-products obtained using the BE4-gam variant, we decided to perform a next-generation sequencing (NGS) analysis to sequence a high volume of fragments for some of our targets. In collaboration with Dr. Concordet and Dr. Giovannangeli (*MNHN, Paris*), 4 different loci that were targeted using the BE4-gam in zebrafish and analyzed by NGS: *tp53 (Q170\*)*, *tek*, *cbl* and *ctnnb1*. For each gene, the NGS was done on the pool of all of the injected embryos reported in the article 1 in table 2<sup>187</sup>, including those for which mutations were not detectable by Sanger sequencing (n=8 injected embryos for *ctnnb1*, *tek*, *tp53* and 10 injected embryos for *cbl*). As control, non-injected embryos were sequenced (n=8). After the analysis of at least 2733 reads for the wild-type embryos and 6227 reads for the injected embryos, we observed precise C:G to G:A conversions (Fig. 9). The INDELS frequencies obtained were very close to those found in the control samples (Fig. 9). Moreover, as previously observed, these NGS results confirmed in a precise manner that the presence of several C bases in the editing window can result in the generation of different edited by-products, as shown for *tek*, *cbl* and *ctnnb1* (Fig. 9.c, e, g).

Finally, for *ctnnb1* gene a polymorphism site was present in the injected fish from our fish facility (Fig. 9.g, in blue). The sgRNA used was designed to recognize the alleles containing an A base at the target site (Fig. 9.g, wild-type 2) representing less than 20% of the total wild-type alleles in the pool of injected embryos analyzed, and 95% of the base edited sequences contained the A base. This final observation suggests that one single mismatch in the sgRNA sequence can drastically decrease the base editing

efficiency, highlighting the robustness of this technique with respect to DNA off-targets. All together, these results support the specificity of BE4-gam in zebrafish.



**Figure 9: BE4-gam high fidelity of base editing in zebrafish. (a, h)** NGS analysis from pools of injected embryos from our previous study<sup>141</sup> where C-to-T conversions were induced in *tp53* (a, b), *tek* (c, d), *cbl* (e, f) and *ctnnb1* (g, h) genes using the BE4-gam base editor. (b, d, f, h) Percentages of the wild-type sequences and sequences containing the pure C-to-T conversions or INDELS found with NGS sequencing. (a, c, e, g) Most sequences obtained are represented with edits highlighted in yellow. In red are the targeted bases, in green is the PAM and in blue is the polymorphism in *ctnnb1* target site (g). The red arrows are the nick sites.

## II.2. ESTABLISHMENT OF THE NEAR-PAMLESS BASE EDITOR STRATEGY FOR MULTIPLEX GENE EDITING AND BASE EDITING CO-SELECTION APPROACH.

### II.2.1. SUMMARY ARTICLE 2.

So far base editors have been used in different model systems including animal organisms with a high degree of efficiency but their applications are still limited by the need of an appropriate PAM present at the correct distance of the C base that will potentially be modified. This problem can be particularly complex for multiplex gene editing. In the article 2, we aimed to overcome this limitation by developing the near-PAMless base editor approach for the first time in animals, called the CBE4max-SpRY<sup>214</sup>. We were able to induce precise C-to-T base conversions at a remarkable efficiency in zebrafish embryos using non-5'-NGG-2' PAMs. We could show:

- C-to-T conversions up to 100% of efficiency for different loci using non-NGG PAMs and a high rate of germline transmissions. We also report for some targets a base editing efficiency higher than the one obtained using the AncBE4max variant for the same sgRNA upstream of a 5'-NGG-3' PAM.
- A base conversion for the C located 12 bases away from the PAM.
- The possibility to test several sgRNAs and modulate the localization within the editing window of the targeted C base to increase the C-to-T conversion efficacy and to exclude other Cs present in the target site, avoiding the generation of other bystander mutations.
- Multiplex gene editing by targeting different loci using 5'NAN-3' and 5'-NGN-3' PAMs simultaneously.
- A simple and efficient visual co-selection method allowing to identify the most edited embryos following injections of *CBE4max-SpRY* mRNA, sgRNA inducing a LOF mutation in the *tyrosinase* gene and the sgRNAs targeting the genes of interest. This strategy represents the first example of an *in vivo* selection method in zebrafish with CRISPR/Cas9 based techniques that is here applied for base editing in an animal model.

- Thanks to the flexibility of our tool, we have created a multigenic cancer predisposition model introducing an activating mutation in *nras* oncogene and a stop mutation in *tp53* tumor suppressor gene. These results pave the way to model genetic diseases such as cancers by introducing multiple genetic mutations without performing oncogenic overexpression by transgenesis. This can complement the past and current models that could rely so far mainly on the overexpression of human mutated oncogene to generate cancer models.

## II.2.2. ARTICLE 2: DISEASE MODELING BY EFFICIENT GENOME EDITING USING A NEAR-PAMLESS BASE EDITOR *IN VIVO*.

Rosello M, *et al.*. Disease modeling by efficient genome editing using a near PAM-less base editor *in vivo*. *bioRxiv*, 2021.2006.2028.450169 (2021).

1 **Disease modeling by efficient genome editing using a near**  
2 **PAM-less base editor *in vivo*.**

3

4 Marion Rosello<sup>1</sup>, Malo Serafini<sup>1</sup>, Marina C Mione<sup>2</sup>, Jean-Paul Concordet<sup>\*3</sup> and Filippo  
5 Del Bene<sup>\*1</sup>

6

7 **Affiliations**

8 <sup>1</sup>Sorbonne Université, INSERM U968, CNRS UMR 7210, Institut de la Vision, Paris,  
9 France.

10 <sup>2</sup>Department of Cellular, Computational and Integrative Biology – CIBIO, University of  
11 Trento, Trento, Italy.

12 <sup>3</sup>Muséum National d'Histoire Naturelle, INSERM U1154, CNRS UMR 7196, Paris,  
13 France.

14

15 **Corresponding authors:**

16 Jean-Paul Concordet\*: [jean-paul.concordet@mnhn.fr](mailto:jean-paul.concordet@mnhn.fr)

17

18 Filippo Del Bene\*: [filippo.del-bene@inserm.fr](mailto:filippo.del-bene@inserm.fr)

19

20

21 **Abstract**

22

23 **Base Editors are emerging as an innovative technology to introduce point**  
24 **mutations in complex genomes. So far, the requirement of an NGG Protospacer**  
25 **Adjacent Motif (PAM) at a suitable position often limits the editing possibility to**  
26 **model human pathological mutations in animals. Here we show that, using the**  
27 **CBE4max-SpRY variant recognizing the NRN PAM sequence, we could**  
28 **introduce point mutations for the first time in an animal model and achieved up**  
29 **to 100% efficiency, thus drastically increasing the base editing possibilities.**  
30 **With this near PAM-less base editor we could simultaneously mutate several**  
31 **genes and developed a co-selection method to identify the most edited embryos**  
32 **based on a simple visual screening. Finally, we applied our method to create a**  
33 **new zebrafish model for melanoma predisposition based on the simultaneous**  
34 **editing of multiple genes. Altogether, our results considerably expand the Base**  
35 **Editor application to introduce human disease-causing mutations in zebrafish.**

## 36 **Introduction**

37 The CRISPR (Clustered Regularly Interspaced Short Palindromic Repeats)/Cas9  
38 system is a very powerful gene editing tool to perform mutagenesis in zebrafish<sup>1</sup>. The  
39 sgRNA-Cas9 complex first binds to target DNA through a Protospacer Adjacent Motif  
40 (PAM) and Cas9 next cleaves DNA upon stable annealing of the sgRNA spacer  
41 sequence to the sequence immediately upstream of the PAM. The NGG PAM  
42 sequence is thus required for the *Sp*Cas9 protein to introduce a DNA double-strand  
43 break (DSB). This technique is now extensively used in zebrafish to produce knock-  
44 out alleles<sup>2</sup>. Additionally, recent studies showed that exogenous DNA and Single  
45 Nucleotide Polymorphisms (SNPs) can be introduced in the genome using  
46 CRISPR/Cas9 mediated homology-directed (HDR) repair with variable efficiency<sup>3, 4, 5</sup>.  
47 In order to complement these HDR-based strategies to introduce specific point  
48 mutation, second-generation gene editing tools called base editors (BEs) have recently  
49 been developed. The Cytidine Base Editor (CBE) is composed of a Cas9-D10A  
50 nickase fused to a cytidine deaminase and converts C-to-T bases in a restricted  
51 window of 13 to 19 nucleotides (nt) upstream of the PAM sequence without generating  
52 DSBs<sup>6, 7, 8</sup>. In zebrafish, several CBE variants have been shown to work with different  
53 gene editing efficiencies<sup>9, 10, 11, 12</sup>. In a previous study we tested several CBE variants  
54 and we were able to reach a C-to-T conversion efficacy of 91% in zebrafish, showing  
55 many applications from signaling pathway activation to human disease modeling<sup>12</sup>. It  
56 has been shown that potentially all C bases within the PAM [-19, -13] bp window can  
57 be edited with these tools, although a higher efficiency was generally achieved for the  
58 Cs located in the middle of this editing window, highlighting the importance of the C  
59 distance to the PAM for efficient base editing<sup>13</sup>. Therefore, the restriction of the base  
60 modification to the PAM-dependent window is still a critical intrinsic limitation to the  
61 base editor techniques. Due to this restriction, these tools cannot be applied to any  
62 gene and any mutation of interest. To overcome this limitation, extensive works have  
63 been done in cultured cells to engineer CBEs recognizing other PAM sequences than  
64 the classical NGG and thereby significantly expand the range of C bases that can be  
65 converted. Based on the technological advances made in cultured cells, we tested  
66 several novel CBEs in zebrafish to overcome the PAM limitation. Among them, the  
67 most recent and flexible variant is the CBE4max-SpRY, reported as a near PAM-less  
68 base editor recognizing almost all PAM sequences in cultured cells<sup>14</sup>. Here we  
69 established the CBE4max-SpRY CBE variant for the first time in an animal model.

70 Using this variant, we were able to perform in zebrafish C-to-T conversions at an  
71 unprecedented high efficiency, reaching up to 100%. We show that the CBE4max-  
72 SpRY converts C bases efficiently using NG or NA PAM in zebrafish and that it is  
73 possible to mutate several genes using NG and NA PAMs at the same time, increasing  
74 drastically the base editing possibilities. Based on these results, we developed a co-  
75 selection method to phenotypically identify the most edited embryos following the  
76 CBE4max-SpRY, by co-targeting for the *tyrosinase* gene and selecting embryos for  
77 lack of pigmentation. Finally, using this approach, we could for the first time in zebrafish  
78 simultaneously introduce a loss-of-function mutation in a tumor suppressor gene  
79 together with a gain-of-function mutation in an endogenous oncogene. We targeted  
80 *tp53* tumor suppressor and *nras* oncogene and generated a new zebrafish model with  
81 an abnormal melanocyte growth which is an hallmark of melanoma formation  
82 susceptibility<sup>15</sup>, without over-expressing mutated oncogenes which has been the main  
83 strategy used in zebrafish to model cancer so far<sup>16, 17, 18, 19</sup>.

## 84 **Results**

### 85 **Evaluation of several CBE variants recognizing non-NGG PAM sequences in** 86 **zebrafish.**

87 Base editing requires the presence of a PAM at 13 to 19 bp downstream of the targeted  
88 C base. This limitation is critical and often makes CBEs unable to introduce the desired  
89 point mutation in the genome. To overcome this limitation particularly important when  
90 trying to model disease causing mutations in animal models, we first tested three  
91 different CBE variants recognizing the NG PAM: xCas9-BE4<sup>20</sup>, CBE4max-SpG<sup>14</sup> and  
92 CBE4max-SpRY<sup>14</sup>. In order to analyze their efficacy in zebrafish, we started by  
93 injecting into one-cell stage embryos sgRNAs acting upstream of an NGG PAM that  
94 we previously found to be very efficient with the original CBEs<sup>12</sup>. First we tested *tp53*  
95 *Q170\** sgRNA, with which we got up to 86% of efficiency using BE4-gam<sup>12</sup>. After  
96 injection with the *xCas9-BE4* mRNA, we were unable to detect any base conversion  
97 by Sanger sequencing of PCR products of the target region. We then tested CBE4max-  
98 SpG and CBE4max-SpRY with *rb1W63\*NGG* sgRNA targeting the *rb1* tumor  
99 suppressor gene upstream of an NGG PAM (Fig. 1a) to introduce the W63\* mutation  
100 with which we previously got up to 91% of base conversion using the ancBE4max  
101 variant<sup>12</sup>. We sequenced the target region from 14 single embryos and we did not  
102 detect any C-to-T conversion using CBE4max-SpG whereas we got C-to-T  
103 conversions for 4 out of 11 single embryos analyzed with up to 33% of efficiency using  
104 CBE4max-SpRY (Fig. 1b). The CBE4max-SpRY has been reported as a near PAM-  
105 less base editor in cultured cells<sup>14</sup> and we next chose to analyze its flexibility. We  
106 designed two sgRNAs, one upstream of an NG PAM (*rb1W63\*NG* sgRNA) and the  
107 second upstream of an NA PAM (*rb1W63\*NA* sgRNA), in order to introduce the same  
108 mutation in the *retinoblastoma1* (*rb1*) gene (Fig. 1a). After injection into one-cell stage  
109 of each sgRNAs with the *CBE4max-SpRY* mRNA, we sequenced two pools of 9  
110 injected embryos. With both sgRNAs we obtained high base editing efficiency rates,  
111 up to 58% of efficiency with the NG PAM and up to 48% of efficiency with the NA PAM  
112 (Fig. 1c). Additionally, we noticed that another C was also edited, leading to the R64K  
113 mutation but it was 3' to the premature STOP codon introduced (C16 for the NG sgRNA  
114 and C17 for the NA sgRNA). In our previous study using the ancBE4max CBE, this C  
115 was at 19 bp from the NGG PAM and we obtained less than 22% of base editing<sup>12</sup>  
116 whereas here we show that this C can be converted up to 55% of efficiency (Fig. 1c).

117 This last result shows that for some C base targets, thanks to the PAM flexibility of the  
118 CBE4-SpRY, we can now increase the C-to-T conversion efficacy compared to the use  
119 of the classical ancBE4max. We demonstrated here for the first time in animals that  
120 the near PAM-less CBE4max-SpRY converts a C base into a T base efficiently using  
121 NA or NG PAMs.

122

### 123 **Total base conversion in F0 embryos using an NA PAM.**

124 In order to explore the efficiency of the near PAM-less CBE variant in zebrafish, we  
125 decided to target another locus, the *tyrosinase* gene. We designed 3 different guides  
126 upstream, respectively, of an NGG, NA and NG PAM in order to introduce the W273\*  
127 mutation in Tyrosinase, an enzyme necessary for the production of melanin pigments  
128 (Fig. 2a). Upon independent injections of these sgRNAs with the *ancBE4max* mRNA  
129 for the NGG sgRNA and the *CBE4max-SpRY* mRNA for the NGG, NA and NG  
130 sgRNAs, the embryos showed a range of pigmentation defects. We divided these  
131 phenotypes in four groups depending on the severity of the pigmentation defects: wild-  
132 type like, mildly depigmented, severely depigmented and albino. Representative  
133 pictures of 2 days post-fertilization (2 dpf) embryos for each group are illustrated in Fig.  
134 2c. Upon the injection of the *tyrW273\* NGG* sgRNA and the *ancBE4max* mRNA, we  
135 obtained 50% of mildly depigmented embryos with a small proportion of severely  
136 affected embryos whereas surprisingly almost 100% of the injected fish with the  
137 *CBE4max-SpRY* mRNA were depigmented (Fig. 2b, columns 2 and 3). Next, using the  
138 *tyrW273\* NA* sgRNA and the *CBE4max-SpRY*, we obtained 100% of injected embryos  
139 showing pigmentation defects with almost 50% exhibiting a total lack of pigmentation  
140 (Fig. 2b, column 4). Finally, with the NG PAM sgRNA, 50% of the injected embryos  
141 were poorly affected, the base editing efficiency reached only 20% (Fig. 2b columns 5,  
142 d). Remarkably in the pool of 9 albino embryos from the injection of the *tyrW273\* NA*  
143 sgRNA and *CBE4max-SpRY* mRNA, we obtained 100% of C-to-T conversion for the  
144 C16 and 95% for the C15 (Fig. 2d, e), an efficiency rate never reached so far even with  
145 the use of the classical CBEs recognizing the NGG PAM. We did not obtain this C-to-  
146 T efficiency neither using the *tyrW273\* NG* sgRNA nor the *ancBE4max* with the  
147 *tyrW273\* NGG* sgRNA (Fig. 2b, d). We could speculate that the difference of base  
148 conversion efficiency is due to the distance of the C from the PAM or to the sgRNA  
149 sequence for which the shift of one base would drastically impact the gene editing  
150 efficiency (Fig. 2a). Moreover, we found that the mutation was transmitted to 100% of

151 the offspring by screening only 4 F0 adult fish presenting pigmentation defect at the  
152 adult stage (Supplementary Fig. 1, Founder 1). Other 3 F0 fish also transmitted the  
153 mutation (Supplementary Fig. 1). The screening was performed by Sanger sequencing  
154 of PCR products of the *tyrosinase* region in random single F1 embryos from an  
155 outcross of each founder. Moreover, by crossing 2 founders, Founder 1 and 2 in  
156 supplementary Fig. 1, we were able to obtain 51,9% of albino embryos (n=28/54).  
157 Together these results highlight the CBE4max-SpRY as a very powerful CBE variant,  
158 showing that for many targets this CBE could be a better choice than the classical  
159 CBEs and could edit 100% of the alleles of the injected embryo using NA or NG PAMs.

160

### 161 **Simultaneous editing of two different bases.**

162 To test if we can take advantage to the high flexibility and efficacy of the CBE4max-  
163 SpRY variant in zebrafish for multiplex editing, we targeted two loci at the same time  
164 using a sgRNA upstream of an NG PAM and one upstream of an NA PAM. We injected  
165 the *CBE4max-SpRY* mRNA and the two synthetic *tyrW273\* NA* and *rb1W63\* NG*  
166 sgRNAs into the cell at one-cell stage. Among the injected embryos, 100% showed  
167 pigmentation defects and at least 50% exhibited total absence of pigmentation (Fig.  
168 3a). It can be noted that this proportion is almost the same as the one after the use of  
169 the *tyrW273\* NA* sgRNA only and the *CBE4max-SpRY* mRNA (Fig. 2b, column 4),  
170 meaning that the addition of a second guide did not affect the base conversion efficacy  
171 at the *tyrosinase* locus. Sequencing of the two loci was performed on three different  
172 pools of embryos separated according to the severity of their pigmentation defects. As  
173 expected, the severity of the pigmentation phenotype follows the base conversion  
174 efficiency of the *tyrosinase* gene (Fig. 3a, b). In particular, we found that in the pool of  
175 35 albino embryos, we could no longer detect the wild type C by Sanger sequencing,  
176 suggesting almost 100% of C-to-T conversion in the *tyrosinase* gene. Remarkably, in  
177 the albino pool, up to 44% of C-to-T conversion of the *rb1* gene was detected, revealing  
178 that double *tyrosinase* and *rb1* mutations were present in a large proportion of cells  
179 (Fig. 3a, b), while in mildly depigmented zebrafish, editing of *rb1* was up to 16%. In  
180 addition, we found that both mutations were transmitted to the offspring with high  
181 transmission rates, as shown by screening the progeny of only one F0 adult fish (Fig.  
182 3c, d). We obtained different combinations of mutations, and we showed that some F1  
183 embryos were mutated for the *tyrosinase* gene alone. At the same time, we noticed  
184 that some embryos were mutated for different Cs of the editing windows for each gene

185 (Fig. 3d). With these results we could demonstrate that, using this approach, it is now  
186 possible to mutate simultaneously two different genes with high efficiency by  
187 combining two different PAM sequences in zebrafish, NA and NG PAMs. Additionally,  
188 we can observe that the highest efficiency for *rb1* mutation was found in the albino  
189 embryo pool, embryos for which the *tyr* mutation is generated in 100% of the alleles  
190 as measured by Sanger sequencing.

191

### 192 **Co-selection strategy to prescreen phenotypically the most edited embryos.**

193 The efficiency of CBE4max-SpRY achieved here above opens the possibility to  
194 perform multiplex mutagenesis in zebrafish. Such experiments, however, usually  
195 involve time consuming screening to obtain a founder carrying all the desired mutations  
196 or long crossing protocols to genetically combine multiple mutations identified in  
197 different animals. For this reason, we have decided to take advantage of the high base  
198 editing rate of the *tyrW273\*NA* sgRNA and developed a method for co-selection of  
199 base editing to rapidly identify the most edited F0 embryos following *CBE4max-SpRY*  
200 mRNA injections. We first injected the *rb1W63\*NG* and *nras NA* sgRNAs with the near  
201 PAM-less *CBE4max-SpRY* mRNA and found C-to-T conversion rates up to 3% at *rb1*  
202 and 50% at *nras* targets in a pool of 50 injected embryos (Fig. 4a). Moreover, we can  
203 observe that the C12 has been edited at 23% of efficiency and the C18 has not been  
204 edited on the contrary of what we usually observe using the other CBE4 variants (Fig.  
205 4a). From these observations, we can speculate that the editing windows of the  
206 CBE4max-SpRY is slightly different than the usual [-19, -13 bp] PAM window and that  
207 with the use of this CBE it is possible to target the C12 bases. We then added to the  
208 same micro-injection mix the *tyrW273\*NA* sgRNA and obtained embryos showing  
209 pigmentation defects that we split in four groups as above (Fig. 2c and Fig. 4b). By  
210 sequencing the 3 loci in each pool of embryos, we could show that the editing efficiency  
211 is higher or lower to the same extent in the 3 targeted loci (Fig. 4c). The mildly  
212 depigmented embryos were edited up to 41% for *tyr*, 61% for *nras* and no detectable  
213 mutation for *rb1* whereas the albino embryos were edited up to 95% for *tyr*, 100% for  
214 *nras* and 22% for *rb1* (Fig. 4c). Base editing co-selection strategies were recently  
215 demonstrated in cultured cells<sup>21</sup> and have not been reported in animals so far. Using  
216 the *tyrW273\* NA* sgRNA, we show here a powerful strategy to readily obtain the most  
217 efficiently C base edited embryos at targeted loci of interest by simply selecting for  
218 albino embryos resulting from co-targeting the *tyrosinase* gene.

219

220 **Genetic disease modeling by generating combination of mutations.**

221 We next investigated the ability to use CBE4max-SpRY to introduce in zebrafish  
222 mutations found in human cancers. The zebrafish has become a powerful *in vivo* model  
223 to study a high variety of human cancer types<sup>16, 17, 18, 19</sup>. However, studies to assess  
224 the activity of mutations in oncogenes in this animal model have relied so far on  
225 transgenic approaches that express human oncogenes using tissue specific or  
226 constitutive promoters. In order to more accurately mimic the impact of cancer  
227 mutations as observed in somatic alterations in humans, we decided to use the  
228 CBE4max-SpRY to directly generate combination of mutations in endogenous  
229 zebrafish genes preserving their normal transcriptional regulation. To test this  
230 important innovative approach, we aimed at generating cancer models with *Nras*  
231 oncogene activation combined with loss-of-function mutations in the *tp53* tumor  
232 suppressor gene. We thus injected into one-cell stage embryos the *CBE4max-SpRY*  
233 mRNA, *nras* NAA and *tp53* Q170\* sgRNAs. Upon this injection we found 50% of the  
234 injected fish with an over-all increase of pigmentation at 3 dpf (n=36/70), a phenotype  
235 never seen upon separate injection of each sgRNA alone (Fig. 5a). To verify that the  
236 absence of phenotype when injecting only *nras* NAA sgRNA is not due to a low  
237 efficiency of gene editing, we sequenced a pool of 80 injected embryos and we  
238 detected high efficiency of base conversion (Fig. b). We next randomly selected 4  
239 embryos presenting a pigmentation similar to control embryos and 8 embryos showing  
240 an increase of pigmentation after the injection of the sgRNAs targeting *nras* and *tp53*  
241 genes. After sequencing both loci, we found indeed a correlation between the pigment  
242 phenotype and the multi-mutagenesis efficiency, embryos with an increase of  
243 pigmentation were more edited for *nras* and *tp53* genes (Fig. 5c, d). We thus have  
244 shown for the first time that endogenous activation of *nras* oncogene and knock-out of  
245 *tp53* tumor suppressor gene leads to an increase of melanocyte numbers in zebrafish,  
246 an early evidence of abnormal melanocyte growth which could led to melanoma  
247 formation (Fig. 5). Indeed, it has been reported that fish over-expressing human  
248 mutated HRAS oncogene in melanocytes were hyperpigmented at 3 dpf and  
249 developed melanoma at the adult stage<sup>15</sup>. Moreover, other reports using zebrafish  
250 transgenic lines have suggested a role of p53 and Ras oncogenes in melanoma  
251 formation<sup>22, 23</sup>. We have developed here for the first time a hyper-pigmentation

252 zebrafish model by generating endogenous activating mutation in *nras* oncogene and  
253 loss-of-function mutation in *tp53* tumor suppressor gene.

## 254 Discussion

255 While the Base Editor technology is emerging as a revolutionary method to introduce  
256 precise single mutation in the genome, the presence at the good localization of the  
257 NGG PAM sequence is necessary and often a constrain. This has restricted its  
258 potential applications as the absence of the PAM puts the CBE unusable for many  
259 targets of interest. Here, we addressed these limitations in zebrafish by testing several  
260 base editor variants recognizing other PAM sequences. We unfortunately did not  
261 obtain any C-to-T conversions in zebrafish embryos using the previously published  
262 xCas9-BE4<sup>20</sup> and CBE4max-SpG<sup>14</sup>. Nevertheless we could introduce point mutations  
263 with a remarkable high efficiency rate using the CBE4max-SpRY variant, a recently  
264 described near PAM-less CBE variant engineered and validated in cultured cells but  
265 never reported working in an animal model so far<sup>14</sup>. We thus significantly expand the  
266 base conversion possibilities in zebrafish and open the possibility to convert non NGG-  
267 targetable C bases which could not be achieved so far. Through our results, we could  
268 demonstrate that the CBE4max-SpRY can be extremely efficient, reaching 100% of C-  
269 to-T conversion in at least 50% of the injected embryos (Fig. 2d, e, 3b and 4c). We  
270 screened only 4 F0 fish and the 4 transmitted the mutation to the offspring with a high  
271 germline transmission rate (Supplementary Fig. 1). These last results are particularly  
272 remarkable as this efficiency rate has never been reached previously in zebrafish, even  
273 with the use of the classical CBEs recognizing the NGG PAM<sup>9, 10, 11, 12</sup>. Moreover, for  
274 one of our targets the CBE4max-SpRY was even more efficient than the ancBE4max  
275 (Fig. 2b-e). This could be due also to the fact that the CBE4max-SpRY might have a  
276 slightly different editing window that the usual PAM [-19, -13 bp] window as we show  
277 base editing for the C12 and no conversion for the C18 in *nras* targeting (Fig. 4). These  
278 last results and the PAM flexibility of the CBE4max-SpRY allow now to test several  
279 sgRNAs for the mutation of interest and play with the C base localization within the  
280 editing window to increase the C-to-T conversion efficacy compared to the use of the  
281 classical ancBE4max. These properties allow also to exclude other Cs present in the  
282 editing window to avoid the generation of other unwanted mutations near the targeted  
283 C. We furthermore demonstrated that using NA and NG PAMs we were able to  
284 precisely and simultaneously perform the Tyr (W273\*) and Rb1 (W63\*) mutations with  
285 high efficiency rates and both mutations were transmitted to the germline (Fig. 3). In  
286 this line, we also performed simultaneously 3 different and precise mutagenesis using

287 3 different PAM sequences (Fig. 4). This ability now in zebrafish is extremely useful if  
288 several mutations need to be introduced in order to model a human genetic disease  
289 such as cancers, especially if some mutations are located on the same chromosome.  
290 Different CRISPR co-selection methods have been engineered in *Drosophila*, *C.*  
291 *elegans* and cultured cells in order to phenotypically detect and enrich the cells or  
292 animals in which more mutagenesis events are taking place, by adding an sgRNA  
293 conferring a phenotypical read-out if the mutagenesis occurred<sup>24, 25, 26, 27, 28</sup>. Base  
294 editing co-selection strategies were recently demonstrated in cultured cells<sup>21</sup> but have  
295 not been reported in animals so far. Moreover, these time-saving strategies have never  
296 been developed in zebrafish. We here developed a co-selection method for base  
297 editing in zebrafish to prescreen phenotypically injected embryos based on the  
298 detection of pigmentation defects. The method is based on the addition of the *tyrW273\**  
299 *NA* sgRNA to the micro-injection mix. Importantly mutations in the *tyr* gene are viable  
300 and do not affect most developmental processes in zebrafish. We indeed have shown  
301 in this study that using this strategy we could select the albino embryos which were  
302 the most C-to-T converted for *nras* and *rb1* genes (Fig. 4b, c). We also developed a  
303 zebrafish model combining activation mutation for *nras* oncogene and knock-out of  
304 *tp53* tumor suppressor gene revealing an increase of melanocytes (Fig. 5), a clear  
305 melanoma predisposing phenotype. Indeed, fish over-expressing human mutated  
306 HRAS oncogene in melanocytes were hyperpigmented at 3dpf and developed  
307 melanoma at the adult stage<sup>15</sup>. The high efficiencies of CBE4max-SpRY obtained in  
308 this study and the possibility to precisely mutate simultaneously several genes using  
309 different PAMs pave the way for future applications in a tissue specific manner and for  
310 genetic disease modeling. For example, it could be implemented in the MAZERATI  
311 (Modeling Approach in Zebrafish for Rapid Tumor Initiation) system<sup>29</sup> in order to rapidly  
312 model and study *in vivo* combinations of endogenous mutations occurring in complex  
313 multigenic disorders. Finally, the high flexibility and efficiency of our method to induce  
314 combination of specific mutations will allow to rapidly create zebrafish cancer models  
315 combining the precise set of mutations found in individual patients. In the long term,  
316 these models could be used for rapid and patient specific drug screening for advanced  
317 personalized medicine<sup>17, 30</sup>.

318 **Acknowledgements**

319 We thank Panagiotis Antoniou and Annarita Miccio for sharing the *pCAG-CBE4max-*  
320 *SpG-P2A-EGFP* and *pCAG-CBE4max-SpRY-P2A-EGFP* plasmids<sup>14</sup>. M.R. was  
321 supported by the Fondation pour la Recherche Médicale (FRM grant number  
322 ECO20170637481) and la Ligue Nationale Contre le Cancer. Work in the Del Bene  
323 laboratory was supported by ANR-18-CE16 "iReelAx", UNADEV in partnership with  
324 ITMO NNP/AVIESAN (national alliance for life sciences and health) in the framework  
325 of research on vision and IHU FOReSIGHT [ANR-18-IAHU-0001] supported by French  
326 state funds managed by the Agence Nationale de la Recherche within the  
327 Investissements d'Avenir program.

328

329 **Author contributions**

330 M.R. and M.S. did the experimental works and analyzed the results. M.C.M. helped  
331 the design of targets, reviewed and edited the manuscript. M.R., J.P.C., F.D.B.  
332 conceived the project and wrote the article. J.P.C., F.D.B. co supervised the study.

333

334 **Competing interests statement**

335 The authors declare that they have no conflict of interest.

336

337 **Subjects**

338 Prime editing, base editors, zebrafish, genome modifications

339

340 **References**

341

- 342 1. Sander JD, Joung JK. CRISPR-Cas systems for editing, regulating and  
343 targeting genomes. *Nat Biotechnol* **32**, 347-355 (2014).  
344
- 345 2. Hwang WY, *et al.* Efficient genome editing in zebrafish using a CRISPR-Cas  
346 system. *Nat Biotechnol* **31**, 227-229 (2013).  
347
- 348 3. Wierson WA, *et al.* Efficient targeted integration directed by short homology in  
349 zebrafish and mammalian cells. *Elife* **9**, (2020).  
350
- 351 4. Prykhozhij SV, *et al.* Optimized knock-in of point mutations in zebrafish using  
352 CRISPR/Cas9. *Nucleic Acids Res* **46**, e102 (2018).  
353
- 354 5. Tessadori F, *et al.* Effective CRISPR/Cas9-based nucleotide editing in zebrafish  
355 to model human genetic cardiovascular disorders. *Dis Model Mech* **11**, (2018).  
356
- 357 6. Komor AC, Kim YB, Packer MS, Zuris JA, Liu DR. Programmable editing of a  
358 target base in genomic DNA without double-stranded DNA cleavage. *Nature*  
359 **533**, 420-424 (2016).  
360
- 361 7. Komor AC, *et al.* Improved base excision repair inhibition and bacteriophage  
362 Mu Gam protein yields C:G-to-T:A base editors with higher efficiency and  
363 product purity. *Sci Adv* **3**, eaao4774 (2017).  
364
- 365 8. Koblan LW, *et al.* Improving cytidine and adenine base editors by expression  
366 optimization and ancestral reconstruction. *Nat Biotechnol* **36**, 843-846 (2018).  
367
- 368 9. Zhang Y, *et al.* Programmable base editing of zebrafish genome using a  
369 modified CRISPR-Cas9 system. *Nat Commun* **8**, 118 (2017).  
370
- 371 10. Carrington B, Weinstein RN, Sood R. BE4max and AncBE4max Are Efficient in  
372 Germline Conversion of C:G to T:A Base Pairs in Zebrafish. *Cells* **9**, (2020).  
373
- 374 11. Zhao Y, Shang D, Ying R, Cheng H, Zhou R. An optimized base editor with  
375 efficient C-to-T base editing in zebrafish. *BMC Biol* **18**, 190 (2020).  
376
- 377 12. Rosello M, *et al.* Precise base editing for the *in vivo* study of  
378 developmental signaling and human pathologies in zebrafish. *bioRxiv*,  
379 2020.2012.2012.422520 (2020).  
380
- 381 13. Gaudelli NM, *et al.* Programmable base editing of A\*T to G\*C in genomic DNA  
382 without DNA cleavage. *Nature* **551**, 464-471 (2017).  
383
- 384 14. Walton RT, Christie KA, Whittaker MN, Kleinstiver BP. Unconstrained genome  
385 targeting with near-PAMless engineered CRISPR-Cas9 variants. *Science* **368**,  
386 290-296 (2020).  
387

- 388 15. Santoriello C, *et al.* Kita driven expression of oncogenic HRAS leads to early  
389 onset and highly penetrant melanoma in zebrafish. *PLoS One* **5**, e15170 (2010).  
390
- 391 16. Cayuela ML, *et al.* The Zebrafish as an Emerging Model to Study DNA Damage  
392 in Aging, Cancer and Other Diseases. *Front Cell Dev Biol* **6**, 178 (2018).  
393
- 394 17. Cagan RL, Zon LI, White RM. Modeling Cancer with Flies and Fish. *Dev Cell*  
395 **49**, 317-324 (2019).  
396
- 397 18. Callahan SJ, *et al.* Cancer modeling by Transgene Electroporation in Adult  
398 Zebrafish (TEAZ). *Dis Model Mech* **11**, (2018).  
399
- 400 19. Casey MJ, Stewart RA. Pediatric Cancer Models in Zebrafish. *Trends Cancer*  
401 **6**, 407-418 (2020).  
402
- 403 20. Hu JH, *et al.* Evolved Cas9 variants with broad PAM compatibility and high DNA  
404 specificity. *Nature* **556**, 57-63 (2018).  
405
- 406 21. Li S, *et al.* Universal toxin-based selection for precise genome engineering in  
407 human cells. *Nat Commun* **12**, 497 (2021).  
408
- 409 22. Yen J, *et al.* The genetic heterogeneity and mutational burden of engineered  
410 melanomas in zebrafish models. *Genome Biol* **14**, R113 (2013).  
411
- 412 23. Dovey M, White RM, Zon LI. Oncogenic NRAS cooperates with p53 loss to  
413 generate melanoma in zebrafish. *Zebrafish* **6**, 397-404 (2009).  
414
- 415 24. Ewen-Campen B, Perrimon N. ovo(D) Co-selection: A Method for Enriching  
416 CRISPR/Cas9-Edited Alleles in Drosophila. *G3 (Bethesda)* **8**, 2749-2756  
417 (2018).  
418
- 419 25. Ge DT, Tipping C, Brodsky MH, Zamore PD. Rapid Screening for CRISPR-  
420 Directed Editing of the Drosophila Genome Using white Coconversion. *G3*  
421 *(Bethesda)* **6**, 3197-3206 (2016).  
422
- 423 26. Kim H, *et al.* A co-CRISPR strategy for efficient genome editing in  
424 *Caenorhabditis elegans*. *Genetics* **197**, 1069-1080 (2014).  
425
- 426 27. Liao S, Tammaro M, Yan H. Enriching CRISPR-Cas9 targeted cells by co-  
427 targeting the HPRT gene. *Nucleic Acids Res* **43**, e134 (2015).  
428
- 429 28. Agudelo D, *et al.* Marker-free coselection for CRISPR-driven genome editing in  
430 human cells. *Nat Methods* **14**, 615-620 (2017).  
431
- 432 29. Ablain J, *et al.* Human tumor genomics and zebrafish modeling identify SPRED1  
433 loss as a driver of mucosal melanoma. *Science* **362**, 1055-1060 (2018).  
434
- 435 30. Letrado P, de Miguel I, Lamberto I, Diez-Martinez R, Oyarzabal J. Zebrafish:  
436 Speeding Up the Cancer Drug Discovery Process. *Cancer Res* **78**, 6048-6058  
437 (2018).

- 438  
439 31. Kluesner MG, *et al.* EditR: A Method to Quantify Base Editing from Sanger  
440 Sequencing. *CRISPR J* 1, 239-250 (2018).  
441  
442

## 443 **Methods**

444

### 445 **Fish lines and husbandry**

446 Zebrafish (*Danio rerio*) were maintained at 28 °C on a 14 h light/10 h dark cycle. Fish  
447 were housed in the animal facility of our laboratory which was built according to the  
448 respective local animal welfare standards. All animal procedures were performed in  
449 accordance with French and European Union animal welfare guidelines. Animal  
450 handling and experimental procedures were approved by the Committee on ethics of  
451 animal experimentation.

452

### 453 **Molecular cloning**

454 To generate the *pCS2+\_CBE4max-SpG* and the *pCS2+\_CBE4max-SpRY* plasmids,  
455 each *CBE4max-SpG* and *CBE4max-SpRY* sequence has been inserted into *pCS2+*  
456 *plasmid* linearized with EcoRI using the Gibson Assembly Cloning Kit (New England  
457 Biolabs). The fragment has been amplified using the primers F-5'-  
458 TGCAGGATCCCATCGATTCCGCCACCATGAAACGGACAG -3' and R-5'-  
459 TAGAGGCTCGAGAGGCCTTGTCAGACTTTCCTCTTCTTCTTGG -3') from the  
460 *pCAG-CBE4max-SpG-P2A-EGFP* plasmid (Addgene plasmid #139998)<sup>14</sup> and from  
461 the *pCAG-CBE4max-SpRY-P2A-EGFP* plasmid (Addgene plasmid #139999)<sup>14</sup>.

462

### 463 **mRNAs synthesis**

464 *pCS2+\_CBE4max-SpG* plasmid has been used to generate *CBE4max-SpG* mRNA *in*  
465 *vitro*. *pCS2+\_CBE4max-SpRY* plasmid has been used to generate *CBE4max-SpRY*  
466 mRNA *in vitro*. Each plasmid was linearized with NotI restriction enzyme and mRNAs  
467 were synthesized by *in vitro* transcription with 1 µL of GTP from the kit added to the  
468 mix and lithium chloride precipitation (using the mMESSAGING mMACHINE sp6 Ultra kit  
469 #AM1340, Ambion).

470

### 471 **sgRNA design**

472 The *sequenceParser.py* python script<sup>12</sup> was used to design *tyr* sgRNAs. All the  
473 synthetic sgRNAs were synthesized by IDT as Alt-R® CRISPR-Cas9 crRNA. Prior  
474 injections, 2 µL of the crRNA (100 pmol/µL) and 2 µL of Alt-R® CRISPR-Cas9 tracrRNA

475 (100 pmol/ $\mu$ L) from IDT were incubated at 95°C for 5 min, cooled down at room  
476 temperature and then kept in ice.

477 **Micro-injection**

478 To make the mutagenesis with base editing, a mix of 1 nL of *CBE* mRNA and synthetic  
479 sgRNAs was injected into the cell at one-cell stage zebrafish embryos. For the single  
480 mutagenesis, the final concentration was 600 ng/μL for *CBE* mRNA and 43 pmol/μL  
481 for sgRNA. For the double *rb1* and *tyr* mutations, the final concentration was 600 ng/μL  
482 for *CBE* mRNA and 21 pmol/μL for each sgRNAs. For the double *rb1* and *nras*  
483 mutations, the final concentration was 600 ng/μL for *CBE* mRNA and 8,6 pmol/μL for  
484 *nras* sgRNA and 34,4 pmol/μL for *rb1* sgRNA. For the double *p53* and *nras* mutations,  
485 the final concentration was 600 ng/μL for *CBE* mRNA and 8,6 pmol/μL for *nras* sgRNA  
486 and 34,4 pmol/μL for *tp53* sgRNA. For the *tyr*, *nras* and *rb1* mutations, the final  
487 concentration was 600 ng/μL for *CBE* mRNA, 8,6 pmol/μL for *nras* sgRNA and 8,6  
488 pmol/μL for *tyr* sgRNA and 25,8 pmol/μL for *rb1* sgRNA.

489

490 **Whole-embryos DNA sequencing**

491 For genomic DNA extraction, embryos were digested for 1 h at 55°C in 0.5 mL lysis  
492 buffer (10 mM Tris, pH 8.0, 10 mM NaCl, 10 mM EDTA, and 2% SDS) with proteinase  
493 K (0.17 mg/mL, Roche Diagnostics) and inactivated 10 min at 95°C. To sequence and  
494 check for frequency of mutations, each target genomic locus was PCR-amplified using  
495 Phusion High-Fidelity DNA polymerase (Thermo Scientific). For the *tyr* locus, a PCR  
496 was performed with primers Fwd-5'-ATCGGGTGTATCTGCTGTTTTGG-3' and Rev-  
497 5'-CCATACCGCCCCTAGAACTAACATT-3'. For the *rb1* locus, the primers used were  
498 Fwd-5'-TCTGTCAACTGTTGTTTTCCAGAC-5' and Rev-5'-  
499 CAATAAAAAGACAAGCTCCCCACTG-5'. For the *nras* locus, the primers used were  
500 Fwd-5'- CTTTTCTCTTTTTGTCTGGGTG-5' and Rev-5'-  
501 CGCAATCTCACGTTAATTGTAGTGT-5'. For the *tp53* locus, the primers used were  
502 Fwd-5'- ATATCTTGTCTGTTTTCTCCCTGCT-5' and Rev-5'-  
503 GTCCTACAAAAGGCTGTGACATAC-5'. The DNAs have been extracted on an  
504 agarose gel and purified (using the PCR clean-up gel extraction kit #740609.50,  
505 Macherey-Nagel) and the sanger sequencings have been performed by Eurofins. The  
506 sequences were analyzed using ApE software and quantifications of the mutation rates  
507 done using editR online software<sup>31</sup>.

508

509 **Imaging**

510 Embryos were oriented in egg solution with an anesthetic (Tricaine 0,013%). Leica  
511 MZ10F was used to image them. Adult fish were imaged using a net and an Iphone  
512 xs.

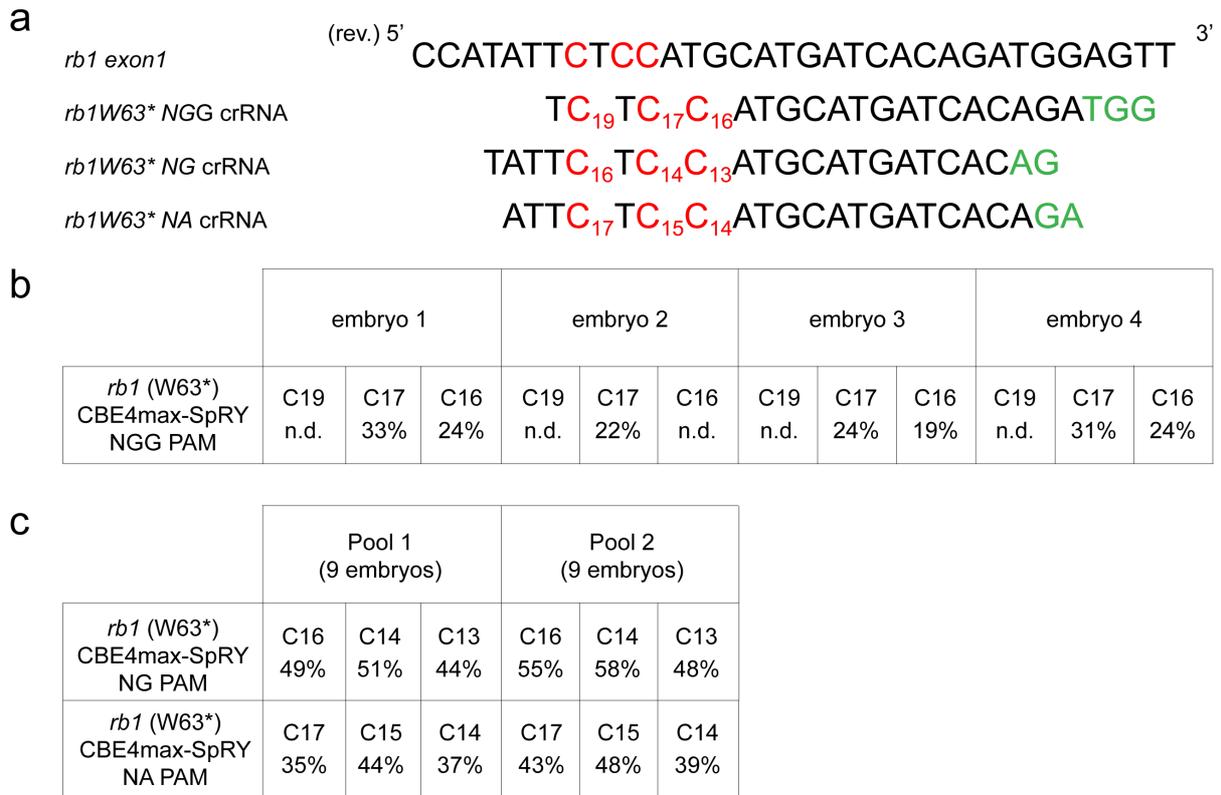
513

#### 514 **Statistics**

515 A non-parametric t-test with the Mann–Whitney correction was applied to determine  
516 significance in base editing. The software used was Prism 7 (GraphPad).

517 **Figures**

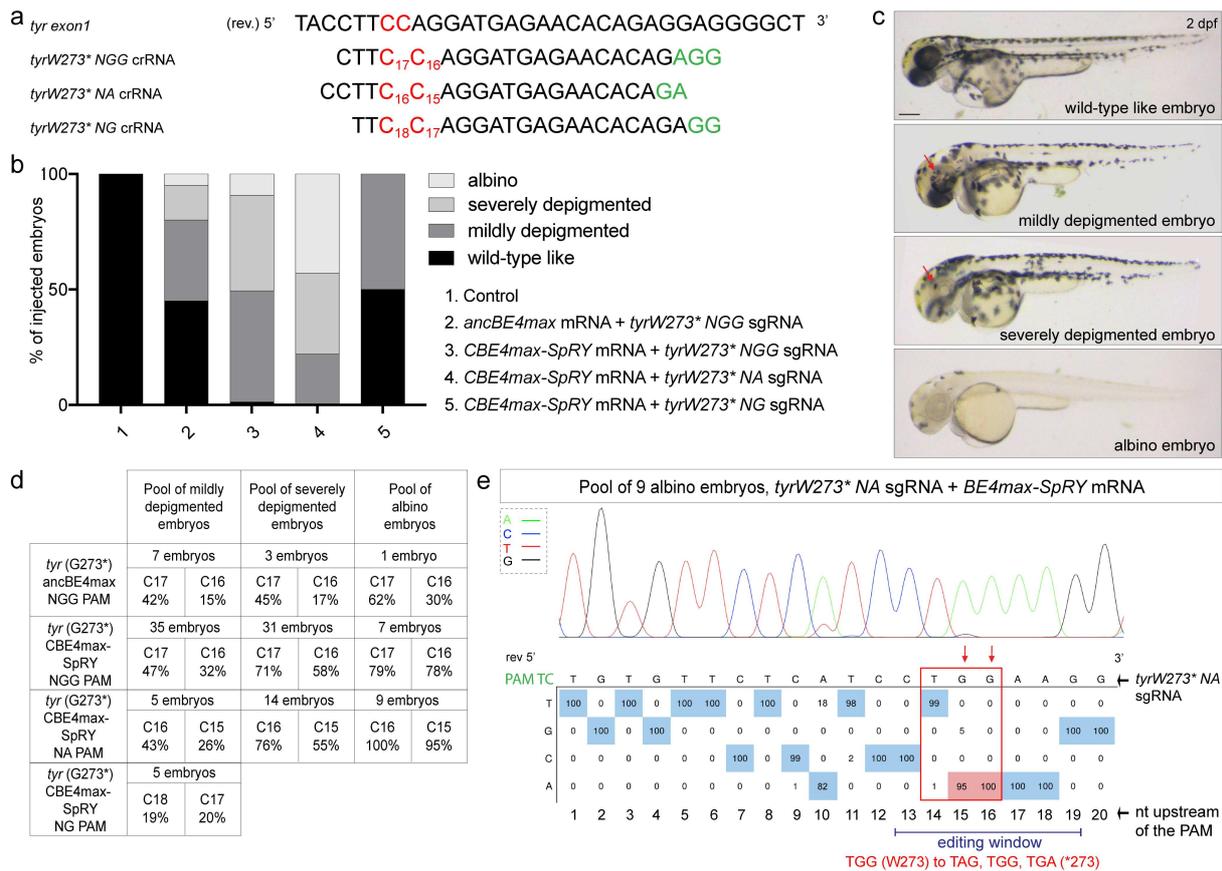
518



519

520 **Figure 1. Efficient C-to-T conversion using the CBE4max-SpRY variant with NA**  
 521 **and NG PAMs.**

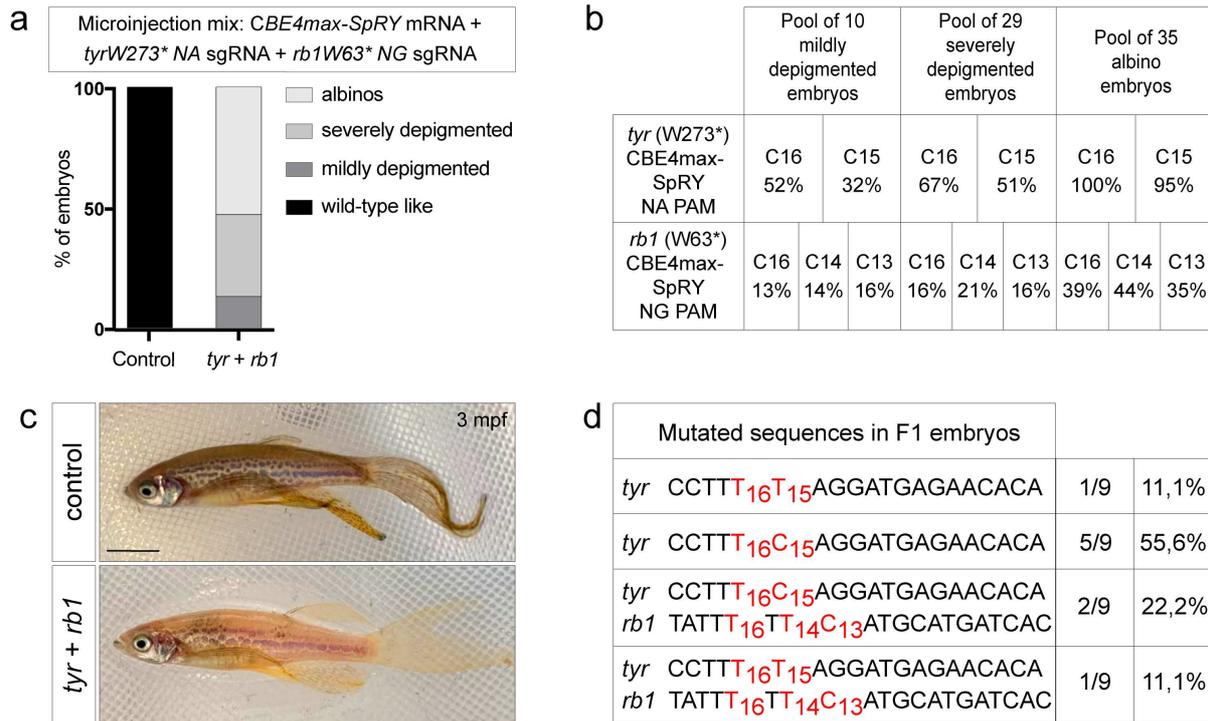
522 **a.** Targeted genomic sequence of the exon 1 of the *rb1* tumor suppressor gene and  
 523 the 3 different sgRNAs used to introduce the W63\* mutation. For each sgRNA, the  
 524 targeted Cs are in red and the PAM sequence is in green. **b.** C-to-T conversion  
 525 efficiency for each targeted Cs in the 4 single edited embryos out of 11 analyzed  
 526 embryos injected with *CBE4max-SpRY* mRNA and *rb1W63\** NGG sgRNA. **c.** C-to-T  
 527 conversion efficiency for each targeted Cs in 2 pools of 9 embryos injected with  
 528 *CBE4max-SpRY* mRNA and *rb1W63\** NGG sgRNA.



529  
530  
531  
532  
533  
534  
535  
536  
537  
538  
539  
540  
541  
542  
543  
544  
545  
546  
547  
548  
549  
550  
551  
552  
553

## Figure 2. Total base conversion in the F0 embryos with *CBE4max-SpRY* and a NA PAM.

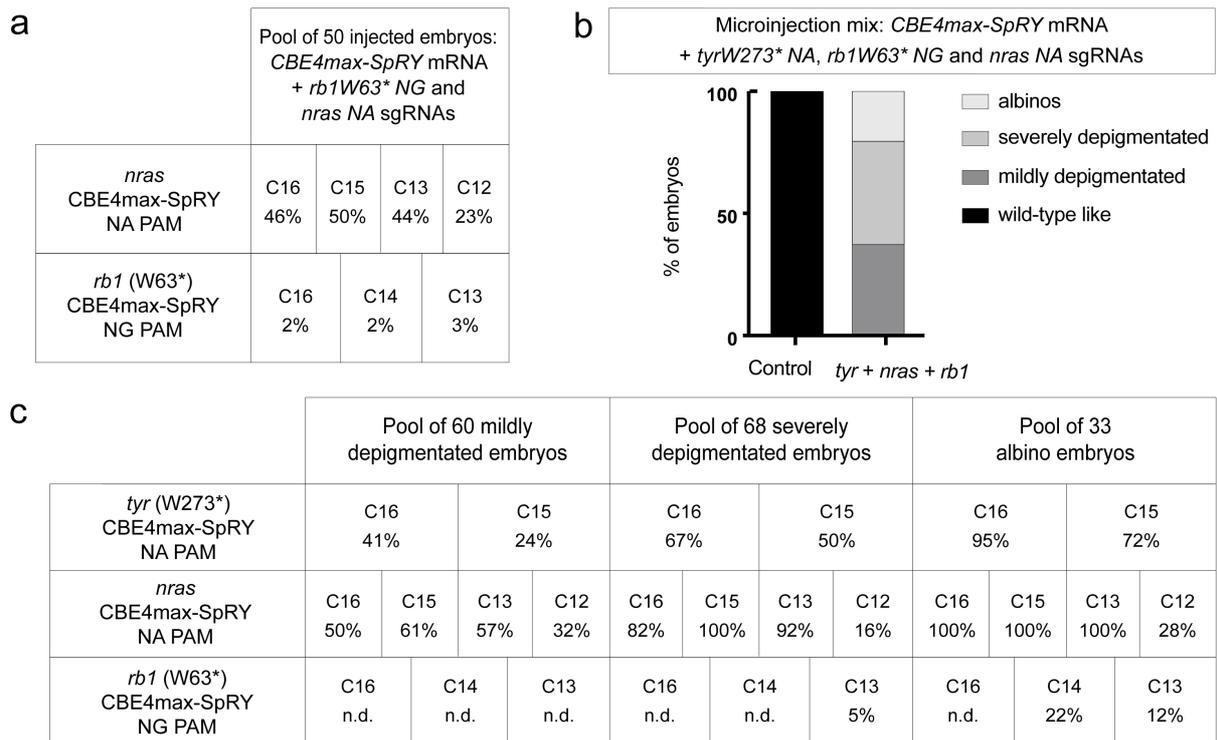
**a.** Targeted genomic sequence of the exon 1 of the *tyrosinase* gene and the 3 different sgRNAs used to introduce the W273\* mutation. For each sgRNA, the targeted Cs are in red and the PAM sequence is in green. **b.** Proportion of the 4 groups based on the pigmentation defects described in Fig. 2c for each injection: the *ancBE4max* mRNA and the *tyr*W273\*NGG sgRNA (column 2, 19 embryos in total), the *CBE4max-SpRY* mRNA and the *tyr*W273\*NGG sgRNA (column 3, 74 embryos in total), the *CBE4max-SpRY* mRNA and the *tyr*W273\*NA sgRNA (column 4, 28 embryos in total), the *CBE4max-SpRY* mRNA and the *tyr*W273\*NG sgRNA (column 5, 10 embryos in total). **c.** Lateral view of representative 2 days post-fertilization (dpf) embryos showing different severity of pigmentation defects (wild-type like, mildly depigmented, severely depigmented and albino embryos). Scale bar = 100  $\mu$ m. **d.** C-to-T conversion efficiency for each targeted Cs and each pool of embryos showing pigmentation defects presented in Fig. 2b. **e.** DNA sequencing chromatogram of the targeted *tyr* gene from a pool of 9 albino embryos obtained after the injection of the *CBE4max-SpRY* mRNA with the *tyr*W273\*NA sgRNA. W273\* mutation in Tyr upon C-to-T conversion in *tyr* reached 100% for the C16 base and 95% for the C15 base of gene editing efficiency. Numbers in the boxes represent the percentage of each base at that sequence position. In red are highlighted the base substitutions introduced by base editing while the original bases are in blue. The color code of the chromatogram is indicated in the upper left corner (Adenine green, Cytosine blue, Thymine red, Guanine black). The distance from the PAM sequence of the targeted C base is indicated below the chromatogram<sup>31</sup>.



554  
555  
556  
557  
558  
559  
560  
561  
562  
563  
564  
565  
566

**Figure 3. Germline transmission of two mutations generated simultaneously using NA and NG PAMs.**

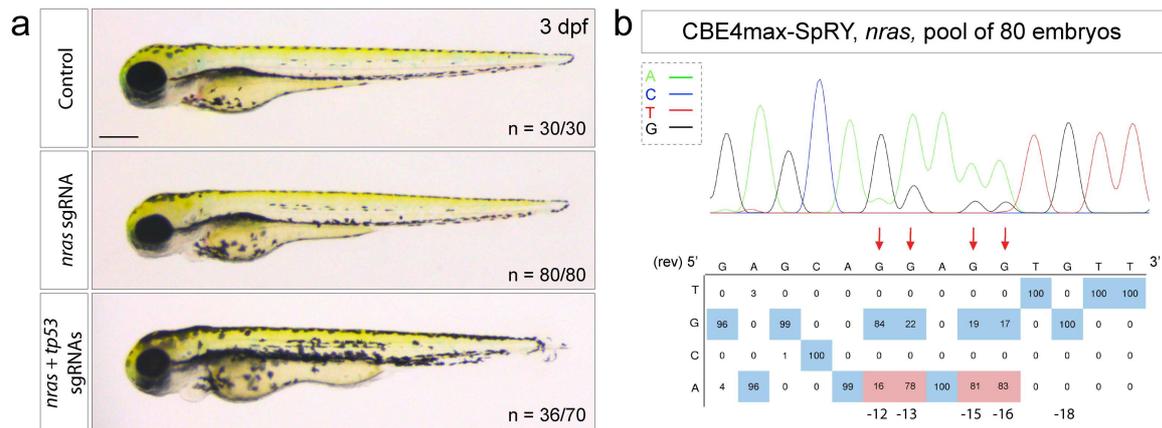
**a.** Proportion of the 4 groups based on the pigmentation lack defects described in Fig. 2c for embryos injected with the *CBE4max-SpRY* mRNA, the *tyrW273\*NA* and the *rb1W63\*NG* sgRNA (column 2, 74 embryos in total). **b.** C-to-T conversion efficiency for each targeted Cs and each pool of embryos showing pigmentation defects presented in Fig. 3a for the *tyrosinase* and *rb1* genes. **c.** Lateral view of a 3 months F0 fish injected at one-cell stage embryo with the *CBE4max-SpRY* mRNA, the *tyrW273\*NA* and the *rb1W63\*NG* sgRNAs and showing pigmentation defects. Scale bar = 5 mm. **d.** Sequenced *tyr* and *rb1* loci of 9 F1 single embryos from the founder fish in Fig. 3c. 6 embryos out of 9 were edited for *tyrosinase* and 3 embryos out of 9 were double edited for *tyrosinase* and *rb1*.



567  
568  
569  
570  
571  
572  
573  
574  
575  
576

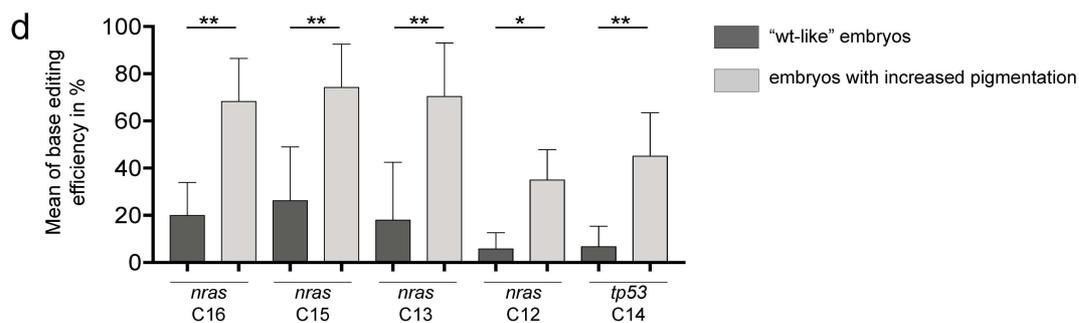
**Figure 4. Base editing co-selection method.**

**a.** C-to-T conversion efficiency for each targeted Cs in *nras* and *rb1* genes in a pool of 50 embryos injected with *CBE4max-SpRY* mRNA and *nras NA* and *rb1 W63\** sgRNAs. **b.** Proportion of the 4 groups based on the pigmentation defects described in Fig. 2c for the embryos injected with the *CBE4max-SpRY* mRNA, the *tyrW273\*NA*, *nras NAA* and *rb1W63\*NG* sgRNAs (column 2, 161 embryos in total). **c.** C-to-T conversion efficiency for each targeted Cs and each pool of embryos showing pigmentation defects presented in Fig. 4b for the *tyrosinase*, *nras* and *rb1* genes. The albino embryos are the most edited for the 3 different loci.



**c**

embryos injected with <i>CBE4max-SpRY</i> mRNA + <i>nras</i> and <i>tp53</i> sgRNAs	<i>nras</i>				<i>tp53</i>
	C16	C15	C13	C12	C14
“wt-like” embryo 1	27%	29%	6%	n.d.	9%
“wt-like” embryo 2	3%	2%	n.d.	n.d.	n.d.
“wt-like” embryo 3	15%	18%	12%	9%	n.d.
“wt-like” embryo 4	35%	56%	54%	14%	18%
embryo 1 with increased pigmentation	63%	68%	55%	29%	28%
embryo 2 with increased pigmentation	61%	66%	59%	41%	40%
embryo 3 with increased pigmentation	36%	42%	32%	8%	29%
embryo 4 with increased pigmentation	80%	97%	100%	47%	49%
embryo 5 with increased pigmentation	67%	70%	72%	33%	36%
embryo 6 with increased pigmentation	75%	80%	73%	32%	54%
embryo 7 with increased pigmentation	64%	72%	72%	46%	40%
embryo 8 with increased pigmentation	100%	99%	100%	44%	85%



577  
578  
579  
580  
581  
582  
583  
584  
585  
586  
587  
588  
589

**Figure 5. Endogenous activation of *nras* oncogene and knock-out of *tp53* tumor suppressor gene led to an increase of melanocyte numbers.**

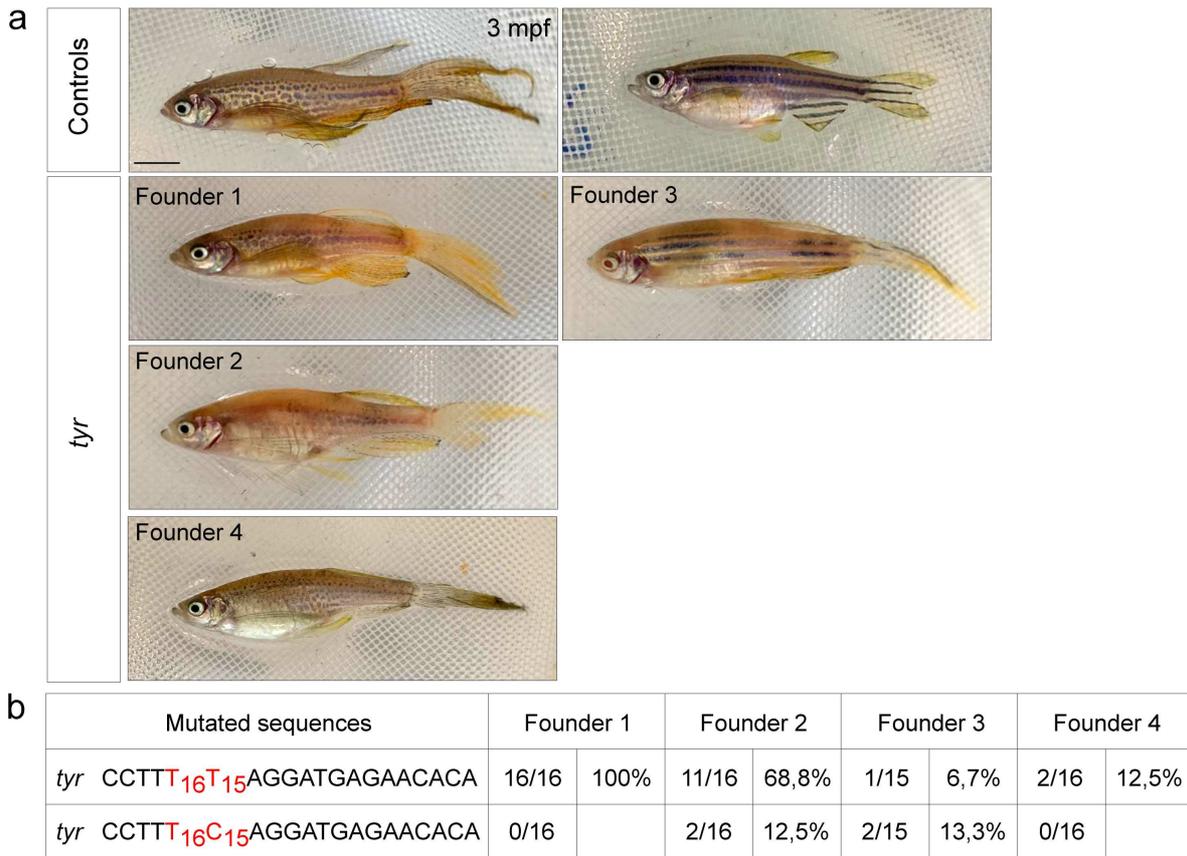
**a.** Lateral view of 3 dpf embryos. The injected embryos edited only for *nras* do not present any defects whereas 50% of the injected embryos with *nras* NA and *tp53* Q170\* sgRNAs show an increase of pigmentation. Scale bar = 100  $\mu$ m. **b.** DNA sequencing chromatogram of the targeted *nras* gene from a pool of 80 injected embryos with the *CBE4max-SpRY* mRNA and the *nras* NA sgRNA. C-to-T conversion shows 83% of efficiency for the C16, 81% for the C15, 78% for the C13 and 16% for the C12. Numbers in the boxes represent the percentage of each base at that sequence position. In red are highlighted the base substitutions introduced by base editing while the original bases are in blue. The color code of the chromatogram is indicated in the upper left corner (Adenine green, Cytosine blue, Thymine red, Guanine

590 black). The distance from the PAM sequence of the targeted C base is indicated below  
591 the chromatogram<sup>31</sup>. **c.** C-to-T conversion efficiency for each targeted Cs in *nras* and  
592 *tp53* genes in single embryos injected with *CBE4max-SpRY* mRNA and *nras NA* and  
593 *tp53 Q170\** sgRNAs. 4 single embryos which did not show an increase of pigmentation  
594 named as “wt-like” embryos and 8 single embryos showing an increase of pigmentation  
595 have been sequenced. **d.** Histogram showing the mean with standard deviation of the  
596 base editing efficiency for the 4 “wt-like” embryos and the 8 embryos with increased  
597 pigmentation for Fig. 5c. Mann–Whitney test.  
598  
599

Supplementary information

Disease modeling by efficient genome editing using a near PAM-less base editor  
*in vivo*.

(Rosello et al.)



**Supplementary Figure 1. High germline transmission rate.**

**a.** Lateral view of a 3 months F0 fish injected at one-cell stage with the *CBE4max-SpRY* mRNA and the *tyrW273\*NA* sgRNA and showing pigmentation lack defect. Scale bar = 5 mm. **b.** Sequenced *tyrosinase* locus of F1 single embryos randomly chosen from each founder. 16 embryos out of 16 were edited for *tyrosinase* for the Founder 1, 13 embryos out of 16 for the Founder 2, 3 embryos out of 15 for the Founder 3 and 2 embryos out of 16 for the Founder 4.

## II.3. THE USE OF BASE EDITORS IN ZEBRAFISH TO STUDY MUTATIONS IN *PIP5K1C* GENE FOUND IN PATIENTS WITH A POLY-MALFORMATIVE SYNDROME.

### II.3.1. CONTEXT.

Three different heterozygous *de novo* missense mutations in *PIP5K1C* gene were identified in 6 young patients, all presenting developmental delay, neurological defects and varying dysmorphic features.

*PIP5K1C* gene encodes for the Phosphatidyl-4-Phosphate 5-Kinase Type 1 Gamma, a kinase of the PIPK kinase family. These kinases regulate the levels of the second messenger phosphoinositides, which are important regulators of diverse cellular processes such as cell migration, adhesion, polarity and division<sup>215</sup>. Alterations of genes involved in the production and regulation of phosphoinositides led to numerous neurodegenerative human diseases<sup>216</sup> and psychiatric disorders<sup>217, 218, 219, 220, 221</sup>. Increase of PIPK activity is also involved in various types of cancer<sup>222, 223, 224, 225, 226, 227, 228, 229</sup>. *PIP5K1C* is a member of the PIP5K subfamily which phosphorylates the phosphatidylinositol 4-phosphate (PI(4)P) to form the phosphatidylinositol 4,5-biphosphate (PI(4,5)P<sub>2</sub>)<sup>230</sup>.

The single mutations found in the *PIP5K1C* locus of the six different subjects coming from independent families identified by our collaborators (*Dr. Morleo and colleagues*) are:

- c.436G>A in exon 5 leading to PIP5K1C(E146K). Subject 1, a 6-year-old male and Subject 2, a 5-year-old male.
- c.662A>G at exon 7 leading to PIP5K1C(Y221C). Subject 3, a 7 year-old-female and Subject 4, a 7-year-old male.
- c.614A>G at exon 6 leading to PIP5K1C(Y205C). Subject 5, an 8-year-old female
- c.700C>T at exon 7 PIP5K1C(R234C). Subject 6: a 4.5-year-old female.

These patients exhibit abnormalities including developmental delay, microcephaly, several seizures, structural brain abnormalities, dysmorphic and musculoskeletal features, optic atrophy or vision abnormalities, dentition and growth problems. Using

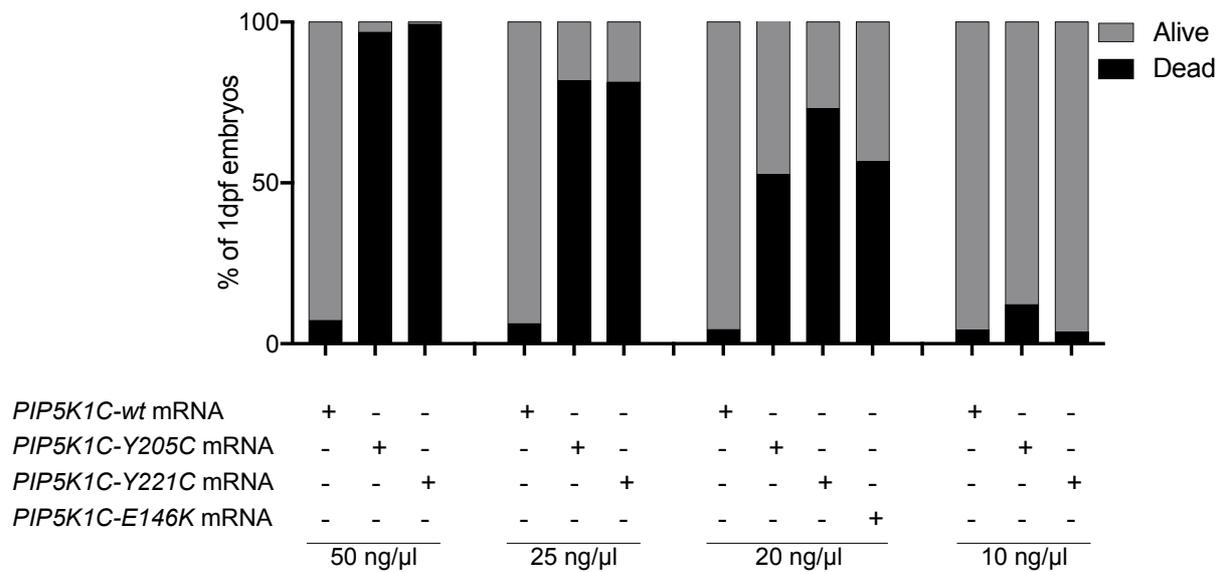
fibroblasts from these patients, an increase of PIP2 level and of the calcium flux has been found, suggesting that these mutations are activating mutations. Abnormalities in the actin cytoskeleton were also demonstrated in fibroblasts from one of the patients.

In this context, our aim was to elucidate *in vivo* whether these activating mutations of *PIP5K1C* are responsible for the developmental delay and the neurological phenotypes observed in the patients.

### II.3.2. CAUSALITY LINK BETWEEN MUTATIONS FOUND IN *PIP5K1C* OF PATIENTS AND THE DISEASE.

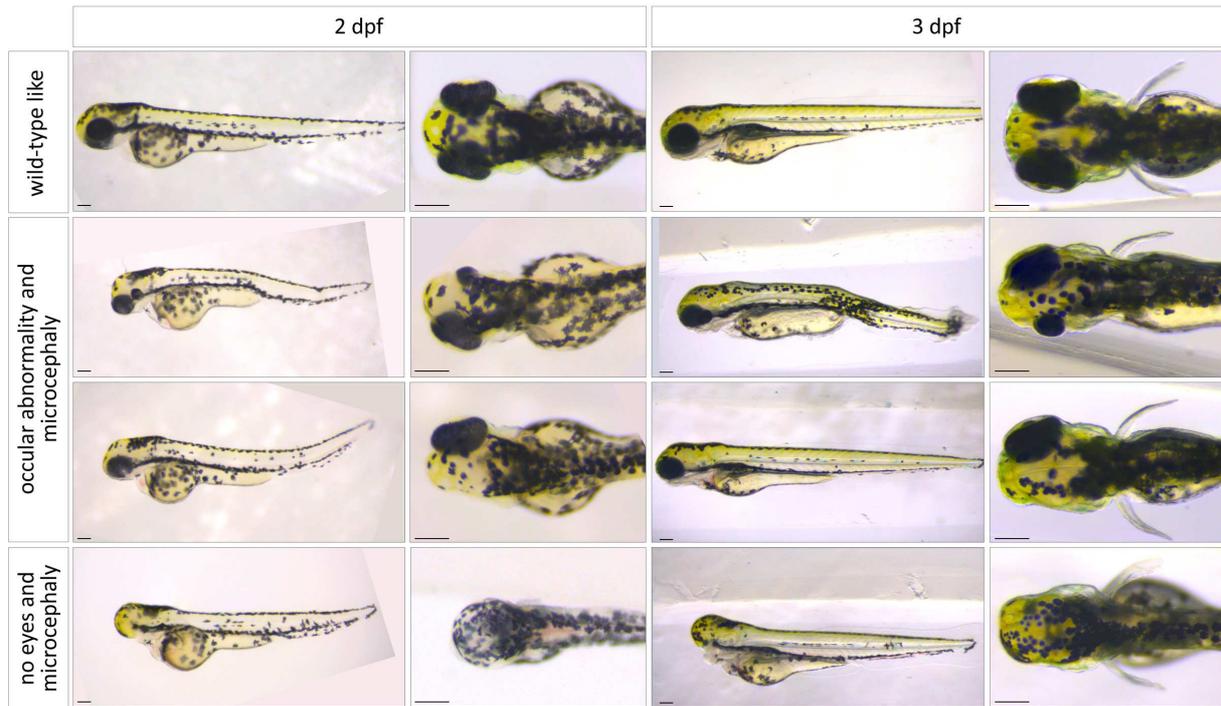
In order to investigate *in vivo* if there is a causal link between the neurological, ocular and craniofacial defects of the patients and their mutations in *PIP5K1C* gene, we have decided to overexpress the 3 different human variants as well as the wild-type *PIP5K1C* as a control by injecting their mRNA at one-cell stage zebrafish embryos.

We first codon optimized for zebrafish the human coding sequence of *PIP5K1C* using the codon optimization IDT tool and we cloned the sequence to make the mRNA synthesis *in vitro*. We then generated the same plasmids with different human mutations to get *PIP5K1C* (E146K), (Y205C) and (Y221C) variants. Upon the injections of the different mRNA of these variants, we obtained a high rate of mortality for the 3 mutated variants whereas almost all of the embryos injected the wild-type *PIP5K1C* mRNA survived (Fig. 10). We indeed first injected mRNAs at 50 ng/ $\mu$ l which resulted in more than 95% of mortality of the injected embryos for the Y205C and Y221C variants and less than 8% with the wild-type mRNA injection at 1 dpf. We then decreased the concentration to 25 ng/ $\mu$ l, 20 ng/ $\mu$ l and 10 ng/ $\mu$ l. We did not obtain any increased mortality and morphological defects at 10 ng/ $\mu$ l whereas we got between 27% and 57.5% of survivors with a range of different specific phenotypes at 2 dpf by injecting the mRNA at 20 ng/ $\mu$ l (Fig. 10). We therefore decided to use this concentration to further our analysis of the obtained phenotypic defects.



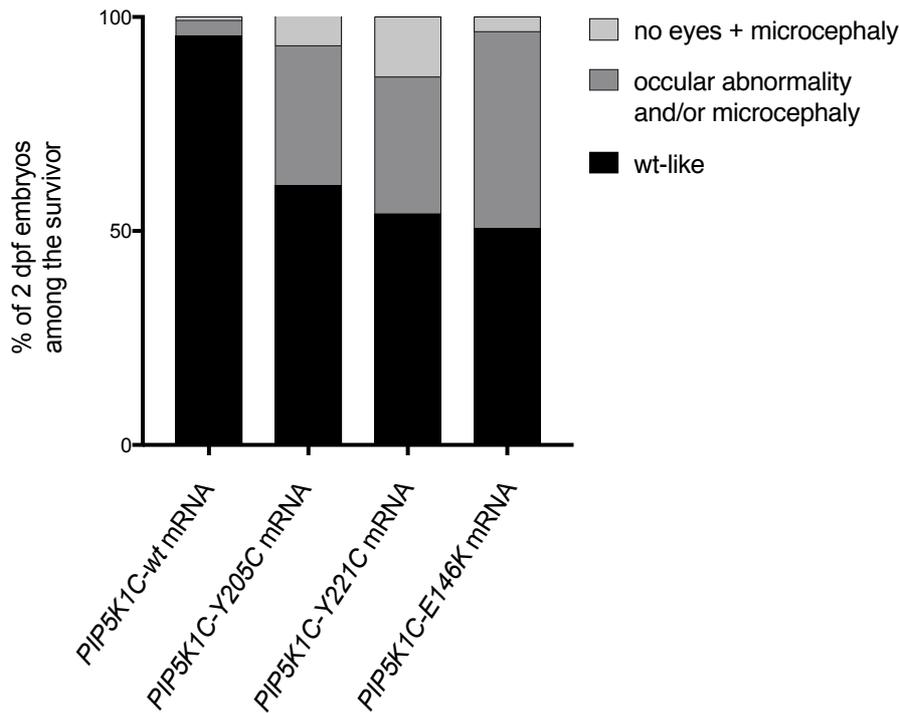
**Figure 10: High mortality rate observed by expressing the human PIP5K1C (E146K), (Y205C) and (Y221C) variants.** Quantifications of dead and alive embryos at 1 dpf after injection of *PIP5K1C-wild type*, *Y205C*, *Y221C* and *E146K* mRNAs at one cell-stage with different concentrations. The total number of embryos per column is: 224, 250, 149, 96, 176, 102, 243, 623, 545, 410, 119, 132 and 112.

After injections of the different mRNAs at 20 ng/μl, we defined 3 different groups in which we classified the injected embryos at 2 dpf based on the severity of the observed morphological defects: a first one composed by morphologically wild-type like embryos presenting no obvious phenotypes, a second one in which there were embryos with ocular abnormalities such as the development of one eye only or a small eye and/or microcephaly and finally a third group with embryos without any eye development and showing extreme microcephaly (Fig. 11).



**Figure 11: Embryos expressing PIP5K1C (Y205C), (Y221C) and (E146K) present a spectrum of ocular abnormalities and microcephaly at 2 and 3 dpf.** Panel of the phenotypes obtained in the embryos at 2 and 3 dpf after injections of *PIP5K1C* (Y205C), (Y221C) and (E146K) mRNAs at 20 ng/ $\mu$ l in one cell-stage embryos. Scale bar= 50  $\mu$ m.

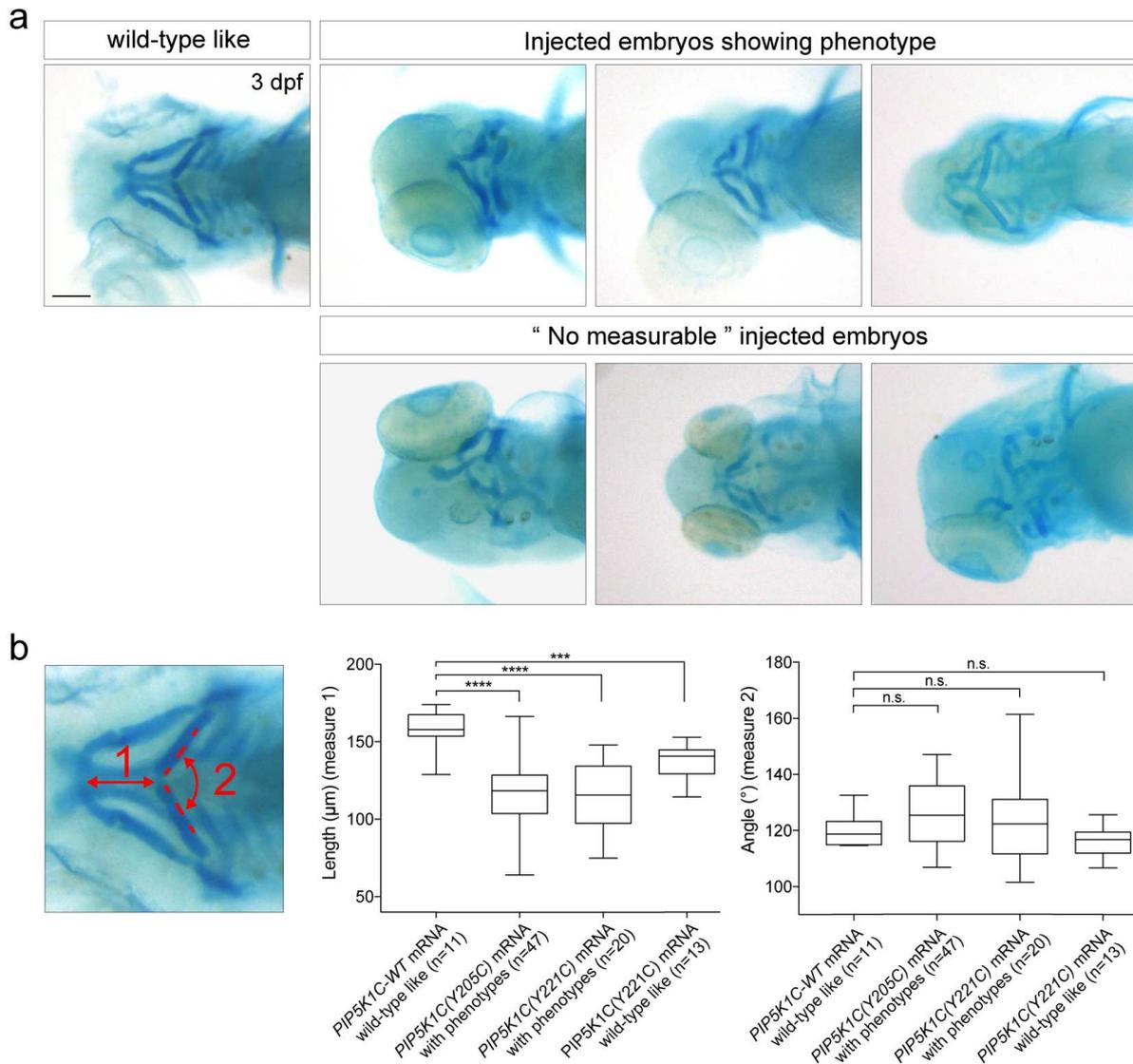
After quantifications, we could observe that with the expression of PIP5K1C (Y205C) variant 32.7% of the alive embryos showed at 2 dpf ocular abnormalities and/or microcephaly and 6.7% no eyes and microcephaly. With the expression of PIP5K1C (Y221C), 32% exhibited ocular abnormalities and/or microcephaly and 14% no eyes and microcephaly. With the expression of PIP5K1C (E146K), 46% presented ocular abnormalities and/or microcephaly and 3.4% no eyes and microcephaly whereas with the expression of the wild-type PIP5K1C, less than 4% of the injected embryos showed a phenotype, even when we increased the concentration up to 50 ng/ $\mu$ l (Fig. 12). Together our results suggest that the tested PIP5K1C mutant variants are functionally relevant and that the neurodevelopmental defects observed in the patients may be a result of the mutations identified in *PIP5K1C* gene of the patients.



**Figure 12: Ocular abnormality and microcephaly phenotypes are found upon the expression of PIP5K1C (Y205C), (Y221C) and (E146K).** Quantifications of the embryos showing the different phenotypes among the alive 2 dpf embryos injected with *PIP5K1C-wild-type*, *Y205C*, *Y221C* and *E146K* mRNAs at 20 ng/ $\mu$ l in one cell-stage embryos. The total number of embryos per column is: 235, 401, 157 and 173.

We next assessed whether these larvae have head structure defects. We thus performed an Alcian blue staining to visualize the cartilage structures present in the head of the injected larvae at 3 dpf for each group (Fig 12 and 13.a). For the “no eyes and microcephaly” group of each variant, no staining was observed, suggesting a total absence of cartilage formation in the head. We measured the distance between Meckel’s and ceratohyal cartilages along the antero-posterior axis (Fig. 13.b, measure 1) of larvae injected with the *PIP5K1C-wt* mRNA (n=11), larvae with ocular abnormalities injected with the *PIP5K1C(205C)* mRNA (n=47), larvae injected with the *PIP5K1C(221C)* mRNA with ocular abnormalities (n=20) and wild-type like (n=13). Among the larvae with ocular abnormalities injected with the *PIP5K1C(205C)* mRNA, 2 larvae did not show any cartilage structures at all and 12 larvae were not measurable (Fig. 13.a). Among the larvae with ocular abnormalities injected with the *PIP5K1C(221C)* mRNA, two larvae were not measurable. Overall, we can observe that the distance between Meckel’s and ceratohyal cartilages was reduced in the larvae injected with the *PIP5K1C(205C)* and (*Y221C*) mRNA with or without phenotype

compared to the one of the larvae injected with *PIP5K1C-wt* mRNA. These results suggest a specific compaction of the head in these larvae (Fig. 13.b, measure 1). We also measured the angle at the ceratohyal cartilage intersection but no difference was observed between the different tested conditions and the wt control larvae (Fig. 13.b, measure 2).



**Figure 13: Larvae expressing *PIP5K1C(205C)* and *PIP5K1C(Y221C)* show head cartilage defects. (a)** Alcian blue staining of 3 dpf injected larvae. 12 larvae injected with the *PIP5K1C(205C)* mRNA and 2 larvae injected with the *PIP5K1C(221C)* mRNA with ocular abnormalities were not measurable. Scale bar=150 $\mu\text{m}$ . **(b)** Quantification of the distance between Meckel's and ceratohyal cartilages (measure 1), and the ceratohyal angle (measure 2). Wilcoxon-Mann-Whitney test, n.s.: non-significant, \*\*\*,  $p < 0.001$ , \*\*\*\*,  $p < 0.0001$ .

Using zebrafish, we were able to demonstrate that the mutations identified in the patients cause defects of the forebrain, including eye development, as well as head cartilage abnormalities.

### II.3.3. ENDOGENOUS MUTATION OF *PIP5K1C* ZEBRAFISH ORTHOLOGUES USING BASE EDITORS: DISEASE MODELING AND LIMITATIONS.

We have previously shown by overexpressing the human *PIP5K1C* mutated variants mRNAs this induction of specific phenotypes that could support the role of these mutations in the origin of the neurological and craniofacial defects found in the patients. In order to validate our overexpression results and to develop a more physiological relevant zebrafish model for this poly-malformative syndrome, we have decided to introduce these mutations genetically resulting to equivalent endogenous modifications at the protein level in the zebrafish orthologue Pip5k1c.

In zebrafish, *PIP5K1C* gene has been duplicated resulting on the existence of *pip5k1ca* and *pip5k1cb* genes. Indeed, 320–350 million years ago, a teleost-specific whole genome duplication occurred at the root of the teleost lineage generating two paralogues for a vast number of fish genes. Both copies can then be kept, one or both genes can specialize each with a specific function (subfunctionalization), can acquire a novel function (neofunctionalization) and they can partition their territories of gene expression. The loss of one of the two duplicated genes can also occur<sup>231</sup>. For these reasons, we first aligned the protein sequences found in zebrafish to the human one in order to select the ortholog that is the most likely to carry the same function as the human *PIP5K1C* and thus the one that we should target to introduce the human mutations into the zebrafish genome. By alignment using clustal omega program on uniprot.org website, we found that both proteins are close to the human one, Pip5k1ca being the closest one (Fig. 14). Indeed, Pip5k1ca shows 85% of similarity and 71% of identity and Pip5k1cb 75% of similarity and 62% of identity with *PIP5K1C*. For both genes, the E146K mutation could be introduced in zebrafish after C-to-T conversion using CBE among the 3 mutations found in the patients presenting the poly-malformative syndrome (Fig. 14 in red). The human Y205C cannot be reproduced after C-to-T conversion and neither could the human Y221C as no Y orthologue is present in zebrafish (Fig. 14, red Y221C). Therefore, we decided first to introduce E146K mutation into the genome of the zebrafish using the CRIPSR/Cas9-based CBE technology. In zebrafish, the corresponding amino acid of the human E146 is E151 for Pip5k1ca and E144 for Pip5k1cb.

```

human_PIP5K1C      MELEVPDEAES-AEA--GAVPSEAAWAESGAAAGL----AQKKAAPTEVLSMTAQPGPG 53
zebrafish_Pip5k1ca MEAAAAEGAASLSEAGDGSPLSGATASEDAGSADDVDADSISKKAFITEM--PSSSGLPS 58
zebrafish_Pip5k1cb MEAAAAGCSALAEADAVLGC-----MEPAEIEAALKKTCLEMPSSSCHSSSSG 51
** . . : :* . . ** : :* :. .

human_PIP5K1C      HGKKLGHARGVDASGETTYKKTTSSTLKGAIQLGIGYTVGHLSSKPERDVLMDQDFYVVESI 113
zebrafish_Pip5k1ca HGKKIGHARGVDASGETTYKKTTSALKGAIQLGIGYTVGNLSSKPERDVLMDQDFYVVESI 118
zebrafish_Pip5k1cb PEKKIGHRRVDASGETTYKKTTSALKGAIQLGIGYTVGNLSSKPERDVLMDQDFYVVESI 111
**:* ** *****:*****:*****:*****

human_PIP5K1C      FFPSEGSNLTPAHHFQDFRFKTYAPVAFRYFRELFE146KFGIRPDDYLYSLCNEPLIELSNPGAS 173
zebrafish_Pip5k1ca FFPSEGSNLTPAHHFPDFRFKTYAPVAFRYFRELFE146KFGIRPDDYLYSLCNEPLIELSNPGAS 178
zebrafish_Pip5k1cb FFPSEGSNLTPAHHYPDFRFKTYAPVAFRYFRELFE146KFGIRPDDYLYSLCNEPLIELSNPGAS 171
*****:*****:*****:*****:*****

human_PIP5K1C      GSLFYVTSDDDEFIIKTVMHKEAEFLQKLLPGYYMNLNQNPRY205CTLLPKFY221CGLYCVQSGGKNI 233
zebrafish_Pip5k1ca GSVFYLTKDDEFIIKTVMHKEAEFLQKLLPGYYMNLNQNPRY205CTLLPKFY221CGLYCVQSGGKNI 238
zebrafish_Pip5k1cb GSIY205CFYVTRDDEFILKTVMHKEAEFLQKLLPGYYMNLNQNPRY205CTLLPKFY221CGLYCVQSGGKNI 231
**:* ** *****:*****:*****:*****:*****

human_PIP5K1C      RVVVMNY205CILPRVVMHLKFDLKGSTYKRRASKKEKESFPTYKDLDFMQDMPEGLLLDAD 293
zebrafish_Pip5k1ca RMVVMNY205CNVLPY205CRVY205CVMHLKYDLKGSTYKRRASKKEREKAKPTFKDLDFMQELPDGLMLDY205CTD 298
zebrafish_Pip5k1cb RIVVMNY205CNVLPY205CRVY205CVMHLKFDLKGSTYKRRASKKEREKRNY205CPTYKDLDFMQDY205CVQDGLLLDY205CVD 291
*:* ** *****:*****:*****:*****:*****

human_PIP5K1C      TFSALVKTY205CLQRDCLVLESFKIMDYSLLLGVHNIY205CDQHERERQAQGAQSTSDEKRPVQKAL 353
zebrafish_Pip5k1ca TYNALVKTY205CLQRDCLVLESFKIMDYSLLLGVHNIY205CDQAAKEQOMEGSQGNSDEKRPLAQKAL 358
zebrafish_Pip5k1cb TYNALLKTY205CLQRDCLVLESFKIMDYSLLLGVHY205CVEQAERERQMEGSQY205CG-SDEKRPAQY205CRAL 350
*:* ** *****:*****:*****:*****:*****

human_PIP5K1C      YSTAMESIQGGAARGEAIY205CESDDTMGGIPAVNGRGERLLLHIGIIDILQSYRFIKKLEHTW 413
zebrafish_Pip5k1ca YTTAMESIQGASACGEGIDY205CDDTMGGIPAVNGRGERLLLYIGIIDILQSYRLIKKLEHTW 418
zebrafish_Pip5k1cb YSTAMESIQGGAACGDSIDY205CDDTMGGIPAVSGKGERLLLFIGIIDILQSYRLIKKLEHTW 410
*:* ** *****:*****:*****:*****:*****

human_PIP5K1C      KALVHDGDTVSVHRPSFYAERFFKY205CFMSNTVFRKNSSLKSSPSKY205CKGRGGALLAVKPLGPTA 473
zebrafish_Pip5k1ca KALVHDGDTVSVHRPSFYADRFLRFMSNTVFRKTSSLKSSPSKY205CKRGRGLAV-GKYCGPGA 477
zebrafish_Pip5k1cb KALVHDGDTVSVHRPNFYADRFFRY205CMSANVFKKSTSLKASPAKY205CKGRSVLAV-PKFSVPGA 469
*****:*****:*****:*****:*****

human_PIP5K1C      AFSASQIPSEREEAQYDLRGARSYPTLEDEGRDLLPCTPPSFEEATTASY205CIATLSS-TS 532
zebrafish_Pip5k1ca AWSASQLPY205CMRDENIYDLRGARSFPTLEDDGRADVLY205CPCTPPSFEEATTASY205CIATLSSY205CTTS 537
zebrafish_Pip5k1cb AWSASQLPY205CSERDENIYDLRGARSFPVLENDG--LQY205CSCTPPSFEEATTASVATLSSY205CNY205CTS 526
*:* ** *****:*****:*****:*****:*****

human_PIP5K1C      LSIPERSPSETSEQPRYRRRTQSSGQDGRPQ--EPPAEEDLQY205CITVQVEPACS----- 584
zebrafish_Pip5k1ca LSIPERSPY205CSDTSEHPRYRRRTQSSHEETMQD-----EDQQTITVEVEVEGRYDSEPT 589
zebrafish_Pip5k1cb ISIPERSPY205CSDTSEHPRYRRRTLY205CSLKDAEGRTHEVVDVHEEDQY205CITVQVEVY205CKRETEEDED 586
*****:*****:*****:*****:*****

human_PIP5K1C      --V----EIVVPKEEDA-----GVEASPAGASAY205CAVEVETASQASDEEGAPA---- 624
zebrafish_Pip5k1ca LVA----PQVSPEISEA--AETIPEASSSSVPASPRIVVESDGGY205CSQAS--GCTSR---- 636
zebrafish_Pip5k1cb DRETY205CDASQY205CTHTDPETHIAAETEMVPPPPADSDPA--DDLETQTASSAPDDQENPPPADS 643
. *: * .: ** *: :.*

human_PIP5K1C      -----SQASDEEDAPATDIYFPTDERS 646
zebrafish_Pip5k1ca -----ASVDEEDVPIY205CTDIYFPPEDKS 658
zebrafish_Pip5k1cb EQQTAAY205CDPPEAESRSAESAEANPQTQLPPADPLSQSRESLEEELAAY205CNTDIYF*----- 698
. *: . *****

human_PIP5K1C      WVYSPLHYSAQAP--PASDGESDT* 669
zebrafish_Pip5k1ca WVYSPLHFSSGPKTLPEDEGESET* 683
zebrafish_Pip5k1cb -----

```

Figure 14: Amino acid alignment of the human PIP5K1C with the zebrafish Pip5k1ca and Pip5k1cb proteins. Clustal omega program was used on uniprot website to align the sequences. In red are indicated the human mutations found in patients.

As previously explained duplicated genes can also be expressed in separate tissues. We thus analyzed the expression pattern of both paralogs to select the one expressed in the developing central nervous system as the patients present neurological features and visual deficiencies. By *in situ* hybridization we could show that *pip5k1ca* and *pip5k1cb* display a similar pattern of expression. Indeed, both genes are expressed in the central nervous system and in the developing eyes at 1 dpf and in the brain at 2 dpf (Fig. 15).



**Figure 15:** *pip5k1ca* and *pip5k1cb* are expressed in the developing central nervous system in zebrafish. *In situ* hybridization for *pip5k1ca* and *pip5k1cb* at 1 dpf and 2 dpf larvae. Dorsal view for the 2 pictures on the right of the panel, lateral view for the 4 other pictures, all oriented with the anterior pole on the left. Scale bar = 50  $\mu$ m.

Fibroblasts derived from the patients has shown an increase of PI(4,5)P<sub>2</sub>, the final product of PIP5K1C, suggesting that the mutations found in the genome of the patients could act as activating mutations (Observations made by our collaborator Dr. Morleo and colleagues). Together with these results and the similar expression pattern of *pip5k1ca* and *pip5k1cb* that we observed, we hypothesized that targeting only one of these genes to analyze the effect of the mutation would be sufficient as it would result in a dominant and constitutively active form of the protein. We have thus designed different sgRNAs to target both genes independently and to select the most efficient sgRNA for the generation of the E146K mutation using the base editing technology.

In both loci, near to the C base target, two other Cs that could also be edited are present. The C located in 5' of the target C would give rise to a silent mutation if converted into a T base whereas the C present in 3' would give the unwanted R150K and R143K mutations for Pip5k1ca and Pip5k1cb respectively (Fig. 16.a, b). We thus

have designed different sgRNAs to shift the C base to target inside the editing window and to exclude the unwanted C located in 3' as we can use the CBE4max-SpRY working efficiently with NGN and NAN PAMs. After testing several sgRNAs, we found that the C12 and the C11 could also be edited in zebrafish using the near PAM-less CBE4max-SpRY (Fig. 16, *pip5k1ca* sgRNA 3). This result (conferring new possibilities, which could be useful for other cases) did not allow us to obtain any sgRNA exclusively editing the E151K for Pip5k1ca and E144K for Pip5k1cb (Fig. 16.c, d). Despite this finding, we have decided to target *pip5k1ca* because: 1) It is the closest one to the human *PIP5K1C* and 2) Using the sgRNA 2 we obtained a higher C-to-T conversion efficiency using the CBE4max-SpRY base editor, 7 out of 8 embryos were editing up to 69% (Fig. 16.c). It can be noticed that using the sgRNA 2 we obtained a higher conversion efficiency with CBE4max-SpRY than with the classical AncBE4max (Fig. 16.c) as we have previously observed when targeting the *tyrosinase* gene (*c.f.* Article 2). Moreover, for some embryos, we obtained a C-to-T conversion without inducing the R150K mutation or with a very low conversion rate compared to the use of sgRNA 1 with which we obtained more or less the same rate of efficiency for both E151K and R150K (Fig. 16.c, sgRNA 2 embryos 2 and 4).



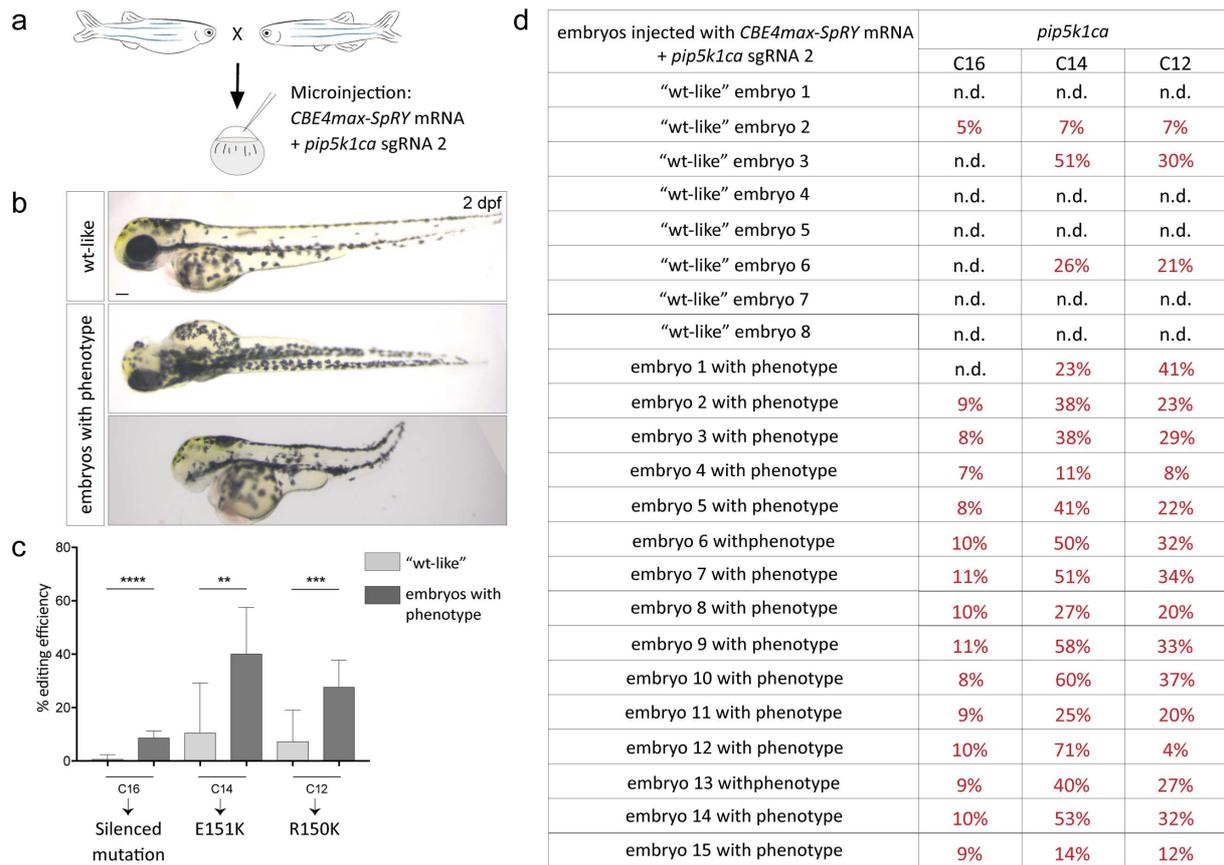
**Figure 16: Introduction of the human E146K mutation in Pip5k1ca and Pip5k1cb after C-to-T conversion using base editors. (a, b)** sgRNAs designed to target the C in yellow giving the E151K mutation in Pip5k1ca (a) and E144K mutation in Pip5k1cb after converted into a T base in the corresponding genes. The PAM sequences are

in green, the Cs that could be targeted in red with the position from the PAM mentioned above each Cs. The sequences of both loci are the reverse strand of the coding sequence. **(c, d)** C-to-T conversion efficiency estimated using EditR online software<sup>232</sup> of single embryos injected with each sgRNAs and *ancBE4max* or *CBE4max-SpRY* mRNA. n.d.: no detectable mutation by Sanger sequencing.

Even if we were unfortunately not able to obtain the perfect sgRNA giving exclusively the mutation corresponding to the human E146K, for the majority of the injected embryos we could show here that with the use of the CBE4max-SpRY variant we can design several sgRNAs targeting the same mutation and select the best one. This was not possible using the classical AncBE4max for which only one sgRNA could be designed (Fig. 16.a, b *pip5k1ca* sgRNA 1, *pip5k1cb* sgRNA 1.).

Upon the injection of the *CBE4max-SpRY* mRNA with the *pip5k1ca* sgRNA 2 (Fig. 17.a), we obtained 30% of the injected embryos showing microphthalmia and/or microcephaly, same phenotype obtained with the previous experiments using mRNA overexpression (Fig. 11, Fig. 17.b). We next randomly selected 8 single “wild-type like” and 15 single embryos showing these phenotypes to sequence their *pip5k1ca* locus. We obtained a statistically significant difference of C-to-T conversion efficiency between the two groups, the embryos with the phenotype were more edited than the wild-type like embryos (Fig. 17.c). Indeed, we obtained a C-to-T conversion for all the embryos showing the phenotype, up to 71% of efficiency. Among the “wild-type like” embryos, only 3 out of 8 embryos were edited (Fig. 17.d).

*Pip5k1cb* locus is a potential off-target of *pip5k1ca* sgRNA 2 with only two mismatches. We thus have sequenced by Sanger sequencing *pip5k1cb* in the 15 embryos with phenotype, no mutations have been found in all the sequenced embryos. This result is another example supporting the fidelity of the CBE4max-SpRY base editor.



**Figure 17: Mutations in *pip5k1ca* using base editor induce a similar phenotype to the one obtained by the human *PIP5K1C(E146K)* mRNA overexpression. (a)** Injection of *CBE4max-SpRY* mRNA and *pip5k1ca* sgRNA 2 in embryos at one cell-stage. **(b)** 30% of the injected embryos show microphthalmia and/or microcephaly. Scale bar= 50µm. **(c)** Mean of the C-to-T conversion efficiency for the two groups of injected embryos. Wilcoxon-Mann-Whitney statistical test, \*\*= P values ≤ 0,01, \*\*\*= P values ≤ 0,001, \*\*\*\*= P values ≤ 0,0001. **(d)** C-to-T conversion efficiency estimated using EditR online software<sup>232</sup> of single injected embryos. n.d.: no detectable mutation by Sanger sequencing.

## II.5. THE VERSATILE PRIME EDITING TECHNOLOGY IN ZEBRAFISH.

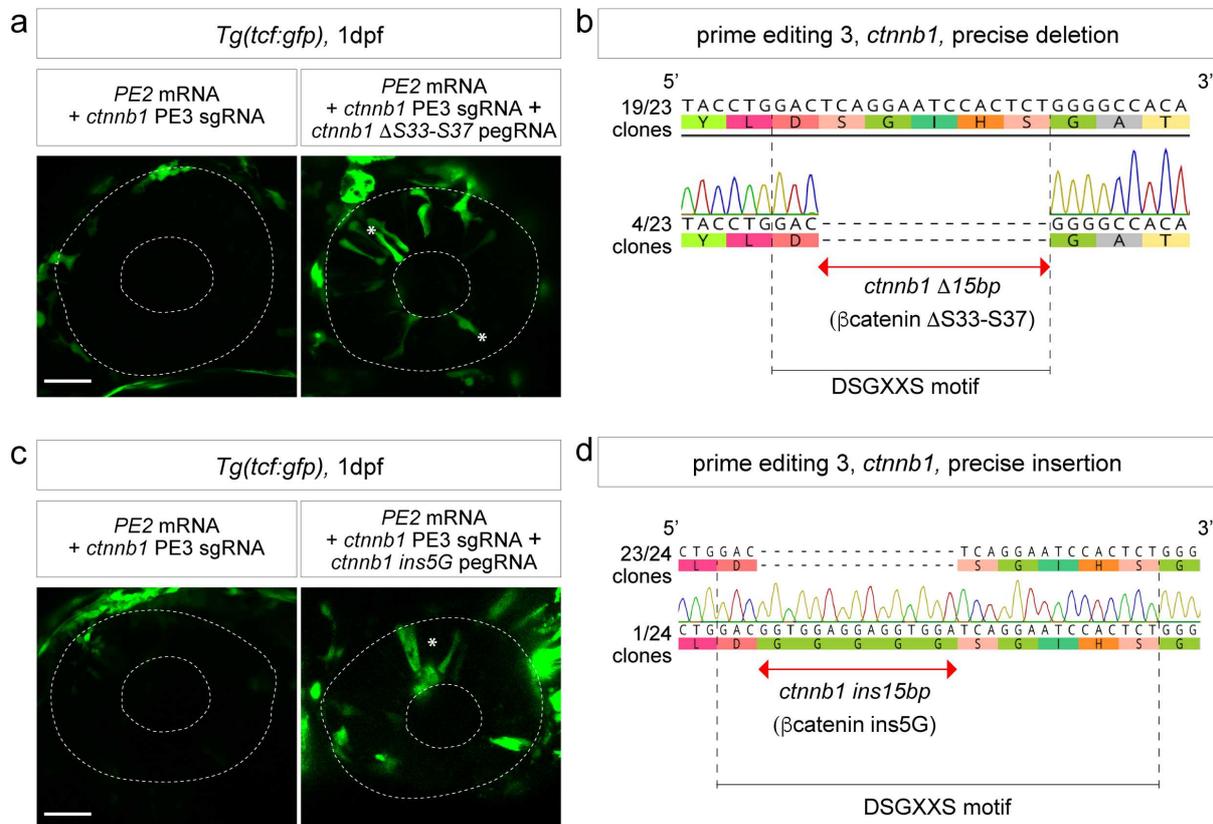
The emergence of base editors had a big impact for disease modeling and gene therapies, increasing drastically the precise genome editing possibilities. After several optimizations, this technology is now a tool of choice to perform precise transitions at high efficiency in many systems, including zebrafish. However, there are still some limitations concerning the editing window in which bystander edits can occur and the restrict conversion possibilities. In this context, prime editing technology is an elegant tool with high potential to perform small deletions, small insertions and all types of conversion when it is not feasible to use base editors. Although several proofs of concept have been reported in several models, this technology is still at its beginning and optimizations are needed in order to fine-tune its editing specificity and efficiency.

### II.5.1. ESTABLISHMENT OF THE TECHNOLOGY.

In order to expand the ability to induce genetic modifications in zebrafish we investigated the potential of prime editing. To do so, we first cloned the available PE2 sequence (addgene #132775) into a *pCS2+* plasmid in order to add a polyA during the *in vitro* mRNA synthesis, important for its stability. We then decided to start first with the PE3 strategy as it was reported to be the most efficient approach in cultured cells even though it could create a high unwanted mutation rate compared to PE2 strategy. Thus, the *PE2* mRNA, PE3 sgRNA and pegRNA were injected directly in one-cell stage zebrafish embryos. The pegRNA and sgRNA were synthetic RNAs ordered from IDT.

We first applied this method to constitutively activate  $\beta$ -catenin as we have previously done with the BE4-gam (*c.f.* Article 1). We used prime editing to precisely delete the sequence coding from S33 to S37 of the DSGXXS motif of  $\beta$ -catenin, required for the binding of the beta-TrCP ubiquitin ligase which triggers its degradation. In order to precisely generate this 15 bp deletion, we co-injected *PE2* mRNA with *ctnnb1*  $\Delta$ S33-S37 pegRNA, and PE3 sgRNA in one-cell stage *Tg(tcf:gfp)* embryos. Using this new editing strategy, an overall increase of GFP-positive cells and ectopic GFP-positive retinal progenitors at 1 dpf were observed, similarly to our previous results using the BE4-gam (Fig. 18.a). By sequencing clones after PCR made on injected embryos, we

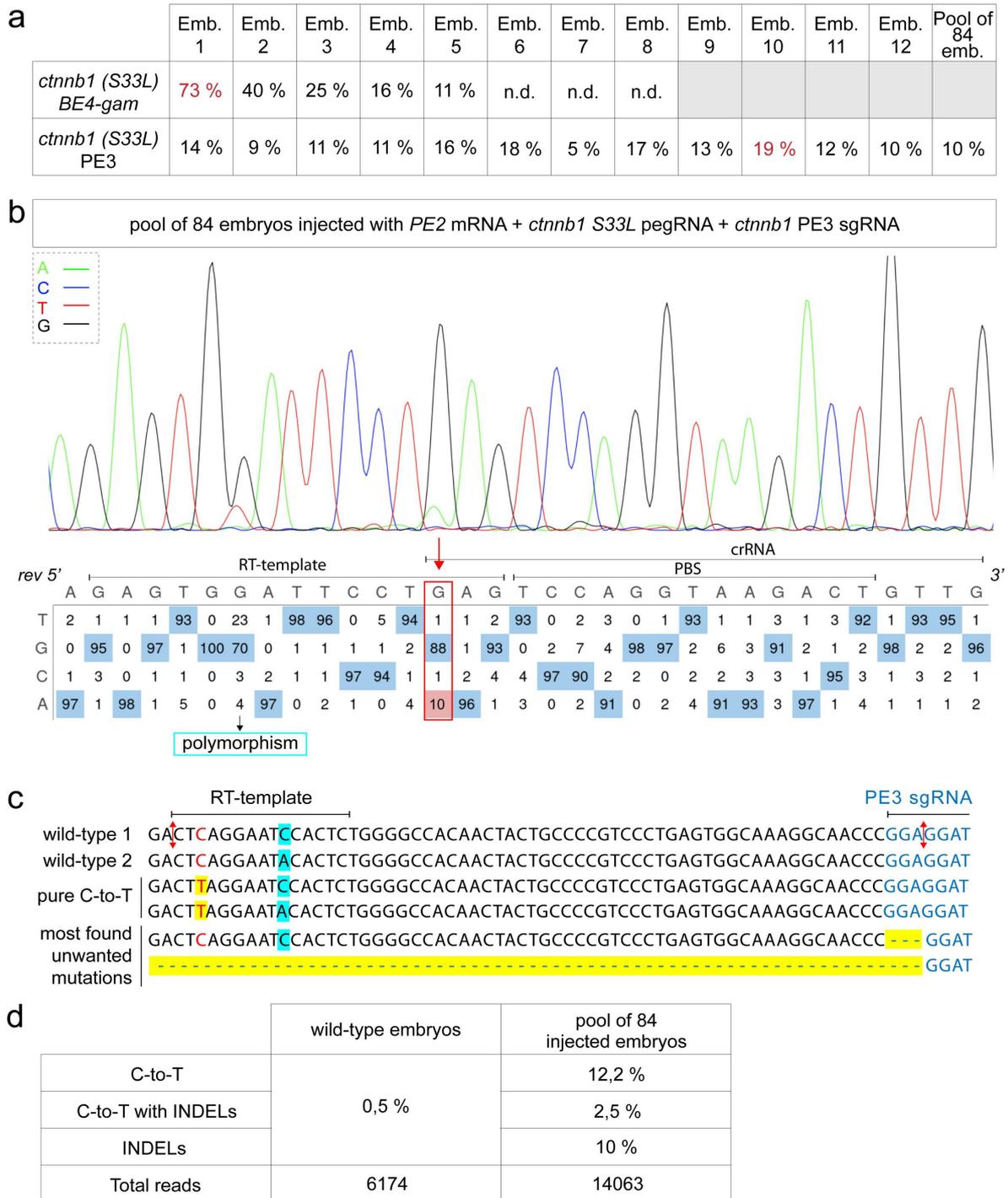
could demonstrate that the 15 bp deletion took place with up to 17% of efficiency (Fig. 18.b). We then investigated the capacity of prime editing to perform precise insertions. We aimed at introducing 15 bases coding for 5G into the DSGXXS motif of  $\beta$ -catenin, interfering with the motif and thus activating Wnt signaling pathway. We co-injected *PE2* mRNA with *ctnnb1 ins5G* pegRNA, and *ctnnb1* PE3 sgRNA in one-cell stage *Tg(tcf:gfp)* embryos and observed ectopic GFP-positive retinal progenitors at 1 dpf (Fig. 18.c). By sequencing single DNA clones after PCR, we were able to correlate this observation to the insertion of the 15 bp sequence. Although at less than 5% of efficiency, no other unwanted mutations were detected (Fig. 18.d).



**Figure 18: Prime editing technology in zebrafish for small insertion and deletion. (a, c)** Activation of Wnt signaling via  $\Delta$ S33-S37 deletion in (a) and ins5G insertion in (c) in  $\beta$ -catenin. 1 dpf *Tg(tcf:gfp)* embryo injected with *PE2* mRNA, *ctnnb1* PE3 sgRNA and in (a) *ctnnb1*  $\Delta$ S33-S37 pegRNA, in (c) *ctnnb1 ins5G* pegRNA. GFP-positive retinal progenitor cells (white stars) were found in the retina (inside the dash lines) which are normally GFP negative in *Tg(tcf:gfp)* embryos. Maximal z-projection. Scale bar= 50 $\mu$ m. **(b)** Sequencing of individual clones, precise 15 bp deletion was found in 4 out of 23 clones leading to the deletion of S33 to S37 of  $\beta$ -catenin. The other 19/23 clones were wild-type. **(d)** Sequencing of individual clones identified precise 15 bp insertion in 1 out of 24 clones leading to the insertion of 5G in  $\beta$ -catenin. The other 23/24 clones were wild-type.

Also targeting the *ctnnb1* gene, we then compared the efficiency of prime editing with the efficiency of base conversion obtained using BE4-gam. To do so, *PE2* mRNA, *ctnnb1 S33L* pegRNA, and *ctnnb1* PE3 sgRNA were co-injected at one-cell stage and 12 single embryos and a pool of 84 injected embryos were sequenced. C-to-T mutations were observed for all the analyzed embryos up to 19%, an editing efficiency lower than with BE4-gam (Fig. 19.a). Compared to the BE4-gam we did not observe bystander edits. However, 10% of INDELS were generated as revealed by NGS sequencing and 2,5% of the sequence contained the C-to-T conversion with INDELS (Fig. 19.c, d). The fish from our fish facility carried a polymorphism at the locus targeted by the RT-template sequence of the pegRNA as previously observed (Fig. 9.d, Fig. 19.b, c, in blue). The pegRNA used was designed to recognize the alleles containing the C base at that site (Fig. 19.c, wild-type 1). We could observe that 96% of the C-to-T sequences has the C base at the polymorphism site, suggesting that one single mismatch in the pegRNA sequence, even in the RT-template, could decrease the efficiency.

Moreover, to compare the PE2 and PE3 strategies, we injected the *PE2* mRNA and the *ctnnb1 S33L* pegRNA. No C-to-T conversions were observed in the injected embryos by Sanger sequencing using the PE2 strategy, meaning that in our approach the PE3 is more efficient than the PE2 system.

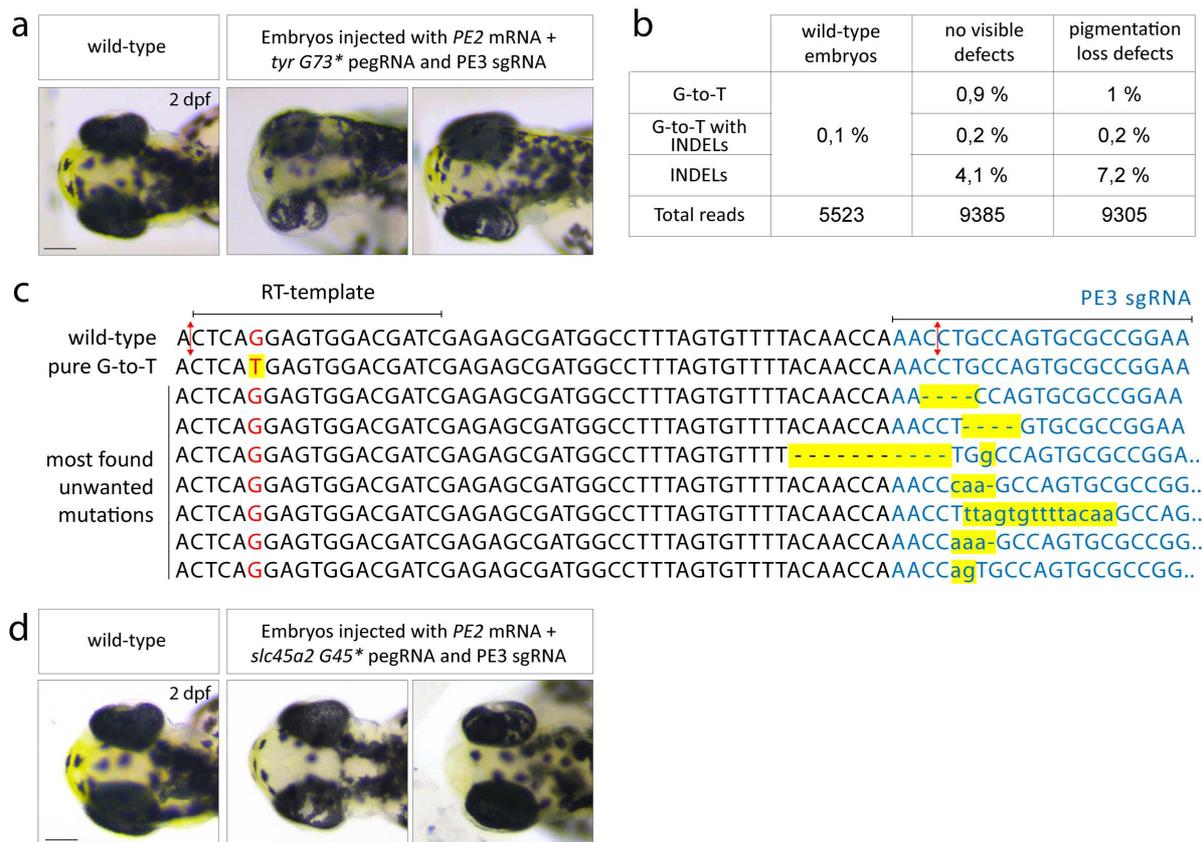


**Figure 19: Prime editing versus base editing for C-to-T conversion in zebrafish. (a)** C-to-T conversion efficiency obtained with BE4-gam<sup>141</sup> and using PE3 approach in order to introduce the S33L mutation in  $\beta$ -catenin. n.d.: non-detectable by Sanger sequencing. In red the highest efficiency obtained. **(b)** DNA sequencing chromatogram of the targeted *ctnnb1* gene from a pool of 84 injected embryos with the PE2 mRNA, the *ctnnb1* S33L pegRNA and *ctnnb1* PE3 sgRNA. S33L mutation in  $\beta$ -catenin upon C-to-T conversion in *ctnnb1* reached 10% of efficiency. Numbers in the boxes represent the percentage of each base at that sequence position. In red is highlighted the base substitutions introduced by prime editing while the original bases are in blue. The color code of the chromatogram is indicated in the upper left corner (Adenine green, Cytosine blue, Thymine red, Guanine black). **(c, d)** NGS sequencing made on the pool of the 84 injected embryos to induce S33L mutation in  $\beta$ -catenin using the PE3 approach. **(c)** Most sequences obtained are represented with highlighted in blue the polymorphism and in yellow the edits. In red the targeted C and in blue the region targeted by the PE3 sgRNA. The red arrows are

the nick sites. **(d)** Percentages of the wild-type sequence and sequences containing the pure C-to-T conversion or INDELS found by NGS sequencing.

We next assessed if we could target other genes and perform conversions that could not be done using BE in zebrafish. We selected the *tyrosinase* gene in order to create a premature stop codon by G-to-T mutations to have an easy read out as previously done with base editors (*c.f.* Chapter II.2). Four different strategies have been designed in order to generate the S50\*, W80\*, G73\* or the G93\* mutation in Tyrosinase using PE3 (Table 2). Only upon the injections to target the G73\*, we observed white patches in the eyes in 15%-20% of the injected embryos in 3 independent experiments (Fig. 20.a). However, no G-to-T conversions were found by Sanger sequencing. We thus performed NGS to analyze more precisely the targeted locus of 3 different pools: non-injected embryos, injected embryos without pigmentation defect and injected embryos with pigmentation loss in the eye. We could observe pure G-to-T conversions in both injected embryo pools, with or without apparent phenotype, and with a similar efficiency (0.9% and 1%) (Fig. 20.b, c). However, we also detected INDELS at a higher rate in the embryos with pigmentation loss defect. The phenotype observed could thus be also a result of INDEL integrations that, as the G-to-T conversion, create a nonsense mutation (Fig. 20.b, c).

In order to perturb the pigmentation formation, we also aimed at introducing a premature stop codon in the *slc45a2* gene by G-to-T transversion using the PE3 approach. Upon injections of *PE2* mRNA, *slc45a2* G45\* pegRNA and PE3 sgRNA, 30% of injected embryos showed pigmentation loss in the eyes at 2 dpf (Fig. 20.d) but no mutations were found by Sanger sequencing. As done for *tyrosinase*, NGS sequencing is required to reveal the frequencies of G-to-T conversion and possible unwanted mutation created upon this injection.



**Figure 20: Pigmentation loss in zebrafish embryos by G-to-T conversions using prime editing technology. (a)** Dorsal view of 2 dpf embryos showing pigmentation loss in the eye, embryos injected with *PE2* mRNA, *tyr G73\** pegRNA and *tyr G73\** PE3 sgRNA. Scale bar= 50µm. **(b, c)** NGS analysis made from 2 pools of injected embryos to induce G73\* mutation in Tyrosinase using the PE3 approach: one containing 8 embryos without phenotype and one with 2 embryos showing pigmentation loss defect, and a pool of uninjected wt embryos. **(b)** Percentages of the wild-type sequence and sequences containing the pure G-to-T conversion or INDELS found with NGS sequencing. **(c)** Most sequences obtained are represented with highlighted in yellow the edits. In red the targeted G and in blue the region targeted by the PE3 sgRNA. The red arrows are the nick sites. **(d)** Dorsal view of 2 dpf embryos showing pigmentation loss in the eye, embryos injected with *PE2* mRNA, *slc45a2 G45\** pegRNA and *slc45a2 G45\** PE3 sgRNA. Scale bar= 50µm.

All together our preliminary data demonstrate that this technology can be applied to zebrafish for conversion, small deletion and insertion with a poorly efficiency. We also show that prime editing is less efficient than base editing in zebrafish. In order to increase the efficiency, we are testing different length of the PBS and RT-template sequences of the pegRNAs used to target *tyrosinase* and *slc45a2* (Table 2). This will give more insight on the design of the guides which remains less straightforward than the one for sgRNAs.

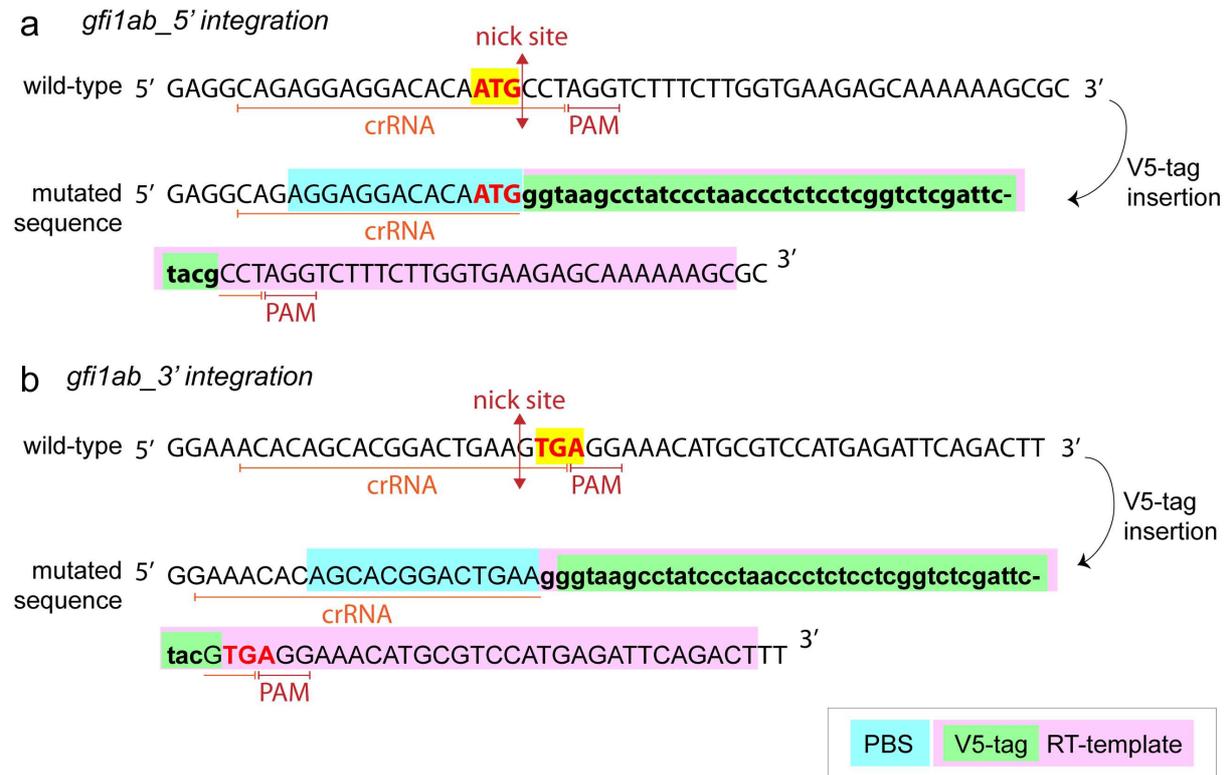
To broaden the scope of this tool we also cloned the SpRY domain from the *CBE4max-SpRY* coding sequence in the *PE2* sequence. As the Cas9-SpRY recognizes

the 5'-NGG-3' PAMs, we first tested the ability of PE2-SpRY to introduce the S33L mutation in  $\beta$ -catenin, being the most efficient and easiest strategy to verify the mutation by Sanger sequencing. Through three independent experiments, we unfortunately did not see any C-to-T conversion by Sanger sequencing.

In parallel, we had in view to develop prime editing in zebrafish in order to insert epitope tags. Indeed, an efficient method to insert a tag in zebrafish would be of great use for conducting cell biology and biochemical experiments as it is often challenging to find an efficient and specific antibody for this animal model. Among the commonly used tags, we chose to insert the V5-tag as it is a small tag consisting in the insertion of 42 bp. Anzalone *et al.* reported insertions up to 44 bp using prime editing in cultured cell. Moreover, the V5-tag was reported to be reliable for performing biochemical assays in zebrafish<sup>130, 233</sup>, but an efficient method to precisely insert it in a locus of choice is still a challenge to date.

We firstly attempted to insert the V5-tag sequence into the *ctnnb1* gene coding for  $\beta$ -catenin in the DSXXSG motif at the same locus targeted previously for insertion, deletion and transition. In this way, only the RT-template of the pegRNA was changed (Table 2). The pegRNA was cloned and synthesized *in vitro* as it was longer than 150 nt, the maximum length for which commercial companies can currently produce the pegRNA as synthetic and stable RNA. After several tests, we did not obtain any insertion by PCR made on injected embryos. We also tested the V5-tag insertion strategy in 5' and 3' of the *gfi1ab* gene. This gene presents two perfect loci to test this strategy thanks to the presence of a PAM at the right localization, leading to a nick site that allows the insertion of only 42 bp, downstream of the ATG start codon and upstream of the TGA stop codon (Fig. 21, Table 2). Also, for the 5' and 3' pegRNAs, we could design a PBS sequence of 14 nt with 50%GC and 13 nt with 54%GC respectively as advised by Anzalone *et al.*<sup>140</sup>. We synthesized the pegRNAs and injected it with the PE2 mRNA and/or the PE3 sgRNA. We did several experiments changing the concentration of each components, trying PE2 and PE3 systems and putting the injected embryos at 32°C instead of 28°C as reported recently to slightly improve efficiency<sup>234</sup>. We aimed at detecting the possible insertions were verified by 3 different PCRs: one amplifying the locus using primers outside of the V5-tag sequence that would give a bigger PCR product if the integration happened, and two with one of the two primers targeting the sequence coding for the V5-tag generating a PCR product only if the insertion occurred using these different strategies. No insertions were

observed. We plan to perform NGS sequencing to analyze the 3 different loci targeted to insert the sequence coding for the V5-tag. It would be interesting to repeat the same assays with synthetic pegRNAs when pegRNA longer than 150 nt will be available and to test several pegRNAs with different length of PBS and RT-template sequences.



**Figure 21: pegRNA design strategies for in-frame insertions of the V5-tag coding sequence in *gfi1ab* gene.** In lower cases the inserted sequence, in green the coding sequence for the V5-tag, in blue the PBS sequence, in pink the RT-template sequence containing the inserted sequence, in orange the crRNA part of the pegRNA. In red and yellow are highlighted, the ATG start codon and the TGA stop codon of the *gfi1ab* gene. The red arrow represents the nick site of the PE2.

pegRNA	Sequence (5'-3'): crRNA, RT-template (blue) with mutation (yellow,) PBS (red)	PBS		RT- tem- plate	Nick from PE3	Wanted edit
		Lenght	%GC			
<i>tyr_550*</i> pegRNA C-to-A conversion	<u>TCCGTGTGCGCGCTCCAGTCGTTTTAGAGCTAGAAATAGCAAGT</u> TAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGGAC CGAGTCGGTCCGAACCCTCGACCTTACTGAGACCCGCACA	13 nt	69%	16 nt	50 bp	No phenotype n.d. by Sanger
<i>tyr_W80*</i> pegRNA G-to-A conversion	<u>GAGTGGACGATCGAGAGCGAGTTTTAGAGCTAGAAATAGCAAGT</u> TAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGGAC CGAGTCGGTCCAACACTAAAGGCTATCGCTCTCGATCGTCCA	14 nt	57%	17 nt	46 bp	No phenotype n.d. by Sanger
<i>tyr_G73*</i> pegRNA G-to-T conversion	<u>GGGCCGAGTATCTCACTCGTTTTAGAGCTAGAAATAGCAAG</u> TAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGGAC CGAGTCGGTCCGATCGTCCACTATGAGTGAGGATACTGCG	13 nt	54%	17 nt	51 bp	Phenotype n.d. by Sanger up to 1% by NGS
<i>tyr_G93*</i> pegRNA G-to-T conversion	<u>AACCAAACCTGCCAGTGCCTGTTTTAGAGCTAGAAATAGCAAGTT</u> AAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGGACCGAGT CGTCCCATGTAGTTTCAGGCCACTGGCAGGTTT	13 nt	54%	17 nt	62 bp	No phenotype n.d. by Sanger
<i>slc45a_G45*</i> pegRNA G-to-T conversion	<u>ACGCCAGTGTGCTGAGCGTGTTTAGAGCTAGAAATAGCAAGTT</u> AAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGGACCGA GTCGGTCCGTCTGGGAAGTCAACGCTCAGCAACTG	13 nt	54%	17 nt	69 bp	Phenotype n.d. by Sanger
<i>gfi1ab_5'_V5</i> pegRNA 42 bp insertion	<u>CAGAGGAGGACACAATGCCTGTTTTAGAGCTAGAAATAGCAAGTTAA</u> AATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGGACCGAGTCG GTCCGCTTTTTGCTCTTCCACCAAGAAAGACCTAGCGTAGAATCGA GACCGAGGAGAGGGTTAGGGATAGGCTTACCATTGTGTCCTCCT	14 nt	50%	74 nt	47 bp (89 bp after edit)	n.d. by PCR
<i>gfi1ab_3'_V5</i> pegRNA 42 bp insertion	<u>ACACAGCACGACTGAAGTGGTTTTAGAGCTAGAAATAGCAAGTTAA</u> AATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGGACCGAGTCGGTC CAGTCTGAATCTCATGGACGATGTTTCTCACGTAGAATCGAGAC CGAGGAGAGGGTTAGGGATAGGCTTACCCTTCAGTCCGTGCT	13 nt	54%	74 nt	41 bp (83 bp after edit)	n.d. by PCR
<i>ctnnb1_V5</i> pegRNA 42 bp insertion	<u>CAACAGTCTTACCTGGACTCGTTTTAGAGCTAGAAATAGCAAGTTAA</u> AATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGGACCGAGTCGGT CCCAGTAGTTGTGGCCCCAGAGTGGATTCTCTGCTAGAAATCGAGA CCGAGGAGAGGGTTAGGGATAGGCTTACCCTCCAGGTAAGACT	13 nt	46%	74 nt	63 bp (105 bp after edit)	n.d. by PCR
<i>ctnnb1_delS33-S37</i> pegRNA 15 bp deletion	<u>CAACAGTCTTACCTGGACTCGTTTTAGAGCTAGAAATAGCAAGTTA</u> AATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGGACCGAGT CGGTCGAGTGGATTCTAAGTCCAGGTAAGACT	13 nt	46%	20 nt	63 bp (48 bp after edit)	4/23 clones
<i>ctnnb1_S33L</i> pegRNA C-to-T conversion	<u>CAACAGTCTTACCTGGACTCGTTTTAGAGCTAGAAATAGCAAGTTA</u> AATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGGACCGAG TCGGTCCGAGTGGATTCTAAGTCCAGGTAAGACT	13 nt	46%	15 nt	63 bp	10% by Sanger 12.5% by NGS
<i>ctnnb1_ins5G</i> pegRNA 15 bp insertion	<u>CAACAGTCTTACCTGGACTCGTTTTAGAGCTAGAAATAGCAAGTTA</u> AATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGGACCGAGTC GGTCCAGTTGTGGCCCCAGAGTGGATTCTGATCCACCTCCTC CACCCTCCAGGTAAGACT	13 nt	46%	43 nt	63 bp (78 bp after edit)	1/24 clones

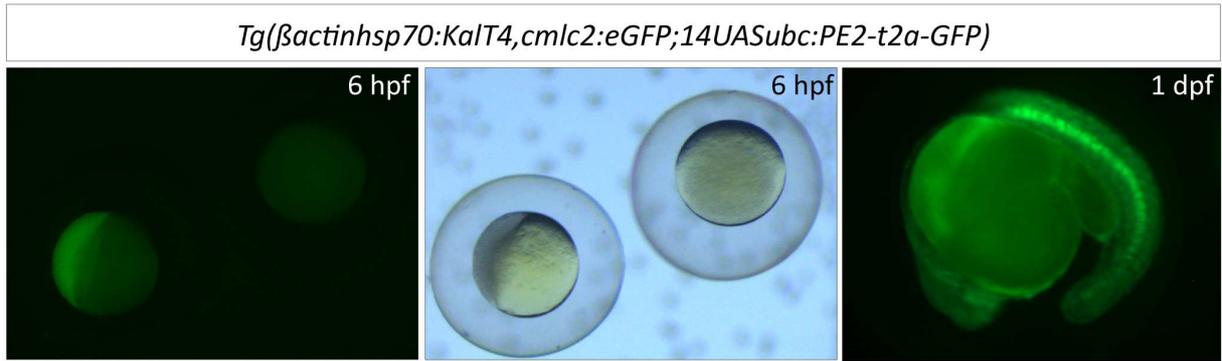
**Table 2: Design of the pegRNAs used in the study and the respective gene editing frequencies obtained.** In the pegRNA sequences, the sequence underlined is the crRNA, in blue is the RT-template containing the mutation in yellow and the PBS is represented in red. The nick from PE3 column show the distance between the nick generated with the pegRNA and the nick generated with the PE3 sgRNA. n.d. = non-detectable.

## II.5.2. MATERNAL CONTRIBUTION OF THE PRIME EDITOR.

In zebrafish, the maternal to zygotic transition occurs at the 10<sup>th</sup> cleavage cycle, embryos initially develop with RNAs and proteins that are generated during oogenesis and supplied by the egg of the mother before the activation of their own zygotic genome<sup>235</sup>. We thus hypothesized that a maternal contribution of the PE2 proteins could increase the gene editing efficiency compared to transient mRNA injections, as producing concentrated recombinant protein *in vitro* for injection is still challenging. This maternal supply could therefore potentially increase gene editing efficiency as it will not be translated from an injected mRNA but readily present in the embryos from the beginning of its development.

In order to get the PE2 protein expressions from one-cell stage of development, we sought of expressing them using a promoter activated during oogenesis. To do so, we generated the *Tg(βactinhsp70:kalt4,cmlc2:eGFP)* transgenic line ubiquitously expressing KalT4, a codon optimized Gal4, from the one-cell stage and having the GFP positive heart transgenesis marker (*cmlc2:eGFP*). Indeed, the expression is driven by the ubiquitous promoter *βactin* with the addition of a small part of the *hsp70* promoter that was reported to boost gene expression<sup>236</sup>.

We then engineered constructs in which the *PE2* coding sequence was cloned downstream of a *14UASubc* enhancer cassette. In this plasmid we also added the *t2a* and *GFP* sequences to have a fluorescent reporter allowing to follow PE2 protein synthesis. The construct was injected in embryos obtained from the cross of the stable *Tg(βactinhsp70:KalT4,cmlc2:eGFP)* transgenic line, and injected embryos were then raised. Upon PE expression, we did not observe any toxicity signs during the growth of the injected embryos. Similar observations were done in *Drosophila* when prime editor was ubiquitously overexpressed in the animals<sup>183</sup>. Adult fish were screened and a founder expressing the GFP was identified. After two generations we obtained transgenic *Tg(βactinhsp70:KalT4,cmlc2:eGFP;14UASubc:PE2-t2a-GFP)* adult females providing GFP-positive eggs, indicating that they maternally express the PE protein (Fig. 22). The injection of pegRNAs and sgRNAs into these eggs will allow to evaluate if gene editing can be achieved using this line. The efficacy of gene editing with those obtained using micro-injections of *PE2* mRNA will then be compared.



**Figure 22: Maternal expression of prime editor proteins.** *Tg( $\beta$ actin<sup>hsp70</sup>:Kalt4,cmlc2:eGFP;14UASubc:PE2-t2 $\alpha$ -GFP)* expresses prime editor protein early and ubiquitously during the embryonic development.

# CHAPTER III: DISCUSSION

## III.1. DISEASE MODELING IN ZEBRAFISH USING BASE EDITORS.

Together with other reports, our work provides a large panel of examples whereby, using the base editing technology, it is now possible to manipulate endogenously signaling pathways, to generate models for human genetic disorders and to introduce precise patient-associated mutations in the zebrafish genome.

By performing a loss-of-function mutation for 7 tumor suppressor genes (*tp53*, *rb1*, *tek*, *bap1*, *cbl*, *dmd* and *sod2*) and a gain-of-function mutation for 3 oncogenes (*kras*, *nras* and *ctnmb1*), we targeted 10 cancer genes in zebrafish using base editors (*c.f.* Article 1)<sup>141</sup>. These new possibilities for generating point mutations at high efficiency in cancer genes can complement current studies for which the main strategy used so far was the overexpression of human mutated oncogenes using tissue-specific promoters<sup>37</sup>. However, the use of these approaches is not optimal as the endogenous regulator sequences driving and controlling the cancer gene expressions are not kept. This is particularly problematic for some experiments like, among others, drug screening for phenotype rescue. Moreover, the occurrence of cancer diseases is generally caused by a combination of mutations occurring in several genes and can now be modeled using the new CBE4max-SpRY which we here demonstrated to be efficient to perform multiplex base editing in zebrafish using NRN PAMs. This tool now allows the generation of cancer models rapidly and easily.

Among our targets, the LOF mutation introduced in the *cbl* tumor suppressor gene allowed us to generate a new zebrafish model for dwarfism, consistently with the reports that found mutations in *CBL* in Noonan syndrome patients notably presenting a short stature and bone malformations<sup>237</sup>. It has been reported that hyperactivation of the fibroblast growth factor receptor 3 (FGFR3) leads to dwarf conditions. FGFR3 triggers intracellular signaling within chondrocytes of the growth plate which terminates its proliferation and bone growth<sup>238</sup>. Interestingly, an *in vitro* study

reported that some of these activating mutations in *FGFR3* disrupt c-Cbl-mediated ubiquitination that serves as a targeting signal for lysosomal degradation and termination of receptor signaling<sup>239</sup>. The dwarf phenotype was observed in 24% of the *MZcbl*<sup>-/-</sup> adult fish and by phase contrast X-ray imaging analyses of the skeleton of these fish and we could observe a smaller skull, a shorter spine and shorter precaudal vertebra ribs in these fish (Fig. 8). These results are in line with the growth defect phenotypes observed in the patients and directly implicate Cbl LOF as a cause for bone malformations in an animal model. A further characterization of these phenotypes as well as rescue experiments are required to fully establish the link between these mutations and the etiology of the disease.

Using the CBE4max-SpRY, we also combined the generation of a GOF mutation for *nras* oncogene with a LOF mutation for *tp53* tumor suppressor gene, revealing an increase of melanocytes in 3 dpf larvae (Fig. 5 in Article 2)<sup>214</sup>, a clear melanoma predisposing phenotype. Indeed, it has been reported that fish overexpressing human mutated *HRAS* oncogene in melanocytes were hyperpigmented at 3 dpf and developed melanoma at the adult stage<sup>240</sup>. The hyperpigmented larvae that we analyzed were edited for both genes and the phenotype was not observed upon single mutation of each gene. In order to validate the effects of the mutation at the protein level, it will be important to analyze the level of expression of the genes downstream of Tp53 such as *p21* and *mdm2*<sup>240</sup> which we expect to decrease upon *tp53* KO as the transcriptional factor Tp53 activates these genes. In addition, to test if the MAP kinase pathway downstream of Ras signaling is activated, a western blot analysis could be performed targeting the phosphorylated ERK1/2<sup>240</sup> and in this case, an increase of phosphorylated ERK should be observed if Nras is activated. Furthermore, it will also be interesting to analyze the fish that survive at later stages as they may develop tumors such as melanoma or present phenotypes linked to RASopathy disorders<sup>241</sup>. Indeed, RASopathy disorders is a group of neurodevelopmental pathologies caused by *de novo* germline mutations in *ras* genes, like Noonan, cardiofaciocutaneous, Costello and LEOPARD syndromes and neurofibromatosis-1. RASopathies are characterized by several phenotypes such as skin spots, cardiac defects and cancer predisposition<sup>242, 243</sup>. As such, our generated fish provide an important resource to help understand aspects of the pathophysiology of these disorders.

Using the same tool, we highlighted a causal link between a poly-malformative phenotype and a genetic mutation found in patients with developmental delay,

neurological defects and varying dysmorphic features by generating the corresponding human E146K mutation endogenously in *pip5k1ca* gene of the zebrafish. The same phenotype as observed with the mRNA overexpression of the mutated human variants was obtained (Fig. 11 and Fig. 17). It can be noted that to generate the Y205C human mutation into the zebrafish genome, the ABE base editor that has the ability to convert A to G bases could be used to model the Y221C human mutation (T**A**T (Y) to T**G**T (C)). For the third mutation, one possibility would be to generate the F221C in zebrafish using CBE which would be the closest to the human Y221C. Unfortunately, despite our different designs we were not able to prevent the conversion of the C giving the R150K mutation in *Pip5k1ca* after C-to-T conversion. Nevertheless, using the near-PAMless CBE and the NAN PAM, this unwanted mutation was not the predominant one identified in the embryos showing a phenotype (Fig. 17), even if we cannot fully exclude to this date that the phenotype observed is not also linked to this unwanted mutation. To address this issue, we intend in the future to express the human *PIP5K1C(R145K)* mRNA to verify if this mutation can potentially give rise to embryos with microphthalmia and/or microcephaly phenotypes or to design different sgRNAs in order to generate the R150K mutation in the zebrafish *Pip5k1ca* alone and to assess its effect.

In the same line of research, other fish models have provided many advantages for disease modeling, for which researchers are currently also developing genome editing tools<sup>244, 245</sup>. For instance, recently, CBE and ABE technologies were established in medaka<sup>246, 247</sup>. Together with the technological advances made on zebrafish, these important advancements bring fish models in the forefront of genetic studies and as an even more attractive system for modeling human genetic diseases. However, although base editing is thus an extremely powerful approach some limitations of the technique need to be considered.

## III.2. BASE EDITORS IN ZEBRAFISH: A POWERFUL APPROACH WITH LIMITATIONS.

### III.2.1. THE POWER AND LIMITATIONS OF BASE EDITORS.

Our results demonstrate that the CBE technology is highly efficient in zebrafish to generate precise C:G to T:A mutations. We could show the specificity of BE4-gam variant in this animal model by NGS analyses. Indeed, we obtained an undesired edit frequency similar to the one of the non-injected embryos, indicating that the unwanted edit frequency may be due to background sequencing mistakes from deep sequencing and that for the 4 analyzed loci using BE4-gam only, specifically C-to-T mutations were generated (Fig. 9). Extended analysis for the genes targeted by the AncBE4max and the CBE4max-SpRY still has to be done in order to validate their specificity as well, which so far appears to be CBE variant and locus -dependent in previously reported studies by cloning analysis. Indeed, using the AncBE4max, Carrington *et al.* have reported 4% of incorrect editing for one locus of F1 zebrafish embryos and found more incorrect edits using the BE4max<sup>248</sup>. Another study reports the use of base editor in zebrafish by the generation and the use of an AncBE4max codon optimized for zebrafish, the zAncBE4max<sup>159</sup>. No unwanted mutations were found in 5 out of the 6 target genes using the zAncBE4max and for the 6<sup>th</sup> locus a low frequency of INDELS (3.2%) and high C-to-T conversion (42%) were obtained<sup>159</sup>, showing that the zAncBE4max is more reliable than BE3. However, a direct comparison of the zAncBE4max with the AncBE4max has not been performed to assess the improvement of the tool by codon optimization.

The development and the use of different or modified Cas9 showing other specificity than the classical 5'-NGG-3' PAM have drastically increased the base editing possibilities, especially to introduce pathogenic mutation for which there is only one locus that can be targeted and often without 5'-NGG-3' PAM at the correct distance of the C base. We compared 4 different CBE4: the *Spymac*-AncBE4max (5'-NAA-3'), the xCBE4 (5'-NGN-3'), the CBE4max-SpG (5'-NGN-3') and the near-PAMless CBE4max-SpRY. No base edits were found using the xCBE4. *Spymac*-AncBE4max and CBE4max-SpG were working at a low efficiency compared to the most flexible CBE,

the CBE4max-SpRY, which allowed us to reach up to 100% of base editing in two different loci using 5'-NAN-3' PAM, *nras* and *tyrosinase* genes. Across all the targeted genes that we analyzed, we found that the CBE4max-SpRY was effectively more efficient with 5'-NRN-3' than 5'-NYN-3', an observation that is in line with what was previously reported in cultured cells<sup>212</sup>. In addition, we found that this tool can be even more efficient than the AncBE4max. Indeed, through the generation of a LOF mutation in *tyrosinase* using the *tyrW273\* NGG* sgRNA, 98,6% of the injected embryos showed pigmentation loss with the use of CBE4max-SpRY whereas with the AncBE4max, the proportion of depigmented embryos reached 57,9% (*c.f.* Fig. 2 of the Article 2). Our work is the first using the CBE4max-SpRY in an animal model. More recently, a preprint study reported the first use in animals of the SpRY-Cas9 and SpG-Cas9 to generate KO alleles by DSBs in zebrafish and HDR-based KIs in *C. elegans*<sup>249</sup>. In contrast to our results that are based on the use of BE versions, they found a higher gene editing efficiency for the SpG-Cas9 than the SpRY-Cas9 and a lower efficiency than the classical SpCas9. These results highlight that the efficiency of the endonuclease Cas9 may not be comparable to the one of the base editors using the same Cas9. This observation was also made in cultured cells by comparing the Cas9 with ABE and CBE<sup>250</sup>.

The power of the use of the CBE4max-SpRY is not restricted to the case where there is no available 5'-NGG-3' PAM. Indeed, using this tool with or without the presence of the classical PAM, the design of several sgRNAs can be done. This is an important advancement as for some targets it allows to select the most efficient sgRNA and to exclude the bystander C bases present in the editing window. In our work, by targeting *pip5k1ca* gene with CBE4max-SpRY, we could indeed design up to 4 sgRNAs in order to introduce the E151K mutation, to test their efficiency and to circumvent C bystander bases (Fig. 16). We could show that in some cases the C positioned 11 or 12 bases away from the PAM can still be converted. This last observation can be useful in DNA region poorly enriched by G or A bases and where the only NRN PAM available is at this distance from the C base target.

The PAM flexibility of this CBE variant allowed us to perform for the first time in zebrafish multiplex base editing, an important progress to model multigenic disorders (as shown in this study by targeting *nras* and *tp53* genes) but also to target two paralogues at the same time. We could indeed target up to 3 genes simultaneously using different PAMs and we developed a co-selection base editing approach to

prescreen injected embryos based on the detection of pigmentation loss. In our study, we developed this strategy by targeting the *tyrosinase* gene but we could also perform the same approach with the generation of a mutation in another gene such as *slc45a*, a gene also necessary for the pigment synthesis. These time-saving strategies, never developed before in zebrafish using CRISPR/Cas9-based approaches, allow to prescreen for the efficiency of several sgRNAs on one hand and, on the other hand, to raise only the most edited embryos. This aspect is particularly important to reduce the number of animals used, to save time and effort for the founder Fo identification carrying the desired mutations and to decrease overall costs.

These latest CBE variants are working with high efficiency in zebrafish and are the tools of choice to perform a precise C-to-T or G-to-A conversion into the genome, especially with the new CBE4max-SpRY which broaden the editing possibilities. However, many other types of conversion are needed to be introduced in the genome to mimic or rescue pathogenic mutations. In this line, the improvement of ABE or the new C-to-G base editors would have a great impact in the community, as well as the development of prime editing in this model which will be discussed in the next section (*c.f.* CHAPTER III.3.).

An important aspect to considered using BE technology is the off-target generation. zABE7.10 did not show any DNA off-target in zebrafish<sup>170</sup> but no studies have been performed to analyze the DNA off-targets of CBEs in this animal model, particularly important for disease modeling. In our studies, we observed with the presence of a SNP in *ctnnb1* gene that one single mismatch in the sgRNA sequence could drastically decrease the base editing efficiency, highlighting the robustness of this technique with respect to DNA off target generation (Fig. 9.d). Using CRISPOR webtool<sup>108</sup>, we found that all the sgRNAs used in our studies for disease modeling did not present DNA off-targets with less than two mismatches. *Pip5k1cb* locus is for example a potential off-target of *pip5k1ca* sgRNA 2 with only two mismatches and by Sanger sequencing of *pip5k1cb* in the 15 embryos showing a phenotype and edited for *pip5k1ca*, no mutations were found in *pip5k1cb*. An analysis of *tp53*, *nras* and *tyrosinase* off-target would further support the specificity of CBE in zebrafish on this matter. If off-targets with less than two mismatches are present for the locus of interest, the strategy that we developed using the CBE4max-SpRY that provides the ability to design several sgRNAs to target the same mutation would also help to avoid any off-target base editing production.

Finally, one last important concern is the RNA off-target generation. Indeed, it has been reported that the deaminase activity of CBE can convert the C into U in RNAs present in the cells that it can bind independently of the Cas9(D10A) specificity<sup>251</sup>. RNA off-targets should not be a problem in zebrafish when the aim is to create a stable mutant line as the CBE is provided by mRNA injections and is no longer expressed anymore after the first days of development. In these studies, the hyperpigmented phenotype by targeting *nras* and *tp53* and the microphthalmia and/or microcephaly phenotypes through *pip5k1ca* gene targeting were observed in FO injected larvae. At this stage, we cannot fully exclude an RNA off-target effect on these phenotypes in addition to those induced by specific base editing. However, these phenotypes were not observed through the injections of the *CBE* mRNA only neither with other sgRNAs, supporting the specificity of our results. Moreover, the hyperpigmentation phenotype was not observed in larvae base edited exclusively for *nras* or *tp53* gene. Although RNA off-targets may not be an issue if controls are done properly, it may be a real concern for conditional experiments as will be discussed in the next section.

### III.2.2. ADDRESSING RNA OFF-TARGETS CONCERN FOR CONDITIONAL BASE EDITING.

Up to 100% of base editing efficiency in FO zebrafish embryos was reached using the CBE4max-SpRY<sup>214</sup>. This method is of great use to mimic and study *in vivo* any genetic disease of choice in zebrafish where precise genetic alterations need to be introduced into the genome. Nevertheless, if the gene is implicated in early embryonic development or in several tissues, the phenotype of interest can be masked by others that are stronger, broader or lethal and as a consequence the embryos might never reach the age where the tissue of interest forms. For these reasons, establishing a tissue-specific base editing approach is required to address this issue.

In this direction and using the BE4-gam and the bipartite and versatile Gal4/UAS system, we aimed at establishing a precise and tissue-specific base editing method in zebrafish. In this system, the Gal4 transcription factor is expressed under a specific promoter and can in turn drive the expression of CBEs cloned downstream of a *UAS* enhancer cassette to induce CBE expression in a tissue-specific manner. We first developed a *Tg(rx2:gal4;UAS:BE4-gam-t2a-GFP)* transgenic line expressing the

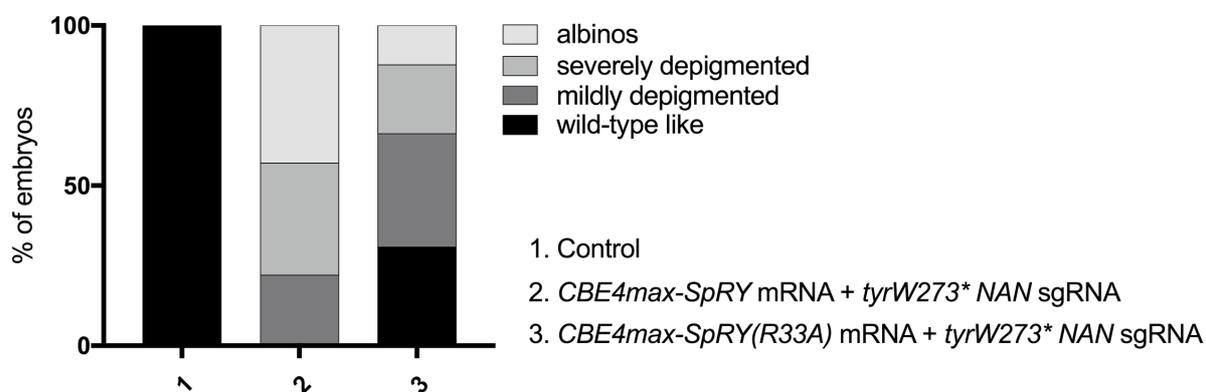
BE4-gam specifically in the retinal progenitor cells. To avoid any steric problem that could interfere with the activity of the CBE and to monitor its expression, a self-cleaving t2a peptide followed by a GFP cassette were added, resulting in the production of two distinct proteins, the BE4-gam and the GFP. However, we observed a delay in the retinal pigmented epithelium formation in the larvae and some of the adult fish developed overgrowth in the eyes, suggesting a toxicity effect of the BE4-gam upon its overexpression through the Gal4/UAS system (Fig. 23).



**Figure 23: Fish overexpressing the BE4-gam using the UAS/Gal4 system display toxicity effect.** Scale bar: 50 $\mu$ m for the 2 dpf larvae and 2mm for the 5 mpf fish. The green heart is the transgenesis marker for the *rx2:gal4* transgene and the GFP positive lens for the *UAS:BE4-gam\_t2a\_GFP*.

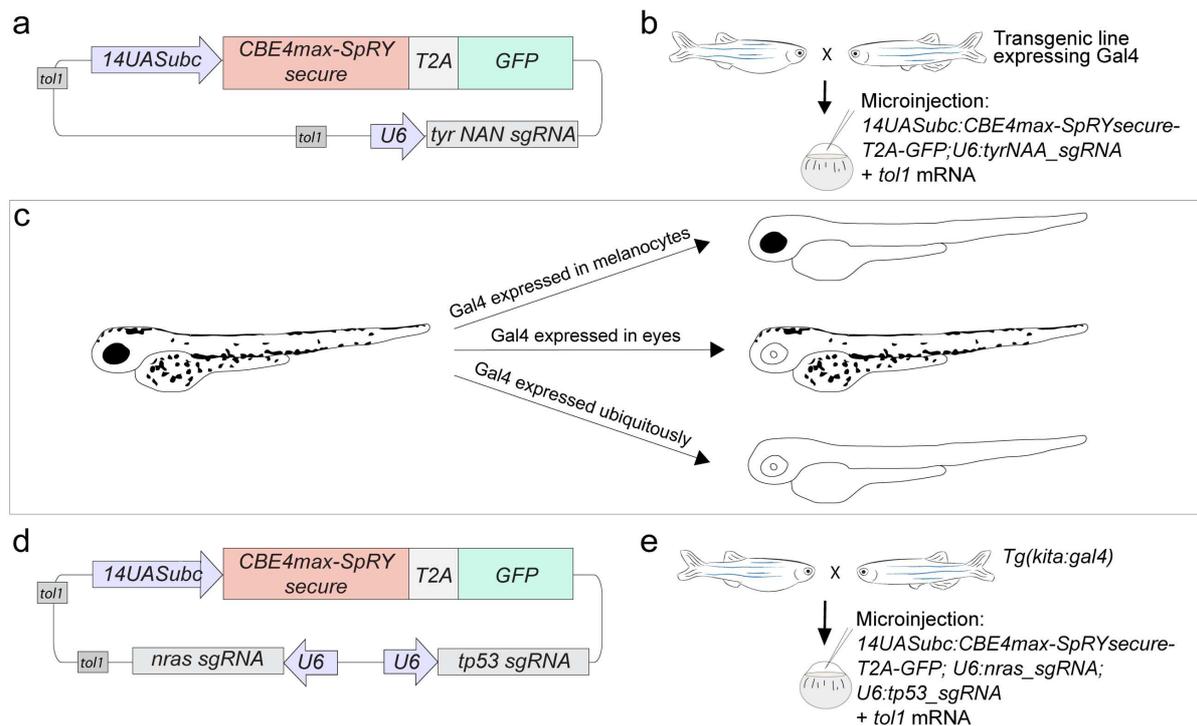
This high toxicity obtained by the overexpression of the BE4-gam might be due to the reported CBE RNA off-target effect<sup>251</sup>. It has been shown that this RNA off-target effect can be almost completely abolished, with a slight decrease of base editing efficiency on DNA on-target, by introducing point mutations on the deaminase, an optimization leading to the production of a “secure CBEs”<sup>251</sup>. Among them, R33A, R33AK34A and E63Q mutations were reported and we introduced them in the latest AncBE4max variant. In parallel, these 3 constructs were tested in cultured cells and efficient base editing was obtained only for the AncBE4max(R33A) (*in collaboration with P. Antoniou and Dr. Miccio, Imagine Institute, Paris*). We thus introduced this mutation in the CBE4max-SpRY and the synthesized *CBE4max-SpRY(R33A)* mRNA was injected with the *tyr W273\* NAN* sgRNA (*c.f.* Article 2) into one-cell stage embryos in order to validate its efficacy in zebrafish (Fig. 24). Upon injections (n=65), embryos showing depigmentation were observed in n=45, and the specific C-to-T conversions were detected by Sanger sequencing (Fig. 24). It is worth noticing that the base editing

efficiency obtained was lower than the one reached with the regular CBE4max-SpRY (Fig. 24).



**Figure 24: CBE4max-SpRY(R33A) generates C-to-T conversions at a lower efficiency than CBE4max-SpRY in zebrafish.** Proportion of the 4 groups based on the pigmentation defects described in Fig. 2c of article 2<sup>214</sup> for each injection: the CBE4max-SpRY mRNA and the tyrW273\* NAN sgRNA (column 2, 28 embryos in total) and the CBE4max-SpRY(R33A) mRNA and the tyrW273\* NAN sgRNA (column 3, 65 embryos in total).

This new secure CBE4max-SpRY(R33A) could now be implemented in the UAS/GAL4 system as described above in order to establish a tissue-specific base editing method that could potentially prevent from these toxicity effects. As a first functional proof of concept, we propose a strategy where the *tyrosinase* gene could be targeted using the *tyr 273\* NAN* sgRNA to have a clear phenotypic read out of the base modification with the loss of pigment in the biallelic mutated cells. As the stability of commercial sgRNA is not well-known and is likely to decrease overtime, the sgRNA would need to be ubiquitously and stably expressed using the *U6* promoter in zebrafish. Cloning the *14UASubc* enhancer cassette downstream of the CBE4max-SpRY would boost the expression of the CBE as the intronic *ubc* cassette has been reported as a remarkable enhancer to boost gene expression<sup>252</sup>. The availability of a large number of Gal4 expressing transgenic lines would allow to test this approach in any tissue of interest. For instance, it would be possible to specifically edit *tyrosinase* using the *Tg(kita:gal4)* transgenic line in melanocytes or in eyes using *Tg(rx2:gal4)* or even ubiquitously using *Tg( $\beta$ actinhsp70:kalt4,cmlc2:GFP)* (Fig. 25.b, c). Furthermore, in the context of cancer disease studies, this method could be applied to introduce somatic mutations in cancer genes and evaluate their oncogenic potential. With these tools, *nras* and *tp53* could be targeted specifically in melanocytes using the *Tg(kita:gal4)* transgenic line (Fig. 25.d, e).



**Figure 25: Establishing the tissue specific base editing approach in zebrafish using the Gal4/UAS system. (a, d)** Plasmid construct containing a *CBE4max-SpRY(R33A)\_t2a\_GFP* cassette under the *14UASubc* enhancer. **(b)** Injection at one cell stage embryos of the *14UASubc:CBE4max* plasmid and *tol1* mRNA to perform transgenesis. The sgRNA is expressed ubiquitously through the *U6* promoter activity and targets the *tyrosinase* locus. The designed strategy introduces a precise mutation that generates a premature stop codon. Using stable transgenic lines expressing Gal4 in specific tissue of interest, the injected UAS construct will only be activated in these tissues. **(c)** Schematic representation of the expected tissue-specific effect in pigment loss at 2 dpf in injected larvae depending on the transgenic line used. **(d, e)** Injection at one cell-stage embryos of the *14UASubc:CBE4max* plasmid and *tol1* mRNA in order to perform transgenesis. The sgRNAs are expressed ubiquitously through the *U6* promoter and target the *nras* and *tp53* loci to generate respectively a gain-of-function and a loss-of-function mutations. In the *Tg(kita:gal4)* transgenic line, the Gal4 is expressed specifically in melanocytes.

The toxicity observed by overexpressing the BE4-gam was also obtained when we aimed at developing a maternal transgenic line to express the BE4-gam protein early in the eggs using the *Tg( $\beta$ actinhsp70:KalT4,cmlc2:eGFP)*. A high mortality was indeed obtained and no founders were identified screening  $\pm 70$  adult fish. The *CBE4max-SpRY(R33A)* that we tested in zebrafish might help to overcome this issue and allow to create a maternal line. Transgenic lines with a more restricted maternal promoter expression than  *$\beta$ actin* to drive the Gal4 could also be used.

### III.3. PRIME EDITING IN ZEBRAFISH.

#### III.3.1. A NEED OF OPTIMIZATION TO ROUTINELY USE PRIME EDITING IN ZEBRAFISH.

All together our preliminary data shows that the prime editing technology can be applied to zebrafish for conversion, small deletion and insertion although with poor efficiency. In addition, our results indicate that PE2 is less efficient than BE and should only be chosen to create mutations that cannot be induced using BEs in zebrafish. In order to increase the efficiency, the length of the PBS and RT-template sequences of the pegRNAs used to target *tyrosinase* and *slc45a2* should be tested as both targets give rise to clear read outs for gene editing. This would give more insights on the design of the guides, a task that is less straightforward than for sgRNAs and that appears to be locus-dependent (Table 2).

Focusing on *tyrosinase* and *slc45a* genes would also allow to develop a prime editing co-selection method that would help to select the most edited embryos based on the depigmentation phenotype, a method that we developed here with the CBE technology.

Having prime editing working efficiently in zebrafish would not only be powerful for disease modeling but also to develop conditional knock-outs and biochemical studies in this animal model. Indeed, *loxP* sites used for conditional purposes are only 34 bp of size and the sequence coding for the V5-tag previously mentioned in this manuscript is 42 bp long. These two short sequences could in theory be integrated using PE2 technology as Anzalone *et al.* reported insertions up to 44 bp using this technique in cultured cell<sup>140</sup>.

Recently, prime editing has been performed in zebrafish by micro-injecting the prime editor protein and homemade pegRNAs. Germline transmission for two substitutions leading to Tyr(P302L) and Kras(G12V) mutations generation were reported. This work supports that prime editing as a very promising genetic tool in zebrafish. Nevertheless, even if this study uses PE protein for the first time, a very low gene editing efficiency was obtained. Targeted substitutions in 8 different genes were observed but with less than 6% of pure edits (wanted mutation without errors) in some cases and for most with less than 2% (6 out of the 8 genes). For deletions, the mean

pure gene editing frequencies obtained was ranging from 4.13% to 33.61% and for insertions from just 0.10% to 18.00%. In the majority of the cases, the obtained frequencies of INDELS were higher than the desired mutation frequencies. Indeed, for the insertion of 9 bp in *adgrf3b* locus for instance, at least 5 times more INDELS than the wanted 9 bp insertion were generated. Interestingly in this study, the injected embryos were incubated either at 28.5 or 32°C and a modest improvement on the gene editing efficiency of the PE2 was found with an incubation at 32°C. In the same study a PBS length of 10 nt rather than 13 nt was used. Surprisingly under these parameters, gene editing efficiency was not higher using PE3 strategy. Therefore, the use of PE2 as protein represents an advancement in the field that may be worth optimizing<sup>234</sup>.

In contrast to these reported results using PE2 protein, we obtained few alleles containing both the wanted mutation and INDEL in the 4 mutations generated using PE3 approach, neither did we observe gene editing by Sanger sequencing using PE2 system. To investigate further these differences, it would be worth to use the same pegRNAs reported in this study with our conditions and to analyze if these differences are due to technicalities such as the use of protein *vs* mRNA and/or homemade *vs* synthetic RNA guides or to biological differences such as the targeted locus. These results would allow us to also analyze the improvement of gene editing using the PE protein.

Overall, some optimizations are still needed to be able to use prime editing in zebrafish as a tool of choice for the generation of precise genetic modifications.

### III.3.2. MS2-PE2 TECHNOLOGY.

To further enlarge the possibilities of gene modification into the zebrafish such as the insertion of large DNA sequences, a novel gene editing tool in close collaboration with Dr. Jean-Paul Concordet (*MNHN, Paris*) is under current development: the MS2-PE2, which couples the PE2 and MS2 systems.

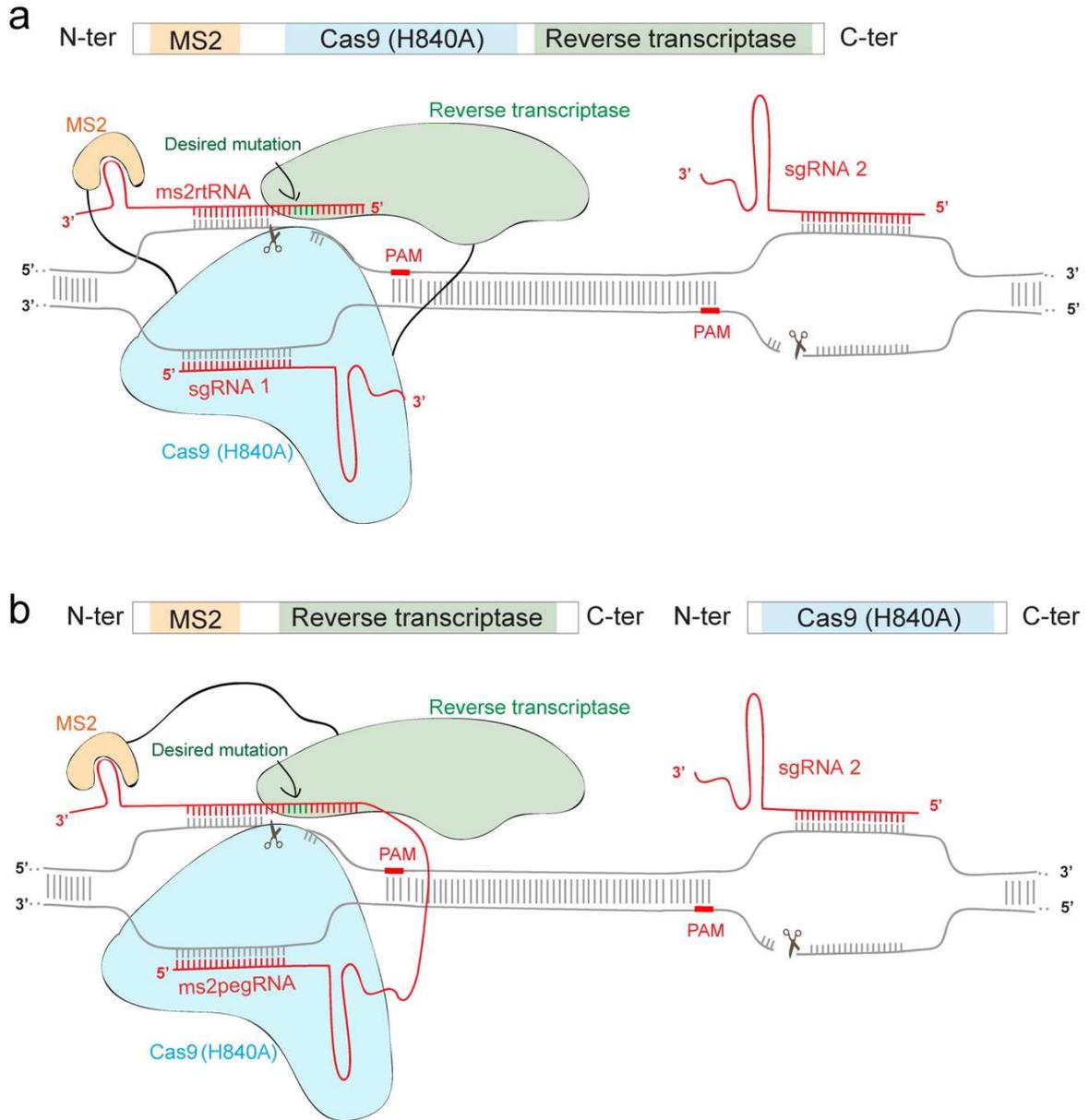
The bacteriophage coat protein MS2 binds with high affinity and specificity the MS2 RNA hairpins by its RNA-binding site. These MS2 RNA stem loop structures are found only in the bacteriophage RNA, conferring the high specificity of this system. The addition of the MS2 RNA loops to the sgRNA in SAM CRISPRa system was shown not to affect its function and was used as such in several models for its high affinity, and specificity (*c.f.* section I.2.1.3). Based on these findings, we decided to test whether

the combination of the MS2 and PE2 systems could increase the efficiency or possibilities of precise gene editing in zebrafish.

To this aim, two approaches were designed and some of the constructs are now available (Fig. 26). The first one is based on the fusion of the MS2 protein on the N-terminal end of the PE2. Here, instead of using a pegRNA, a classic sgRNA serves to generate the nick upstream of the mutation site (Fig. 26.a, sgRNA1) and a ms2rtRNA contains the PBS and RT-template sequences with the MS2 RNA loops at its 3' end (Fig. 26.a, ms2rtRNA). As the RT-template is not followed by the tracrRNA sequence, as it is the case in the pegRNA, we hypothesize that this system would allow to insert larger DNA sequences and will avoid the integration of reverse transcribed tracrRNA sequences at the mutation sites as reported in some studies<sup>140, 234</sup>. The template of the RT will thus be brought to the locus of interest by the MS2 (Fig. 26.a). In the second approach, the MS2 and the RT are fused and the Cas9(H840A) acts alone (Fig. 26.b). The RNA MS2 loop is added to the 3' end of the pegRNA (Fig. 26.b, ms2pegRNA). In this method, an increase of efficiency by keeping both enzymes separate will be tested. In addition, these two approaches could be tested with or without a second guide RNA to force the DNA repair on the wild-type strand of DNA as done in the PE3 approach (Fig. 26, sgRNA2).

Testing the DNA constructs that we have engineered to this end will allow to potentially develop an improved version of prime editing technology where we expect to obtain higher efficiencies and to insert larger sequences.

Overall, the future improvement of prime editing through this or other strategies would not only be an important advancement for zebrafish studies but also in other systems, including gene therapies to even further accelerate gene editing possibilities to rescue pathological mutations in human patients



**Figure 26: Schematic representation of the two MS2-PE2 approaches. (a)** MS2 protein in orange is fused to PE2 and recognizes the MS2 RNA in 3' of the ms2rtRNA containing the PBS and RT-template sequences. The sgRNA 1 is used to generate the nick upstream of the mutation site. **(b)** MS2 protein in orange is fused to RT in green. The Cas9(H840A) acts alone and recognizes the ms2rtRNA to generate the nick upstream of the mutation site. The MS2 RNA in 3' of the ms2pegRNA is recognized by the MS2 which brings the RT at the locus of interest. **(a, b)** The sgRNA2 is added to guide the DNA repair on the non-edited strand of DNA.

# CHAPTER IV: CONCLUSION

Along with previous studies, this work further highlights the power of Zebrafish as a vertebrate model of choice to study developmental biology and for modeling diseases. Since the emergence of CRISPR/Cas9, several disease models have been generated. Although extremely efficient to introduce LOF mutations, this technology is based on the generation of DSBs leading to unpredictable by-products. Furthermore, in zebrafish HDR-based knock-in approaches are to date still challenging, poorly efficient and often lead to the introduction of many unwanted mutations.

In this study, a large panel of CBEs was tested in order to introduce precisely C-to-T mutations in zebrafish<sup>141, 214</sup>: the BE4-gam, AncBE4max, xCas9-BE4, CBE4max-SpG, CBE4max-SpRY and the *Spymac*-ancBE4max that we created based on the reported *Spymac*-BE3<sup>204</sup>. Using the BE4-gam and AncBE4max we were able to reach an unprecedented C-to-T conversion efficiency of 91% and show the specificity of BE4-gam in zebrafish by NGS. We next broadened the base editing possibilities using *Spymac*-ancBE4max (5'-NAA-3' PAMs) and the recent CBE4max-SpRY (5'-NRN-3' PAMs) which has never been reported working in animals so far. CBE4max-SpRY showed a C-to-T conversion efficiency up to 100% in different loci and was reliable to perform multiplex mutagenesis. Based on these results, we developed a base editing co-selection method, the first example developed in an animal model *in vivo*. Using these genetic tools, we activated endogenously Wnt signaling pathway, created disease models for dwarfism and a rare poly-malformative syndrome. Furthermore, we targeted several oncogenes and tumor suppressor genes, demonstrating that BEs is a powerful approach to complement the usual cancer genes overexpression systems employed in zebrafish. In fact, endogenous gene regulations often critical in these diseases are conserved using BE strategies, that therefore should be favored when possible.

We next developed the versatile PE technology in this vertebrate model. We obtained specific gene editing for conversions, 15-bp insertion and 15-bp deletion. However, the gene editing efficiency remains to be improved and a high proportion of INDELS are still found by NGS. In this line, we are engineering a PE maternal

transgenic line in order to provide PE protein in the egg and an alternative strategy based on prime editing to enrich this essential genome editing toolbox, the MS2-PE2 technology. As such, upon optimizations PE, which is highly promising, could be among the forefront gene editing strategies.

To conclude, my PhD work contributed in providing a large panel of genetic tools working in zebrafish without generating DSBs and also assessed their efficiencies and limitations for zebrafish. Through the accessibility and ease of genome manipulations in this animal model, it also shows that zebrafish is an exceptional platform to engineer new genome editing approaches in animals and to study genetic disorders.

# MATERIALS AND METHODS

## FISH LINES AND HUSBANDRY

Zebrafish (*Danio rerio*) were maintained at 28 °C on a 14 h light/10 h dark cycle. Fish were housed in the animal facility of our laboratory in the Institut de la Vision which was built according to the respective local animal welfare standards. All animal procedures were performed in accordance with French and European Union animal welfare guidelines. Animal handling and experimental procedures were approved by the Committee on ethics of animal experimentation.

## X-RAY ANALYSES

The adult fish were fixed 2 days at 4°C in PFA 4% diluted in PBST and then dehydrated using 100% Methanol and stored in 100% Methanol at -20°C. Prior the X-ray scanning, fish were rehydrated by several washes: 75% Methanol/PBST, 50% Methanol/PBST, 25% Methanol/PBST and two washes of PBST. The fish's whole skeletons were X-rayed using a Faxitron MX-20 DC12 (Tucson), 5x magnification (Institut Imagine, Paris, France). The skeleton measures were done using ImageJ software, the rib lengths were determined using the neuronJ plugin of ImageJ.

## NGS ANALYSES

Sequencing data were analyzed with a custom Python pipeline, Crispies<sup>134</sup> that was adapted to characterize base editing and prime editing outcomes (Museum national d'histoire naturelle, Paris, France). The primer sequences used to amplify the targeted loci analyzed by NGS are listed in the following table.

Gene	Primer Forward (5'-3')	Primer Reverse (5'-3')
<i>ctnnb1</i>	CCCAGTGTAAGTTCCTTTACATT	GTGCTTCATATGGGACAAAAACACT
<i>tp53</i>	GTCATGACGGCTGACAACCTTTATTG	GTGCTTCATATGGGACAAAAACACT
<i>tek</i>	ATCTCAGACGTGACTCTGGTGAA	CAGTAGAGTTTCTGATCTGGCAGTA
<i>cbl</i>	CCCATCACAAAGACAAGCCATTAC	TGTTGGAAAGCTCTCGAGATCCAC
<i>tyrosinase</i>	TCAGGTCGAGGGTTCTGTCA	TTTTGGCCAGATTTAGGTACGAGAT

## MOLECULAR CLONING

The pDEST\_βactinhsp70:Kalt4;cmcl2:eGFP plasmid was generated by combining four plasmids using the Multisite Gateway system (Invitrogen): p5E-βactinhsp70, pME-Kalt4, p3E-polyA and pDEST-cmcl2:eGFP containing *Tol2* sites<sup>253</sup>. The βactin promoter was cloned into the *pCR-bluntII-TOPO* vector (Invitrogen) and inserted in the p5E-MCS by ligation using KpnI and XhoI restriction sites. The 3' 638bp part of the *hsp70* promoter<sup>236</sup> was then inserted into the p5E-βactin vector linearized with XhoI restriction enzyme using the Gibson Assembly Cloning Kit (New England Biolabs, E5520S). The optimized Gal4, Kalt4<sup>254</sup> was amplified and inserted in a pDONR221 using the Multisite Gateway system (Invitrogen).

To generate the other plasmids of these studies, NEBuilder HiFi DNA Assembly Cloning Kit (New England Biolabs, E5520S) was used and the PCR fragments were inserted into the plasmid linearized with restriction enzymes. All the inserted fragments were amplified using Phusion High-Fidelity DNA polymerase (ThermoFisher Scientific, F-530XL). The details are listed in the following table.

The gBlock used to generate the four *PIP5K1C* plasmids was the human *PIP5K1C* coding sequence codon optimized for expression in zebrafish using online software from IDT tool and synthesized as gBlock gene fragments from IDT.

The three synthetic gBlock used to generate the pCS2+\_pegRNA\_gfi1ab\_5'\_V5tag, pCS2+\_pegRNA\_gfi1ab\_3'\_V5tag and pCS2+\_pegRNA\_ctnnb1\_V5tag plasmids contained the *T7* promoter and the pegRNA of each respective pegRNAs.

Plasmid generated		Primer Forward (5'-3')	Primer Reverse (5'-3')	DNA template of the PCR	Destination vector linearization
pCS2+_PIP5K1C-WT	1	TGCAGGATCCCATCGATTCCGGC CACCATGGAGTTGGAAG	TAGAGGCTCGAGAGGCCTTGT TACGTGTCAGATTCTCCGTC	gBlock	pCS2+ with EcoRI
pCS2+_PIP5K1C (E146K)	1	TGCAGGATCCCATCGATTCCGGC CACCATGGAGTTGGAAG	ATTCATATAGCAGCC TGGCAAGAGCTTC	gBlock	pCS2+ with EcoRI
	2	TGCCAGGCTGCTATATG AATCTGAACCAAAAC	TAGAGGCTCGAGAGGCCTTGT TACGTGTCAGATTCTCC		
pCS2+_PIP5K1C (Y205C)	1	TGCAGGATCCCATCGATTCCGGC CACCATGGAGTTGGAAG	GTAAGACCCGAGAAT TTCGGAAGGAGCG	gBlock	pCS2+ with EcoRI
	2	CGAAATTCGCGGTCTT TACTGTGTACAG	TAGAGGCTCGAGAGGCCTTGT TACGTGTCAGATTCTCC		
pCS2+_PIP5K1C (Y221C)	1	TGCAGGATCCCATCGATTCCGGC CACCATGGAGTTGGAAG	CCGAACAATTCCTAAATAC CTAAACGCCAC	gBlock	pCS2+ with EcoRI
	2	TATTTAGGAAATTGT TCGGTATCCGTC	TAGAGGCTCGAGAGGCCTTGT TACGTGTCAGATTCTCCGTC		
pCMV_AncBE4max (R33A)	1	CAGATCCGCTAGAGATCCGCGGC CGTAATACGACTCAC	GTACAGCAGGCATGTCT CCTTGGCCAGC	pCMV_ ancBE4max addgene #112094	pCMV_ ancBE4max addgene #112094 with NotI and StuI
	2	AGGAGACATGCCTGC TGTACGAGATCAAG	TAGATCCTCCAGACTTCAGGC CTGTAGCCACAGAATATG		
pCMV_AncBE4max (R33AK34A)	1	CAGATCCGCTAGAGATCCGCGGC CGTAATACGACTCACTATAGG	GTACAGCAGGCATGTCTCGG CGGCCAGC		
	2	CCGAGACATGCCTGCTGTA CGAGATCAAG	TAGATCCTCCAGACTTCAGGCC TGTAGCCACAGAATATG		
pCMV_AncBE4max (E63Q)	1	CAGATCCGCTAGAGATCCGCG GCCGCTAATACGACTCAC	TTCTCGATGAAATTCACTT GCACGTGCTTTG		
	2	CAAGTGAATTCATCGAGA AGTTTACCTC	TAGATCCTCCAGACTTCAGGC CTGTAGCCACAGAATATG		
pCS2+_CBE4max-SpRY (R33A)	1	TTCTTTTGCAGGATCCCATCGA TTCGGCCACCATGAAACGGACAG	ATGTCTCTTGGCCAGCT CCCAGGGGTC	pCS2+_ CBE4max-SpRY	pCS2+_ CBE4max-SpRY with AvrII and ClaI
	2	GGAGCTGGCCAAGGAGAC ATGCCTGCTGTAC	GCACCCACAGGTGTGGGTA CCTAGGCCAGTGGGCCTCG		
pCS2+_PE2	1	TGCAGGATCCCATCGATTCCGG CCACCATGAAACGGACAG	TAGAGGCTCGAGAGGCCTTGT AGACTTTCCTCTTCTTCTTGGG	pCMV_PE2 addgene #132775	pCS2+ with EcoRI
pCS2+_PE2-SpRY	1	TGCGGAAAGTGCTGAGCATGC CCCAAGTGAATATCGTAAAAAAG	ATCCTCCAGAGTACCTC CCAGCTGAGAC	pCS2+_ CBE4max-SpRY	pCS2+_PE2 with AvrII and SphI
	2	GGGAGGTGACTCTGGA GGATCTAGCGGAG	GCGACAGAAGCCTGCCTTCC CTAGGAACTCCCTTAGTTGTC	pCS2+_PE2	
pCS2+_MS2-PE2	1	TGCCTGGGGACGTCGGAGCAA GCTTGATTAGGTGACACTATAG	GAAGCCGAGCCACCGACT TTCCGTTCTTC	pCS2+_PE2	pCS2+_PE2 with SacI and HindIII
	2	AAAGTCGGTGGCTC GGCTTCAAACCTTAC	TTGTCTGAGCCTCCGTAGATA CCTGAGTTAG	MS2-P65-HSF1 addgene #61423	
	3	ATCTACGGAGGCTCAGACAA GAAGTACAGC	GCCATCACTTTCACGA GCTCGTCCACCACCTTCACTG	pCS2+_PE2	
pCS2+_MS2-RT	1	TGCAGGATCCCATCGATTCCGG CCACCATGGCTTCAAAC	ATCCTCCAGATGAGCCTC CGTAGATACC	MS2-P65-HSF1 addgene #61423	pCS2+ with EcoRI
	2	CGGAGGCTCATCTGGA GGATCTAGCGGAG	TAGAGGCTCGAGAGGCCTTGT TAGACTTTCCTCTTCTTCTTGG	pCS2+_PE2	
14UASubc:BE4- gam_t2a_GFP	1	TCTTCTCAGCGTAAAGCCACCA TGGCTAAACCAGCAAAAC	CTCCGCCCGGACTTT CCTCTTCTTCTTGG	pCMV_BE4-gam addgene #100806	14UASubc with NcoI
	2	GAGGAAAGTCGGCGGCGG AGAGGGCAGA	ATTACTAGTCTCGAGATCTCTCA CTTGACAGCTCGTCCATGCCG	p3E-T2A-eGFP	
pCS2+_pegRNA_ gfi1ab-5'-V5tag	1	ATTC AAGGCCTCTCGAGCCTA AGCTAATACGACTCACTATAGG	TAATACGACTCACTATAGTTTCT AGAGGAGGACACAATGG	gBlock	pCS2+ with XbaI
pCS2+_pegRNA_ gfi1ab-3'-V5tag	1	ATTC AAGGCCTCTCGAGCCTA AGCTAATACGACTCACTATAGG	TAATACGACTCACTATAGTTTCT AGAGCACGGACTGAAG	gBlock	
pCS2+_pegRNA_ ctnnb1-V5tag	1	ATTC AAGGCCTCTCGAGCCTAA GCTAATACGACTCACTATAG	TAATACGACTCACTATAGTTTCT TAGAGTCTTACCTGGAC	gBlock	
14UASubc:PE2_t2a_ GFP	1	TCTTCTCAGCGTAAAGCCACCA TGAAACGGACAGCCGAC	CTCCGCCCGGACTTT CCTCTTCTTCTTGGG	pCS2+_PE2	14UASubc with NcoI
	2	GAGGAAAGTCGGCGGCGG GAGAGGGCAGAG	ATTACTAGTCTCGAGATCTCTC ACTTGTACAGCTCGTCCATGCC	UAS:BE4- gam_t2a_GFP	

## WHOLE-MOUNT *IN SITU* HYBRIDIZATION

The primers used to amplify the template for the probe synthesis from cDNA were: Fwd-5'-GTGGCATTCTGCTGAGTCAATGG-3' and Rev-5'-CCAAGCTTCTAATACGACTCACTATAGGGAGAGGCTCTACTTGTACAGCCTGAGG-3' for *pip5k1ca* and Fwd-5'-GCCTTTCGCTACTTCAGAGAGC-3' and Rev-5'-CCAAGCTTCTAATACGACTCACTATAGGGAGACATGGACCAGAGCTTTCCACG-3' for *pip5k1cb*. The reverse primers contained the *T7* sequence. Digoxigenin-labeled riboprobes were prepared as recommended by the manufacturer instructions using the PCR products as templates (Roche). Whole-mount *in situ* hybridization was performed using standard procedures<sup>255</sup>.

## RNAS SYNTHESIS

To synthesize mRNAs *in vitro*, pCS2+\_*PIP5K1C*, pCS2+\_*PIP5K1C(E146K)*, pCS2+\_*PIP5K1C(Y205C)*, pCS2+\_*PIP5K1C(Y221C)*, pCS2+\_*PE2*, pCS2+\_*PE2-SpRY* and pCS2+\_*CBE4max-SpRY(R33A)* plasmids were linearized with NotI restriction enzyme and mRNAs were synthesized by *in vitro* transcription using the mMACHINE mMACHINE Sp6 kit (Ambion, AM1340). The final products were purified using the RNeasy kit (QIAGEN, 74104).

To synthesize pegRNAs *in vitro*, pCS2+\_*T7-gfi1ab-3'-V5tag*, pCS2+\_*T7-gfi1ab-5'-V5tag* and pCS2+\_*T7-ctnnb1-S33-V5tag* are linearized with XbaI restriction enzyme and RNAs were synthesized by *in vitro* transcription using the MEGAscript T7 transcription kit (ThermoFisher, AMB13345). E.Z.N.A.® MicroElute RNA Clean Up Kit (Omega, R6247-01) was used for the purification, following the protocol except on the step 3 where 1x Volume of 100% ethanol was used.

## MICRO-INJECTION

For the injection of the different *PIP5K1C* mRNAs, the mRNAs were diluted in DNase/RNase-free water. One nanoliter of each mix was then injected into the cell at one-cell stage zebrafish embryos.

To form the synthetic sgRNA complex prior injections, a mix of 2 µl of the Alt-R CRISPR-Cas9 crRNA (100 pmol/µl) and 2 µl of Alt-R CRISPR-Cas9 tracrRNA (100 pmol/µl) from IDT was incubated at 95°C for 5 minutes (min), cooled down at room temperature, and then kept on ice. One nanoliter of another mix containing *CBE*

mRNA (600 ng/ $\mu$ l) and the synthetic sgRNA complex (43 pmol/ $\mu$ l) was then injected into the cell at one-cell stage zebrafish embryos. The sgRNA sequences used in this study are listed in the Table 2 and the sgRNAs in the following table.

sgRNA	Sequence (5'-3')
<i>tp53</i> sgRNA	CAATCAGCGAGCAAATTACA
<i>tek</i> sgRNA	GGAGCTCCAGGTGACGGTAG
<i>cbl</i> sgRNA	AGTCCAGTCTGGCATGTTG
<i>ctnnb1</i> sgRNA	CTGGACTCAGGAATACTC
<i>pip5k1ca</i> sgRNA 1	GAGCTCTCTGAAGTAACGGA
<i>pip5k1ca</i> sgRNA 2	AGAGCTCTCTGAAGTAACGG
<i>pip5k1ca</i> sgRNA 3	AAGAGCTCTCTGAAGTAACG
<i>pip5k1ca</i> sgRNA 4	GAAGAGCTCTCTGAAGTAAC
<i>pip5k1cb</i> sgRNA 1	GAGCTCTCTGAAGTAGCGAA
<i>pip5k1cb</i> sgRNA 2	AGAGCTCTCTGAAGTAGCGA
<i>tyrW273*NAN</i> sgRNA	CCTTCCAGGATGAGAACACA
<i>tyr_G73*</i> PE3 sgRNA	TTCGGGCGCACTGGCAGGTT
<i>tyr_G93*</i> PE3 sgRNA	CTCGCTCTCTCTGCGCAGT
<i>slc45a_G45*</i> PE3 sgRNA	CCGAGCCGATGACGGGCTGT
<i>ctnnb1</i> PE3 sgRNA	TATCCACATCGTCATCCTCC
<i>gfi1ab</i> 5' PE3 sgRNA	CCTCAAAGATCGCGGCTGG
<i>gfi1ab</i> 3' PE3 sgRNA	AGTCTTTGTCCAATAGTGTT
<i>tyr_S50*</i> PE3 sgRNA	GTGAGGATACTGCGGCCCGT

## TRANSGENIC LINES GENERATION

The *Tg(UAS:BE4-gam\_t2a\_GFP)* transgenic line was generated by injecting the *UAS:BE4-gam\_t2a\_GFP* plasmid and *tol2* mRNA into one-cell stage *Tg(rx2:gal4;cmlc2:eGFP)* embryos.

The *Tg( $\beta$ actinhsp70:KalT4;cmlc2:eGFP)* transgenic line was generated by injecting the pDEST\_ *$\beta$ actinhsp70:KalT4;cmlc2:eGFP* plasmid and *tol2* mRNA into one-cell stage embryos.

The *Tg(14UASubc:PE2\_t2a\_GFP)* transgenic line was generated by injecting the *14UASubc:PE2\_t2a\_GFP* plasmid and *tol1* mRNA into one-cell stage *Tg( $\beta$ actinhsp70:KalT4;cmlc2:eGFP)* embryos.

## WHOLE-EMBRYO DNA SEQUENCING

For genomic DNA extraction, single or pool embryo were digested for 1 hour at 55°C in 10 µl per embryo of lysis buffer (10 mM Tris, pH 8.0, 10 mM NaCl, 10 mM EDTA, and 2% SDS) with proteinase K (0.17 mg/mL, Roche Diagnostics) and inactivated 10 min at 95°C. To sequence and check for frequency of mutations, each target genomic locus was PCR-amplified using Phusion High-Fidelity DNA polymerase (ThermoFisher Scientific, F-530XL). PCR products have been extracted from an agarose gel and purified using the PCR clean-up gel extraction kit (Macherey-Nagel, 740609.50), and Sanger sequencing was performed by Eurofins. Sequence analyses were achieved using ApE software and quantifications of the mutation rate done using editR online software<sup>232</sup>. The primer sequences used to amplify the targeted loci are listed in the following table.

Gene	Primer Forward (5'-3')	Primer Reverse (5'-3')
<i>pip5k1ca</i>	CCTCTCCAAAGTTGAAGGTAGATGT	CTTCTGTAAAACTCGGCCTCCTT
<i>pip5k1cb</i>	AACAGTGAAGGAGTAATCTGACTC	TCTTATATGCAACCACACTGAAAG
<i>ctnmb1</i>	CCCAGTGTAAGTCCCTTTACATTT	GTGCTTCATATGGGACAAAAACACT
<i>tyrosinase_ G73</i>	TCAGGTCGAGGGTTCTGTCA	TTTTGGCCAGATTTAGGTACGAGAT
<i>tyrosinase_ W273</i>	AACCACCATAAGCCCGATTATATG	CCGCCCTAGAACTAACATTCTTG
<i>slc45a_ G45</i>	AACCATGACTCTTCTTACTGAGGAC	GACCCTGAAACTCATCTACTTCCTT
<i>gfi1ab 3'</i>	CAATCTGATAACGCACAGTCGG	CCTAATGTAAACGCTATTAGGCACC
<i>gfi1ab 5'</i>	CACCTGTCTTCATAACACCTGG	CAAGTCTGTTGAAGTCATCCTCC
<i>V5-tag_ screen-F</i>	GTAAGCCTATCCCTAACCTCTCC	
<i>V5-tag_ screen-R</i>		GGAGAGGGTTAGGGATAGGCTTAC

## ALCIAN BLUE STAINING

The staining was performed using an acid-free protocol<sup>256</sup>. 3 dpf embryos were fixed in 100% Ethanol overnight at 4°C and were rehydrated in 50% Ethanol in diluted in PBS for 10 min at room temperature. Embryos were then transferred overnight at room temperature in staining solution containing 0.02% Alcian Blue solution (Sigma Aldrich, A5268), 40mM MgCl<sub>2</sub> in 70% Ethanol solution. Embryos were washed once with H<sub>2</sub>O and bleached, without being covered, using 3% H<sub>2</sub>O<sub>2</sub> (Sigma Aldrich,

H1009) and 2% KOH (VWR, 26668.296) for 20 min at room temperature. Clearing was achieved using sequential glycerol/KOH washes: 1) 20% glycerol/0.25% KOH for 3 hours, 2) 50% glycerol/0.25% KOH for 3 hours and in 50% glycerol/0.1% KOH. Embryos were stored in the last solution at 4°C until imaging.

Imaging were done using Leica MZ10F microscope. Measurements of the cartilage elements were determined at 3 dpf from ventral views of Alcian blue stained larvae using ImageJ.

## STATISTICAL ANALYSIS

The non-parametric t-test Wilcoxon-Mann-Whitney was applied to determine significance of the larvae cartilage measures, the base editing and the skeleton measures of *cbl* mutant except for the rib measurement for which the parametric t-test was performed. The software used was Prism 7 (GraphPad).

## IMAGING

The inverted laser scanning confocal microscope Zeiss CLSM-LSM780 was used for high-resolution microscopy of *Tg(tcf:gfp)* embryos, employing a 40x water immersion objective. Z-stacks were acquired every 1–2 µm.

Leica MZ10F was used to image the whole embryos and adult fish of these studies.

# REFERENCES

1. Santoriello C, Zon LI. Hooked! Modeling human disease in zebrafish. *J Clin Invest* **122**, 2337-2343 (2012).
2. Bradford YM, *et al.* Zebrafish Models of Human Disease: Gaining Insight into Human Disease at ZFIN. *ILAR J* **58**, 4-16 (2017).
3. Patton EE, Tobin DM. Spotlight on zebrafish: the next wave of translational research. *Dis Model Mech* **12**, (2019).
4. Adamson KI, Sheridan E, Grierson AJ. Use of zebrafish models to investigate rare human disease. *J Med Genet* **55**, 641-649 (2018).
5. Okuda KS, Hogan BM. Endothelial Cell Dynamics in Vascular Development: Insights From Live-Imaging in Zebrafish. *Front Physiol* **11**, 842 (2020).
6. Ignatius MS, Hayes M, Langenau DM. In Vivo Imaging of Cancer in Zebrafish. *Adv Exp Med Biol* **916**, 219-237 (2016).
7. Bin JM, Lyons DA. Imaging Myelination In Vivo Using Transparent Animal Models. *Brain Plast* **2**, 3-29 (2016).
8. Babin PJ, Goizet C, Raldua D. Zebrafish models of human motor neuron diseases: advantages and limitations. *Prog Neurobiol* **118**, 36-58 (2014).
9. Lewis KE, Eisen JS. From cells to circuits: development of the zebrafish spinal cord. *Prog Neurobiol* **69**, 419-449 (2003).
10. Bakkers J. Zebrafish as a model to study cardiac development and human cardiac disease. *Cardiovasc Res* **91**, 279-288 (2011).
11. Schmitt EA, Dowling JE. Early retinal development in the zebrafish, *Danio rerio*: light and electron microscopic analyses. *J Comp Neurol* **404**, 515-536 (1999).
12. Geisler R, Borel N, Ferg M, Maier JV, Strahle U. Maintenance of Zebrafish Lines at the European Zebrafish Resource Center. *Zebrafish* **13 Suppl 1**, S19-23 (2016).
13. Kimmel CB. Genetics and early development of zebrafish. *Trends Genet* **5**, 283-288 (1989).
14. Dahm R, Geisler R. Learning from small fry: the zebrafish as a genetic model organism for aquaculture fish species. *Mar Biotechnol (NY)* **8**, 329-345 (2006).
15. Howe K, *et al.* The zebrafish reference genome sequence and its relationship to the human genome. *Nature* **496**, 498-503 (2013).

16. Zon LI, Peterson RT. In vivo drug discovery in the zebrafish. *Nat Rev Drug Discov* **4**, 35-44 (2005).
17. Mathias JR, Saxena MT, Mumm JS. Advances in zebrafish chemical screening technologies. *Future Med Chem* **4**, 1811-1822 (2012).
18. Miscevic F, Rotstein O, Wen XY. Advances in zebrafish high content and high throughput technologies. *Comb Chem High Throughput Screen* **15**, 515-521 (2012).
19. Tamplin OJ, *et al.* Small molecule screening in zebrafish: swimming in potential drug therapies. *Wiley Interdiscip Rev Dev Biol* **1**, 459-468 (2012).
20. MacRae CA, Peterson RT. Zebrafish as tools for drug discovery. *Nat Rev Drug Discov* **14**, 721-731 (2015).
21. Grunwald DJ, Streisinger G. Induction of recessive lethal and specific locus mutations in the zebrafish with ethyl nitrosourea. *Genet Res* **59**, 103-116 (1992).
22. Haffter P, *et al.* The identification of genes with unique and essential functions in the development of the zebrafish, *Danio rerio*. *Development* **123**, 1-36 (1996).
23. Driever W, *et al.* A genetic screen for mutations affecting embryogenesis in zebrafish. *Development* **123**, 37-46 (1996).
24. Raby L, Volkel P, Le Bourhis X, Angrand PO. Genetic Engineering of Zebrafish in Cancer Research. *Cancers (Basel)* **12**, (2020).
25. Beckwith LG, Moore JL, Tsao-Wu GS, Harshbarger JC, Cheng KC. Ethylnitrosourea induces neoplasia in zebrafish (*Danio rerio*). *Lab Invest* **80**, 379-385 (2000).
26. van Rooijen E, *et al.* Zebrafish mutants in the von Hippel-Lindau tumor suppressor display a hypoxic response and recapitulate key aspects of Chuvash polycythemia. *Blood* **113**, 6449-6460 (2009).
27. Holder N, Xu Q. Microinjection of DNA, RNA, and protein into the fertilized zebrafish egg for analysis of gene function. *Methods Mol Biol* **97**, 487-490 (1999).
28. Hyatt TM, Ekker SC. Vectors and techniques for ectopic gene expression in zebrafish. *Methods Cell Biol* **59**, 117-126 (1999).
29. 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* **225**, 17-22 (1998).
30. 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* **240**, 239-244 (1999).

31. 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 U S A* **97**, 11403-11408 (2000).
32. Suster ML, Kikuta H, Urasaki A, Asakawa K, Kawakami K. Transgenesis in zebrafish with the tol2 transposon system. *Methods Mol Biol* **561**, 41-63 (2009).
33. Grabher C, Wittbrodt J. Recent advances in meganuclease-and transposon-mediated transgenesis of medaka and zebrafish. *Methods Mol Biol* **461**, 521-539 (2008).
34. Langenau DM, *et al.* Myc-induced T cell leukemia in transgenic zebrafish. *Science* **299**, 887-890 (2003).
35. Sager JJ, Bai Q, Burton EA. Transgenic zebrafish models of neurodegenerative diseases. *Brain Struct Funct* **214**, 285-302 (2010).
36. Halpern ME, Rhee J, Goll MG, Akitake CM, Parsons M, Leach SD. Gal4/UAS transgenic tools and their application to zebrafish. *Zebrafish* **5**, 97-110 (2008).
37. McConnell AM, Noonan HR, Zon LI. Reeling in the Zebrafish Cancer Models. *Annual Review of Cancer Biology* **5**, 331-350 (2021).
38. McCammon JM, Doyon Y, Amacher SL. Inducing high rates of targeted mutagenesis in zebrafish using zinc finger nucleases (ZFNs). *Methods Mol Biol* **770**, 505-527 (2011).
39. Foley JE, Maeder ML, Pearlberg J, Joung JK, Peterson RT, Yeh JR. Targeted mutagenesis in zebrafish using customized zinc-finger nucleases. *Nat Protoc* **4**, 1855-1867 (2009).
40. Gupta A, Christensen RG, Rayla AL, Lakshmanan A, Stormo GD, Wolfe SA. An optimized two-finger archive for ZFN-mediated gene targeting. *Nat Methods* **9**, 588-590 (2012).
41. Sander JD, Yeh JR, Peterson RT, Joung JK. Engineering zinc finger nucleases for targeted mutagenesis of zebrafish. *Methods Cell Biol* **104**, 51-58 (2011).
42. Zhu C, *et al.* Evaluation and application of modularly assembled zinc-finger nucleases in zebrafish. *Development* **138**, 4555-4564 (2011).
43. Meng X, Noyes MB, Zhu LJ, Lawson ND, Wolfe SA. Targeted gene inactivation in zebrafish using engineered zinc-finger nucleases. *Nat Biotechnol* **26**, 695-701 (2008).
44. Huang P, Xiao A, Zhou M, Zhu Z, Lin S, Zhang B. Heritable gene targeting in zebrafish using customized TALENs. *Nat Biotechnol* **29**, 699-700 (2011).
45. 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).

46. Zhao F, *et al.* Irf8 regulates the progression of myeloproliferative neoplasm-like syndrome via Mertk signaling in zebrafish. *Leukemia* **32**, 149-158 (2018).
47. Radev Z, *et al.* A TALEN-Exon Skipping Design for a Bethlem Myopathy Model in Zebrafish. *PLoS One* **10**, e0133986 (2015).
48. Gao H, *et al.* Mecp2 regulates neural cell differentiation by suppressing the Id1 to Her2 axis in zebrafish. *J Cell Sci* **128**, 2340-2350 (2015).
49. Hwang WY, *et al.* Efficient genome editing in zebrafish using a CRISPR-Cas system. *Nat Biotechnol* **31**, 227-229 (2013).
50. Naert T, Vleminckx K. CRISPR/Cas9 disease models in zebrafish and Xenopus: The genetic renaissance of fish and frogs. *Drug Discov Today Technol* **28**, 41-52 (2018).
51. Van De Weghe JC, *et al.* Mutations in ARMC9, which Encodes a Basal Body Protein, Cause Joubert Syndrome in Humans and Ciliopathy Phenotypes in Zebrafish. *Am J Hum Genet* **101**, 23-36 (2017).
52. Gui H, *et al.* Whole exome sequencing coupled with unbiased functional analysis reveals new Hirschsprung disease genes. *Genome Biol* **18**, 48 (2017).
53. Hofmeister W, *et al.* Targeted copy number screening highlights an intragenic deletion of WDR63 as the likely cause of human occipital encephalocele and abnormal CNS development in zebrafish. *Hum Mutat* **39**, 495-505 (2018).
54. Tuschl K, *et al.* Mutations in SLC39A14 disrupt manganese homeostasis and cause childhood-onset parkinsonism-dystonia. *Nat Commun* **7**, 11601 (2016).
55. Grone BP, *et al.* Epilepsy, Behavioral Abnormalities, and Physiological Comorbidities in Syntaxin-Binding Protein 1 (STXBP1) Mutant Zebrafish. *PLoS One* **11**, e0151148 (2016).
56. Pena IA, *et al.* Pyridoxine-Dependent Epilepsy in Zebrafish Caused by Aldh7a1 Deficiency. *Genetics* **207**, 1501-1518 (2017).
57. Liu CX, *et al.* CRISPR/Cas9-induced shank3b mutant zebrafish display autism-like behaviors. *Mol Autism* **9**, 23 (2018).
58. Moortgat S, *et al.* Broadening the phenotypic spectrum and physiological insights related to EIF2S3 variants. *Hum Mutat* **42**, 827-834 (2021).
59. Collery RF, Volberding PJ, Bostrom JR, Link BA, Besharse JC. Loss of Zebrafish Mfrp Causes Nanophthalmia, Hyperopia, and Accumulation of Subretinal Macrophages. *Invest Ophthalmol Vis Sci* **57**, 6805-6814 (2016).

60. Minegishi Y, Nakaya N, Tomarev SI. Mutation in the Zebrafish *cct2* Gene Leads to Abnormalities of Cell Cycle and Cell Death in the Retina: A Model of CCT2-Related Leber Congenital Amaurosis. *Invest Ophthalmol Vis Sci* **59**, 995-1004 (2018).
61. Cai S, Chen Y, Shang Y, Cui J, Li Z, Li Y. Knockout of zebrafish interleukin 7 receptor (IL7R) by the CRISPR/Cas9 system delays retinal neurodevelopment. *Cell Death Dis* **9**, 273 (2018).
62. Sanna-Cherchi S, *et al.* Exome-wide Association Study Identifies GREB1L Mutations in Congenital Kidney Malformations. *Am J Hum Genet* **101**, 1034 (2017).
63. Brophy PD, *et al.* A Gene Implicated in Activation of Retinoic Acid Receptor Targets Is a Novel Renal Agenesis Gene in Humans. *Genetics* **207**, 215-228 (2017).
64. Bolar NA, *et al.* Heterozygous Loss-of-Function SEC61A1 Mutations Cause Autosomal-Dominant Tubulo-Interstitial and Glomerulocystic Kidney Disease with Anemia. *Am J Hum Genet* **99**, 174-187 (2016).
65. Anderson BR, *et al.* In vivo Modeling Implicates APOL1 in Nephropathy: Evidence for Dominant Negative Effects and Epistasis under Anemic Stress. *PLoS Genet* **11**, e1005349 (2015).
66. Hofsteen P, Robitaille AM, Chapman DP, Moon RT, Murry CE. Quantitative proteomics identify DAB2 as a cardiac developmental regulator that inhibits WNT/beta-catenin signaling. *Proc Natl Acad Sci U S A* **113**, 1002-1007 (2016).
67. Sasagawa S, *et al.* Downregulation of GSTK1 Is a Common Mechanism Underlying Hypertrophic Cardiomyopathy. *Front Pharmacol* **7**, 162 (2016).
68. Sasagawa S, *et al.* Comparative Transcriptome Analysis Identifies CCDC80 as a Novel Gene Associated with Pulmonary Arterial Hypertension. *Front Pharmacol* **7**, 142 (2016).
69. Zou J, *et al.* An internal promoter underlies the difference in disease severity between N- and C-terminal truncation mutations of Titin in zebrafish. *Elife* **4**, e09406 (2015).
70. Yang J, *et al.* Deletion of Pr130 Interrupts Cardiac Development in Zebrafish. *Int J Mol Sci* **17**, (2016).
71. Liu X, *et al.* The complex genetics of hypoplastic left heart syndrome. *Nat Genet* **49**, 1152-1159 (2017).
72. Wang J, *et al.* Identification of LBX2 as a novel causal gene of atrial septal defect. *Int J Cardiol* **265**, 188-194 (2018).
73. Ellis JL, *et al.* Zebrafish *abcb11b* mutant reveals strategies to restore bile excretion impaired by bile salt export pump deficiency. *Hepatology* **67**, 1531-1545 (2018).

74. Pazhakh V, Clark S, Keightley MC, Lieschke GJ. A GCSFR/CSF3R zebrafish mutant models the persistent basal neutrophil deficiency of severe congenital neutropenia. *Sci Rep* **7**, 44455 (2017).
75. Hoodless LJ, *et al.* Genetic and pharmacological inhibition of CDK9 drives neutrophil apoptosis to resolve inflammation in zebrafish in vivo. *Sci Rep* **5**, 36980 (2016).
76. Ear J, *et al.* A Zebrafish Model of 5q-Syndrome Using CRISPR/Cas9 Targeting RPS14 Reveals a p53-Independent and p53-Dependent Mechanism of Erythroid Failure. *J Genet Genomics* **43**, 307-318 (2016).
77. Stankiewicz P, *et al.* Haploinsufficiency of the Chromatin Remodeler BPTF Causes Syndromic Developmental and Speech Delay, Postnatal Microcephaly, and Dysmorphic Features. *Am J Hum Genet* **101**, 503-515 (2017).
78. Borck G, *et al.* BRF1 mutations alter RNA polymerase III-dependent transcription and cause neurodevelopmental anomalies. *Genome Res* **25**, 609 (2015).
79. Kury S, *et al.* De Novo Disruption of the Proteasome Regulatory Subunit PSMD12 Causes a Syndromic Neurodevelopmental Disorder. *Am J Hum Genet* **100**, 689 (2017).
80. Shaw ND, *et al.* Corrigendum: SMCHD1 mutations associated with a rare muscular dystrophy can also cause isolated arhinia and Bosma arhinia microphthalmia syndrome. *Nat Genet* **49**, 969 (2017).
81. Talbot JC, *et al.* Pharyngeal morphogenesis requires fras1-itga8-dependent epithelial-mesenchymal interaction. *Dev Biol* **416**, 136-148 (2016).
82. Zhang Y, *et al.* ATP6V1H Deficiency Impairs Bone Development through Activation of MMP9 and MMP13. *PLoS Genet* **13**, e1006481 (2017).
83. Bek JW, *et al.* Lrp5 Mutant and Crispant Zebrafish Faithfully Model Human Osteoporosis, Establishing the Zebrafish as a Platform for CRISPR-Based Functional Screening of Osteoporosis Candidate Genes. *J Bone Miner Res* **36**, 1749-1764 (2021).
84. Perles Z, *et al.* A human laterality disorder caused by a homozygous deleterious mutation in MMP21. *J Med Genet* **52**, 840-847 (2015).
85. Lin CY, *et al.* Klf8 regulates left-right asymmetric patterning through modulation of Kupffer's vesicle morphogenesis and spaw expression. *J Biomed Sci* **24**, 45 (2017).
86. Yin X, Hao J, Yao Y. CRISPR/Cas9 in zebrafish: An attractive model for FBN1 genetic defects in humans. *Mol Genet Genomic Med*, e1775 (2021).

87. Ziegler A, *et al.* Bi-allelic variants in IPO8 cause a connective tissue disorder associated with cardiovascular defects, skeletal abnormalities, and immune dysregulation. *Am J Hum Genet* **108**, 1126-1137 (2021).
88. Jinek M, Chylinski K, Fonfara I, Hauer M, Doudna JA, Charpentier E. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* **337**, 816-821 (2012).
89. Barrangou R, *et al.* CRISPR provides acquired resistance against viruses in prokaryotes. *Science* **315**, 1709-1712 (2007).
90. Cong L, *et al.* Multiplex genome engineering using CRISPR/Cas systems. *Science* **339**, 819-823 (2013).
91. Jinek M, *et al.* Structures of Cas9 endonucleases reveal RNA-mediated conformational activation. *Science* **343**, 1247997 (2014).
92. Nishimasu H, *et al.* Crystal structure of Cas9 in complex with guide RNA and target DNA. *Cell* **156**, 935-949 (2014).
93. Palermo G, *et al.* Key role of the REC lobe during CRISPR-Cas9 activation by 'sensing', 'regulating', and 'locking' the catalytic HNH domain. *Q Rev Biophys* **51**, (2018).
94. Halat M, Klimek-Chodacka M, Orleanska J, Baranska M, Baranski R. Electronic Circular Dichroism of the Cas9 Protein and gRNA:Cas9 Ribonucleoprotein Complex. *Int J Mol Sci* **22**, (2021).
95. El-Brolosy MA, *et al.* Genetic compensation triggered by mutant mRNA degradation. *Nature* **568**, 193-197 (2019).
96. Salanga CM, Salanga MC. Genotype to Phenotype: CRISPR Gene Editing Reveals Genetic Compensation as a Mechanism for Phenotypic Disjunction of Morphants and Mutants. *Int J Mol Sci* **22**, (2021).
97. Kunkel GR, Maser RL, Calvet JP, Pederson T. U6 small nuclear RNA is transcribed by RNA polymerase III. *Proc Natl Acad Sci U S A* **83**, 8575-8579 (1986).
98. Ablain J, Durand EM, Yang S, Zhou Y, Zon LI. A CRISPR/Cas9 vector system for tissue-specific gene disruption in zebrafish. *Dev Cell* **32**, 756-764 (2015).
99. Di Donato V, *et al.* 2C-Cas9: a versatile tool for clonal analysis of gene function. *Genome Res* **26**, 681-692 (2016).
100. Yin L, *et al.* Multiplex Conditional Mutagenesis Using Transgenic Expression of Cas9 and sgRNAs. *Genetics* **200**, 431-441 (2015).

101. Ablain J, *et al.* Human tumor genomics and zebrafish modeling identify SPRED1 loss as a driver of mucosal melanoma. *Science* **362**, 1055-1060 (2018).
102. Fu Y, *et al.* High-frequency off-target mutagenesis induced by CRISPR-Cas nucleases in human cells. *Nat Biotechnol* **31**, 822-826 (2013).
103. Hsu PD, Lander ES, Zhang F. Development and applications of CRISPR-Cas9 for genome engineering. *Cell* **157**, 1262-1278 (2014).
104. Pattanayak V, Lin S, Guilinger JP, Ma E, Doudna JA, Liu DR. High-throughput profiling of off-target DNA cleavage reveals RNA-programmed Cas9 nuclease specificity. *Nat Biotechnol* **31**, 839-843 (2013).
105. Kuscu C, Arslan S, Singh R, Thorpe J, Adli M. Genome-wide analysis reveals characteristics of off-target sites bound by the Cas9 endonuclease. *Nat Biotechnol* **32**, 677-683 (2014).
106. Wu X, *et al.* Genome-wide binding of the CRISPR endonuclease Cas9 in mammalian cells. *Nat Biotechnol* **32**, 670-676 (2014).
107. Canver MC, *et al.* Integrated design, execution, and analysis of arrayed and pooled CRISPR genome-editing experiments. *Nat Protoc* **13**, 946-986 (2018).
108. Concordet JP, Haeussler M. CRISPOR: intuitive guide selection for CRISPR/Cas9 genome editing experiments and screens. *Nucleic Acids Res* **46**, W242-W245 (2018).
109. Trevino AE, Zhang F. Genome editing using Cas9 nickases. *Methods Enzymol* **546**, 161-174 (2014).
110. Mali P, *et al.* CAS9 transcriptional activators for target specificity screening and paired nickases for cooperative genome engineering. *Nat Biotechnol* **31**, 833-838 (2013).
111. Ran FA, *et al.* Double nicking by RNA-guided CRISPR Cas9 for enhanced genome editing specificity. *Cell* **154**, 1380-1389 (2013).
112. Dianov GL, Hubscher U. Mammalian base excision repair: the forgotten archangel. *Nucleic Acids Res* **41**, 3483-3490 (2013).
113. Robertson AB, Klungland A, Rognes T, Leiros I. DNA repair in mammalian cells: Base excision repair: the long and short of it. *Cell Mol Life Sci* **66**, 981-993 (2009).
114. Qi LS, *et al.* Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. *Cell* **152**, 1173-1183 (2013).
115. Kaczmarczyk L, Mende Y, Zevnik B, Jackson WS. Manipulating the Prion Protein Gene Sequence and Expression Levels with CRISPR/Cas9. *PLoS One* **11**, e0154604 (2016).

116. Chavez A, *et al.* Comparison of Cas9 activators in multiple species. *Nat Methods* **13**, 563-567 (2016).
117. Albadri S, Del Bene F, Revenu C. Genome editing using CRISPR/Cas9-based knock-in approaches in zebrafish. *Methods* **121-122**, 77-85 (2017).
118. Sung P, Klein H. Mechanism of homologous recombination: mediators and helicases take on regulatory functions. *Nat Rev Mol Cell Biol* **7**, 739-750 (2006).
119. Hruscha A, *et al.* Efficient CRISPR/Cas9 genome editing with low off-target effects in zebrafish. *Development* **140**, 4982-4987 (2013).
120. Armstrong GA, Liao M, You Z, Lissouba A, Chen BE, Drapeau P. Homology Directed Knockin of Point Mutations in the Zebrafish *tardbp* and *fus* Genes in ALS Using the CRISPR/Cas9 System. *PLoS One* **11**, e0150188 (2016).
121. Hwang WY, *et al.* Heritable and precise zebrafish genome editing using a CRISPR-Cas system. *PLoS One* **8**, e68708 (2013).
122. Irion U, Krauss J, Nusslein-Volhard C. Precise and efficient genome editing in zebrafish using the CRISPR/Cas9 system. *Development* **141**, 4827-4830 (2014).
123. Zhang Y, Huang H, Zhang B, Lin S. TALEN- and CRISPR-enhanced DNA homologous recombination for gene editing in zebrafish. *Methods Cell Biol* **135**, 107-120 (2016).
124. Prykhozhiy SV, *et al.* Optimized knock-in of point mutations in zebrafish using CRISPR/Cas9. *Nucleic Acids Res* **46**, 9252 (2018).
125. Tessadori F, *et al.* Effective CRISPR/Cas9-based nucleotide editing in zebrafish to model human genetic cardiovascular disorders. *Dis Model Mech* **11**, (2018).
126. Almeida MP, *et al.* Endogenous zebrafish proneural Cre drivers generated by CRISPR/Cas9 short homology directed targeted integration. *Sci Rep* **11**, 1732 (2021).
127. Almeida MP, *et al.* Zebrafish Cre/*lox* regulated UFlip alleles generated by CRISPR/Cas targeted integration provide cell-type specific conditional gene inactivation. *bioRxiv*, 2021.2006.2018.448732 (2021).
128. Hisano Y, *et al.* Precise in-frame integration of exogenous DNA mediated by CRISPR/Cas9 system in zebrafish. *Sci Rep* **5**, 8841 (2015).
129. Wierson WA, *et al.* Efficient targeted integration directed by short homology in zebrafish and mammalian cells. *Elife* **9**, (2020).
130. Hoshijima K, Jurynek MJ, Grunwald DJ. Precise Editing of the Zebrafish Genome Made Simple and Efficient. *Dev Cell* **36**, 654-667 (2016).

131. Shin J, Chen J, Solnica-Krezel L. Efficient homologous recombination-mediated genome engineering in zebrafish using TALE nucleases. *Development* **141**, 3807-3818 (2014).
132. Zu Y, *et al.* TALEN-mediated precise genome modification by homologous recombination in zebrafish. *Nat Methods* **10**, 329-331 (2013).
133. He MD, Zhang FH, Wang HL, Wang HP, Zhu ZY, Sun YH. Efficient ligase 3-dependent microhomology-mediated end joining repair of DNA double-strand breaks in zebrafish embryos. *Mutat Res* **780**, 86-96 (2015).
134. Renaud JB, *et al.* Improved Genome Editing Efficiency and Flexibility Using Modified Oligonucleotides with TALEN and CRISPR-Cas9 Nucleases. *Cell Rep* **14**, 2263-2272 (2016).
135. Gutierrez-Triana JA, *et al.* Efficient single-copy HDR by 5' modified long dsDNA donors. *Elife* **7**, (2018).
136. Carroll D. Staying on target with CRISPR-Cas. *Nat Biotechnol* **31**, 807-809 (2013).
137. Concordet JP, Giovannangeli C. CRISPR-Cas systems for genome engineering and investigation. *Methods* **121-122**, 1-2 (2017).
138. Amendola M, *et al.* Recent progress in genome editing for gene therapy applications: the French perspective. *Hum Gene Ther*, (2021).
139. Komor AC, Kim YB, Packer MS, Zuris JA, Liu DR. Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage. *Nature* **533**, 420-424 (2016).
140. Anzalone AV, Koblan LW, Liu DR. Genome editing with CRISPR-Cas nucleases, base editors, transposases and prime editors. *Nat Biotechnol* **38**, 824-844 (2020).
141. Rosello M, *et al.* Precise base editing for the in vivo study of developmental signaling and human pathologies in zebrafish. *Elife* **10**, (2021).
142. Landrum MJ, *et al.* ClinVar: public archive of interpretations of clinically relevant variants. *Nucleic Acids Res* **44**, D862-868 (2016).
143. Landrum MJ, *et al.* ClinVar: public archive of relationships among sequence variation and human phenotype. *Nucleic Acids Res* **42**, D980-985 (2014).
144. Rees HA, Liu DR. Base editing: precision chemistry on the genome and transcriptome of living cells. *Nat Rev Genet* **19**, 770-788 (2018).
145. Kim K, *et al.* Highly efficient RNA-guided base editing in mouse embryos. *Nat Biotechnol* **35**, 435-437 (2017).

146. Rees HA, *et al.* Improving the DNA specificity and applicability of base editing through protein engineering and protein delivery. *Nat Commun* **8**, 15790 (2017).
147. Liang P, *et al.* Effective gene editing by high-fidelity base editor 2 in mouse zygotes. *Protein Cell* **8**, 601-611 (2017).
148. Zafra MP, *et al.* Optimized base editors enable efficient editing in cells, organoids and mice. *Nat Biotechnol* **36**, 888-893 (2018).
149. Yang G, *et al.* Base-Editing-Mediated R17H Substitution in Histone H3 Reveals Methylation-Dependent Regulation of Yap Signaling and Early Mouse Embryo Development. *Cell Rep* **26**, 302-312 e304 (2019).
150. Zhang Y, *et al.* Programmable base editing of zebrafish genome using a modified CRISPR-Cas9 system. *Nat Commun* **8**, 118 (2017).
151. Liu Z, *et al.* Highly efficient RNA-guided base editing in rabbit. *Nat Commun* **9**, 2717 (2018).
152. Su X, *et al.* Effective generation of maternal genome point mutated porcine embryos by injection of cytosine base editor into germinal vesicle oocytes. *Sci China Life Sci* **63**, 996-1005 (2020).
153. Shi Z, *et al.* Modeling human point mutation diseases in *Xenopus tropicalis* with a modified CRISPR/Cas9 system. *FASEB J* **33**, 6962-6968 (2019).
154. Park DS, Yoon M, Kweon J, Jang AH, Kim Y, Choi SC. Targeted Base Editing via RNA-Guided Cytidine Deaminases in *Xenopus laevis* Embryos. *Mol Cells* **40**, 823-827 (2017).
155. Komor AC, *et al.* Improved base excision repair inhibition and bacteriophage Mu Gam protein yields C:G-to-T:A base editors with higher efficiency and product purity. *Sci Adv* **3**, eaao4774 (2017).
156. d'Adda di Fagagna F, Weller GR, Doherty AJ, Jackson SP. The Gam protein of bacteriophage Mu is an orthologue of eukaryotic Ku. *EMBO Rep* **4**, 47-52 (2003).
157. Koblan LW, *et al.* Improving cytidine and adenine base editors by expression optimization and ancestral reconstruction. *Nat Biotechnol* **36**, 843-846 (2018).
158. Sasaguri H, *et al.* Introduction of pathogenic mutations into the mouse *Psen1* gene by Base Editor and Target-AID. *Nat Commun* **9**, 2892 (2018).
159. Zhao Y, Shang D, Ying R, Cheng H, Zhou R. An optimized base editor with efficient C-to-T base editing in zebrafish. *BMC Biol* **18**, 190 (2020).
160. Wang F, *et al.* Generation of a Hutchinson-Gilford progeria syndrome monkey model by base editing. *Protein Cell* **11**, 809-824 (2020).

161. Carreras A, *et al.* In vivo genome and base editing of a human PCSK9 knock-in hypercholesterolemic mouse model. *BMC Biol* **17**, 4 (2019).
162. Rossidis AC, *et al.* In utero CRISPR-mediated therapeutic editing of metabolic genes. *Nat Med* **24**, 1513-1518 (2018).
163. Lim CKW, *et al.* Treatment of a Mouse Model of ALS by In Vivo Base Editing. *Mol Ther* **28**, 1177-1189 (2020).
164. Yeh WH, *et al.* In vivo base editing restores sensory transduction and transiently improves auditory function in a mouse model of recessive deafness. *Sci Transl Med* **12**, (2020).
165. Xie J, *et al.* Efficient base editing for multiple genes and loci in pigs using base editors. *Nat Commun* **10**, 2852 (2019).
166. Zhang W, *et al.* Multiplex precise base editing in cynomolgus monkeys. *Nat Commun* **11**, 2325 (2020).
167. Suh S, *et al.* Restoration of visual function in adult mice with an inherited retinal disease via adenine base editing. *Nat Biomed Eng* **5**, 169-178 (2021).
168. Chemello F, *et al.* Precise correction of Duchenne muscular dystrophy exon deletion mutations by base and prime editing. *Sci Adv* **7**, (2021).
169. Zhou S, *et al.* Highly efficient generation of sheep with a defined FecB(B) mutation via adenine base editing. *Genet Sel Evol* **52**, 35 (2020).
170. Qin W, Lu X, Liu Y, Bai H, Li S, Lin S. Precise A\*T to G\*C base editing in the zebrafish genome. *BMC Biol* **16**, 139 (2018).
171. Xie J, *et al.* ACBE, a new base editor for simultaneous C-to-T and A-to-G substitutions in mammalian systems. *BMC Biol* **18**, 131 (2020).
172. Zhao D, *et al.* Glycosylase base editors enable C-to-A and C-to-G base changes. *Nat Biotechnol* **39**, 35-40 (2021).
173. Chen L, *et al.* Programmable C:G to G:C genome editing with CRISPR-Cas9-directed base excision repair proteins. *Nat Commun* **12**, 1384 (2021).
174. Kurt IC, *et al.* CRISPR C-to-G base editors for inducing targeted DNA transversions in human cells. *Nat Biotechnol* **39**, 41-46 (2021).
175. Lin Q, *et al.* Prime genome editing in rice and wheat. *Nat Biotechnol* **38**, 582-585 (2020).

176. Xu W, *et al.* Versatile Nucleotides Substitution in Plant Using an Improved Prime Editing System. *Mol Plant* **13**, 675-678 (2020).
177. Hua K, Jiang Y, Tao X, Zhu JK. Precision genome engineering in rice using prime editing system. *Plant Biotechnol J* **18**, 2167-2169 (2020).
178. Butt H, Rao GS, Sedeek K, Aman R, Kamel R, Mahfouz M. Engineering herbicide resistance via prime editing in rice. *Plant Biotechnol J* **18**, 2370-2372 (2020).
179. Veillet F, *et al.* Prime editing is achievable in the tetraploid potato, but needs improvement. *bioRxiv*, 2020.2006.2018.159111 (2020).
180. Lu Y, *et al.* Precise genome modification in tomato using an improved prime editing system. *Plant Biotechnol J* **19**, 415-417 (2021).
181. Jiang YY, *et al.* Prime editing efficiently generates W542L and S621I double mutations in two ALS genes in maize. *Genome Biol* **21**, 257 (2020).
182. Xu R, Li J, Liu X, Shan T, Qin R, Wei P. Development of Plant Prime-Editing Systems for Precise Genome Editing. *Plant Commun* **1**, 100043 (2020).
183. Bosch JA, Birchak G, Perrimon N. Precise genome engineering in Drosophila using prime editing. *Proc Natl Acad Sci U S A* **118**, (2021).
184. Liu Y, *et al.* Efficient generation of mouse models with the prime editing system. *Cell Discov* **6**, 27 (2020).
185. Schene IF, *et al.* Prime editing for functional repair in patient-derived disease models. *Nat Commun* **11**, 5352 (2020).
186. Surun D, *et al.* Efficient Generation and Correction of Mutations in Human iPS Cells Utilizing mRNAs of CRISPR Base Editors and Prime Editors. *Genes (Basel)* **11**, (2020).
187. Hwang GH, *et al.* PE-Designer and PE-Analyzer: web-based design and analysis tools for CRISPR prime editing. *Nucleic Acids Res* **49**, W499-W504 (2021).
188. Hsu JY, *et al.* PrimeDesign software for rapid and simplified design of prime editing guide RNAs. *Nat Commun* **12**, 1034 (2021).
189. Standage-Beier K, Tekel SJ, Brafman DA, Wang X. Prime Editing Guide RNA Design Automation Using PINE-CONE. *ACS Synth Biol* **10**, 422-427 (2021).
190. Anderson MV, Haldrup J, Thomsen EA, Wolff JH, Mikkelsen JG. pegIT - a web-based design tool for prime editing. *Nucleic Acids Res* **49**, W505-W509 (2021).
191. Chow RD, Chen JS, Shen J, Chen S. A web tool for the design of prime-editing guide RNAs. *Nat Biomed Eng* **5**, 190-194 (2021).

192. Bhagwat AM, *et al.* multicrispr: gRNA design for prime editing and parallel targeting of thousands of targets. *Life Sci Alliance* **3**, (2020).
193. Siegner SM, Karasu ME, Schroder MS, Kontarakis Z, Corn JE. PnB Designer: a web application to design prime and base editor guide RNAs for animals and plants. *BMC Bioinformatics* **22**, 101 (2021).
194. Li Y, Chen J, Tsai SQ, Cheng Y. Easy-Prime: a machine learning-based prime editor design tool. *Genome Biol* **22**, 235 (2021).
195. Hu JH, *et al.* Evolved Cas9 variants with broad PAM compatibility and high DNA specificity. *Nature* **556**, 57-63 (2018).
196. Hua K, Tao X, Han P, Wang R, Zhu JK. Genome Engineering in Rice Using Cas9 Variants that Recognize NG PAM Sequences. *Mol Plant* **12**, 1003-1014 (2019).
197. Kleinstiver BP, *et al.* Engineered CRISPR-Cas9 nucleases with altered PAM specificities. *Nature* **523**, 481-485 (2015).
198. Kim YB, Komor AC, Levy JM, Packer MS, Zhao KT, Liu DR. Increasing the genome-targeting scope and precision of base editing with engineered Cas9-cytidine deaminase fusions. *Nat Biotechnol* **35**, 371-376 (2017).
199. Guyon A, Rousseau J, Begin FG, Bertin T, Lamothe G, Tremblay JP. Base editing strategy for insertion of the A673T mutation in the APP gene to prevent the development of AD in vitro. *Mol Ther Nucleic Acids* **24**, 253-263 (2021).
200. Kim N, *et al.* Prediction of the sequence-specific cleavage activity of Cas9 variants. *Nat Biotechnol* **38**, 1328-1336 (2020).
201. Wang Y, *et al.* Cytosine Base Editor (hA3A-BE3-NG)-Mediated Multiple Gene Editing for Pyramid Breeding in Pigs. *Front Genet* **11**, 592623 (2020).
202. Ran FA, *et al.* In vivo genome editing using Staphylococcus aureus Cas9. *Nature* **520**, 186-191 (2015).
203. Kleinstiver BP, *et al.* Broadening the targeting range of Staphylococcus aureus CRISPR-Cas9 by modifying PAM recognition. *Nat Biotechnol* **33**, 1293-1298 (2015).
204. Chatterjee P, *et al.* A Cas9 with PAM recognition for adenine dinucleotides. *Nat Commun* **11**, 2474 (2020).
205. Hu Z, *et al.* A compact Cas9 ortholog from Staphylococcus Auricularis (SauriCas9) expands the DNA targeting scope. *PLoS Biol* **18**, e3000686 (2020).

206. Chatterjee P, Jakimo N, Jacobson JM. Minimal PAM specificity of a highly similar SpCas9 ortholog. *Sci Adv* **4**, eaau0766 (2018).
207. Jeong YK, Yu J, Bae S. Construction of non-canonical PAM-targeting adenosine base editors by restriction enzyme-free DNA cloning using CRISPR-Cas9. *Sci Rep* **9**, 4939 (2019).
208. Huang TP, *et al.* Circularly permuted and PAM-modified Cas9 variants broaden the targeting scope of base editors. *Nat Biotechnol* **37**, 626-631 (2019).
209. Nguyen Tran MT, *et al.* Engineering domain-inlaid SaCas9 adenine base editors with reduced RNA off-targets and increased on-target DNA editing. *Nat Commun* **11**, 4871 (2020).
210. Qin R, *et al.* Developing a highly efficient and widely adaptive CRISPR-SaCas9 toolset for plant genome editing. *Plant Biotechnol J* **17**, 706-708 (2019).
211. Yang L, *et al.* Increasing targeting scope of adenosine base editors in mouse and rat embryos through fusion of TadA deaminase with Cas9 variants. *Protein Cell* **9**, 814-819 (2018).
212. Walton RT, Christie KA, Whittaker MN, Kleinstiver BP. Unconstrained genome targeting with near-PAMless engineered CRISPR-Cas9 variants. *Science* **368**, 290-296 (2020).
213. Kweon J, *et al.* Engineered prime editors with PAM flexibility. *Mol Ther* **29**, 2001-2007 (2021).
214. Rosello M, Serafini M, Mione MC, Concordet J-P, Del Bene F. Disease modeling by efficient genome editing using a near PAM-less base editor *in vivo*. *bioRxiv*, 2021.2006.2028.450169 (2021).
215. van den Bout I, Divecha N. PIP5K-driven PtdIns(4,5)P<sub>2</sub> synthesis: regulation and cellular functions. *J Cell Sci* **122**, 3837-3850 (2009).
216. McCrea HJ, *et al.* All known patient mutations in the ASH-RhoGAP domains of OCRL affect targeting and APPL1 binding. *Biochem Biophys Res Commun* **369**, 493-499 (2008).
217. Saito T, *et al.* Mutation analysis of SYNJ1: a possible candidate gene for chromosome 21q22-linked bipolar disorder. *Mol Psychiatry* **6**, 387-395 (2001).
218. Saito T, Stopkova P, Diaz L, Papolos DF, Boussemart L, Lachman HM. Polymorphism screening of PIK4CA: possible candidate gene for chromosome 22q11-linked psychiatric disorders. *Am J Med Genet B Neuropsychiatr Genet* **116B**, 77-83 (2003).

219. Stopkova P, *et al.* Polymorphism screening of PIP5K2A: a candidate gene for chromosome 10p-linked psychiatric disorders. *Am J Med Genet B Neuropsychiatr Genet* **123B**, 50-58 (2003).
220. Stopkova P, *et al.* Identification of PIK3C3 promoter variant associated with bipolar disorder and schizophrenia. *Biol Psychiatry* **55**, 981-988 (2004).
221. Stopkova P, Vevera J, Paclt I, Zukov I, Lachman HM. Analysis of SYNJ1, a candidate gene for 21q22 linked bipolar disorder: a replication study. *Psychiatry Res* **127**, 157-161 (2004).
222. Maehama T, Taylor GS, Dixon JE. PTEN and myotubularin: novel phosphoinositide phosphatases. *Annu Rev Biochem* **70**, 247-279 (2001).
223. Wishart MJ, Dixon JE. PTEN and myotubularin phosphatases: from 3-phosphoinositide dephosphorylation to disease. *Trends Cell Biol* **12**, 579-585 (2002).
224. Luo J, Manning BD, Cantley LC. Targeting the PI3K-Akt pathway in human cancer: rationale and promise. *Cancer Cell* **4**, 257-262 (2003).
225. Osaki M, Oshimura M, Ito H. PI3K-Akt pathway: its functions and alterations in human cancer. *Apoptosis* **9**, 667-676 (2004).
226. Bader AG, Kang S, Zhao L, Vogt PK. Oncogenic PI3K deregulates transcription and translation. *Nat Rev Cancer* **5**, 921-929 (2005).
227. Engelman JA, Luo J, Cantley LC. The evolution of phosphatidylinositol 3-kinases as regulators of growth and metabolism. *Nat Rev Genet* **7**, 606-619 (2006).
228. Carracedo A, Pandolfi PP. The PTEN-PI3K pathway: of feedbacks and cross-talks. *Oncogene* **27**, 5527-5541 (2008).
229. Yuan TL, Cantley LC. PI3K pathway alterations in cancer: variations on a theme. *Oncogene* **27**, 5497-5510 (2008).
230. Doughman RL, Firestone AJ, Anderson RA. Phosphatidylinositol phosphate kinases put PI4,5P(2) in its place. *J Membr Biol* **194**, 77-89 (2003).
231. Pasquier J, *et al.* Gene evolution and gene expression after whole genome duplication in fish: the PhyloFish database. *BMC Genomics* **17**, 368 (2016).
232. Kluesner MG, *et al.* EditR: A Method to Quantify Base Editing from Sanger Sequencing. *CRISPR J* **1**, 239-250 (2018).
233. Burg L, *et al.* Internal epitope tagging informed by relative lack of sequence conservation. *Sci Rep* **6**, 36986 (2016).

234. Petri K, *et al.* CRISPR prime editing with ribonucleoprotein complexes in zebrafish and primary human cells. *Nat Biotechnol*, (2021).
235. Tadros W, Lipshitz HD. The maternal-to-zygotic transition: a play in two acts. *Development* **136**, 3033-3042 (2009).
236. Dalgin G, Ward AB, Hao le T, Beattie CE, Nechiporuk A, Prince VE. Zebrafish *mnx1* controls cell fate choice in the developing endocrine pancreas. *Development* **138**, 4597-4608 (2011).
237. Martinelli S, *et al.* Heterozygous germline mutations in the CBL tumor-suppressor gene cause a Noonan syndrome-like phenotype. *Am J Hum Genet* **87**, 250-257 (2010).
238. Harada D, Yamanaka Y, Ueda K, Tanaka H, Seino Y. FGFR3-related dwarfism and cell signaling. *J Bone Miner Metab* **27**, 9-15 (2009).
239. Cho JY, *et al.* Defective lysosomal targeting of activated fibroblast growth factor receptor 3 in achondroplasia. *Proc Natl Acad Sci U S A* **101**, 609-614 (2004).
240. Santoriello C, *et al.* Kita driven expression of oncogenic HRAS leads to early onset and highly penetrant melanoma in zebrafish. *PLoS One* **5**, e15170 (2010).
241. Kobar K, Collett K, Prykhozij SV, Berman JN. Zebrafish Cancer Predisposition Models. *Front Cell Dev Biol* **9**, 660069 (2021).
242. Jafry M, Sidbury R. RASopathies. *Clin Dermatol* **38**, 455-461 (2020).
243. Riller Q, Rieux-Laucat F. RASopathies: From germline mutations to somatic and multigenic diseases. *Biomed J*, (2021).
244. Kirchmaier S, Naruse K, Wittbrodt J, Loosli F. The genomic and genetic toolbox of the teleost medaka (*Oryzias latipes*). *Genetics* **199**, 905-918 (2015).
245. Schulze L, *et al.* Transparent *Danio rerio* as a genetically tractable vertebrate brain model. *Nat Methods* **15**, 977-983 (2018).
246. Thumberger T, *et al.* hei-tag: a highly efficient tag to boost targeted genome editing. *bioRxiv*, 2021.2005.2027.445956 (2021).
247. Cornean A, Gierten J, Welz B, Mateo JL, Thumberger T, Wittbrodt J. Precise *in vivo* functional analysis of DNA variants with base editing using ACEofBASEs target prediction. *bioRxiv*, 2021.2007.2026.453883 (2021).
248. Carrington B, Weinstein RN, Sood R. BE4max and AncBE4max Are Efficient in Germline Conversion of C:G to T:A Base Pairs in Zebrafish. *Cells* **9**, (2020).

249. Vicencio J, *et al.* Genome editing in animals with minimal PAM CRISPR-Cas9 enzymes. *bioRxiv*, 2021.2006.2006.447255 (2021).
250. Song M, *et al.* Sequence-specific prediction of the efficiencies of adenine and cytosine base editors. *Nat Biotechnol* **38**, 1037-1043 (2020).
251. Grunewald J, *et al.* Transcriptome-wide off-target RNA editing induced by CRISPR-guided DNA base editors. *Nature* **569**, 433-437 (2019).
252. Horstick EJ, *et al.* Increased functional protein expression using nucleotide sequence features enriched in highly expressed genes in zebrafish. *Nucleic Acids Res* **43**, e48 (2015).
253. Kwan KM, *et al.* The Tol2kit: a multisite gateway-based construction kit for Tol2 transposon transgenesis constructs. *Dev Dyn* **236**, 3088-3099 (2007).
254. Distel M, Wullimann MF, Koster RW. Optimized Gal4 genetics for permanent gene expression mapping in zebrafish. *Proc Natl Acad Sci U S A* **106**, 13365-13370 (2009).
255. Oxtoby E, Jowett T. Cloning of the zebrafish *krox-20* gene (*krx-20*) and its expression during hindbrain development. *Nucleic Acids Res* **21**, 1087-1095 (1993).
256. Walker MB, Kimmel CB. A two-color acid-free cartilage and bone stain for zebrafish larvae. *Biotech Histochem* **82**, 23-28 (2007).

## ANNEXES

- Bercier V, Rosello M, Del Bene F, Revenu C. Zebrafish as a Model for the Study of Live in vivo Processive Transport in Neurons. *Front Cell Dev Biol* **7**, 17 (2019).
- Desban L, Prendergast A, Roussel J, Rosello M, Geny D, Wyart C, Bardet PL. Regulation of the apical extension morphogenesis tunes the mechanosensory response of microvilliated neurons. *PLoS Biol* **17**, e3000235 (2019).
- Revenu C, Parlato M\*, Rosello M\*, Durore K, Duclaux-Loras R, Nicolle O, Prospéri MT, Stoufflet J, Vouigny J, Lebreton C, Lépine P, Michaux G, Cerf-Benssusan N, Coudrier E, Del Bene F. Myosin-1b interacts with UNC45A and controls intestinal epithelial morphogenesis. *bioRxiv*, 2021.2009.2009.459609 (2021).



# Zebrafish as a Model for the Study of Live *in vivo* Processive Transport in Neurons

Valérie Bercier<sup>1,2</sup>, Marion Rosello<sup>1</sup>, Filippo Del Bene<sup>1\*</sup> and Céline Revenu<sup>1\*</sup>

<sup>1</sup> Institut Curie, PSL Research University, Inserm U934, CNRS UMR3215, Paris, France, <sup>2</sup> Laboratory of Neurobiology, Center for Brain and Disease Research, Research Group Experimental Neurology, Department of Neurosciences, VIB-KU Leuven, Leuven, Belgium

## OPEN ACCESS

### Edited by:

Hongyan Wang,  
Duke-NUS Medical School,  
Singapore

### Reviewed by:

Gary Alan Barclay Armstrong,  
Montreal Neurological Institute, McGill  
University, Canada

Aniket Gore,  
National Institutes of Health (NIH),  
United States

### \*Correspondence:

Filippo Del Bene  
filippo.del-bene@inserm.fr  
Céline Revenu  
celine.revenu@curie.fr

### Specialty section:

This article was submitted to  
Cell Growth and Division,  
a section of the journal  
Frontiers in Cell and Developmental  
Biology

**Received:** 03 October 2018

**Accepted:** 01 February 2019

**Published:** 19 February 2019

### Citation:

Bercier V, Rosello M, Del Bene F  
and Revenu C (2019) Zebrafish as  
a Model for the Study of Live *in vivo*  
Processive Transport in Neurons.  
*Front. Cell Dev. Biol.* 7:17.  
doi: 10.3389/fcell.2019.00017

Motor proteins are responsible for transport of vesicles and organelles within the cell cytoplasm. They interact with the actin cytoskeleton and with microtubules to ensure communication and supply throughout the cell. Much work has been done *in vitro* and *in silico* to unravel the key players, including the dynein motor complex, the kinesin and myosin superfamilies, and their interacting regulatory complexes, but there is a clear need for *in vivo* data as recent evidence suggests previous models might not recapitulate physiological conditions. The zebrafish embryo provides an excellent system to study these processes in intact animals due to the ease of genetic manipulation and the optical transparency allowing live imaging. We present here the advantages of the zebrafish embryo as a system to study live *in vivo* processive transport in neurons and provide technical recommendations for successful analysis.

**Keywords:** neuronal transport, zebrafish, myosin, kinesin, dynein, *in vivo*

## INTRODUCTION

Processive intracellular transport is essential for the distribution of organelles and cellular cargoes within the cell. In the case of neurons, such transport provides communication between different cell compartments and ensures supply to the growing synapse, clearance of detritus and serves as the support for intracellular signaling (Hirokawa et al., 2010). This process relies on the function of motor proteins and their interaction with the cell cytoskeleton, the three components of which play important roles in regulating transport. Indeed, neurofilaments provide structure and regulate axonal caliber, which influences transport metrics. Microtubules are responsible for axonal polarity, a consequence of the stereotyped orientation of their dynamic fast-growing ends, and act as the rails guiding motor proteins within the axon and dendrites. Finally, actin filaments form a structural network supporting the growth cone, pre- and postsynaptic regions, and play an important role in dendrites where they form the spines, essential for synaptic transmission.

Dynein and the kinesin superfamily are the unidirectional molecular motors responsible for transport on microtubules, both in dendrites and axon. In polarized axons, it is split according to the direction relative to the microtubule fast-growing end (+), with kinesins being responsible for ‘anterograde’ movement (toward the synapse) and dynein for ‘retrograde’ movement (toward the cell body). Their movement can in turn be categorized as ‘slow’ or ‘fast’ depending on their transport rate. Slow axonal transport is mainly used for delivery of cytoskeletal components and associated proteins, with kinetics in the range of 0.2–8 mm/day (Lasek et al., 1984). Fast axonal transport is used for organelles and vesicles, but also for mRNA granules (Maday et al., 2014), with kinetics in

the range of 50–400 mm/day (Lasek et al., 1984). This type of transport, in terms of motor complex involved for different types of cargoes, motor adaptor complexes and transport metrics, has been widely studied and is well reviewed elsewhere (Maday et al., 2014).

Unconventional myosins, molecular motors of the actin cytoskeleton, are commonly associated with dynamic shaping of membranes, as well as organelle formation and transport, but their functions in neuronal transport is not well understood. Among this super-family, Myosin5a, 5b, 6 and 10 have been identified as processive transporters in neurons, and participate in local transport of intracellular cargoes over short-range distances. Processive myosins are likely part of a cooperative mechanism which is based on the coordination of actin and microtubule transporters (Wu et al., 2000). This is nicely illustrated with Myo5a binding directly to kinesins (Huang et al., 1999), suggesting that organelles transported in axons along microtubules may be transported by Myo5a in presynaptic terminals, which lack microtubules (Wu et al., 1998; Bridgman, 1999; Lalli et al., 2003; Nalavadi et al., 2012). Processive myosins are implicated in vesicle endocytosis, recycling and exocytosis, and hence participate in receptor transport and localization, regulating neuronal signaling and axonal pathfinding (Wu et al., 2002; Osterweil et al., 2005; Zhu et al., 2007; Correia et al., 2008; Nash et al., 2010; Lazo et al., 2013; Sui et al., 2015). Moreover, processive myosins take part in transport of mRNAs and RNPs in neurons, as demonstrated for Myo5a (Ohashi et al., 2002; Yoshimura et al., 2006; Balasanyan and Arnold, 2014; Calliari et al., 2014; Lindsay and Mccaffrey, 2014).

From recent evidence, it is apparent that *in vivo* axonal transport data do not recapitulate what has been observed *in vitro* (Gibbs et al., 2016; Klinman and Holzbaur, 2016; Knabbe et al., 2018), emphasizing the need for a more physiological context. With this in mind, excellent work has been published reporting *in vivo* axonal transport (reviewed in Sleight et al., 2017) in models such as the *Drosophila* wing (Vagnoni and Bullock, 2017) and larvae (Vukoja et al., 2018), as well as the mouse brain (Knabbe et al., 2018) and sciatic nerve (Gibbs et al., 2016). All of these models have advantages and drawbacks: the mouse model is widely used and as a mammal, has a high genetic conservation of genes of interest but is not translucent and only allows access to axonal transport in a restricted area of the targeted cell population by way of surgery. The *Drosophila* is a model with a fantastic genetic manipulation toolbox, however, it is an invertebrate with reduced conservation to human compared to vertebrate models.

Over the last decades, zebrafish has emerged as a powerful vertebrate model to study the development of the nervous system *in vivo*. Adult zebrafish are small in size and produce a large number of offsprings, with a rapid external development. The embryonic zebrafish are translucent, and recent advances in genetic manipulation have made this model a great option to monitor neurodevelopment by high-resolution live imaging and at single-cell level. In addition, the zebrafish embryo is used extensively for modeling neurodegeneration (Bandmann and Burton, 2010; Kabashi et al., 2010; Santoriello and Zon, 2012; Babin et al., 2014; Patten et al., 2014; Fontana et al., 2018). Some processive motors have been associated with neurological

disorders (Chen et al., 2013) and many studies have reported axonal transport defects in the context of neurodegenerative diseases (Chevalier-Larsen and Holzbaur, 2006; Goldstein, 2012; Liu et al., 2012; Millecamps and Julien, 2013), further outlining the interest of this model. In this article, we thus discuss the advantages of the zebrafish model in the study of live *in vivo* intracellular transport, with a particular focus on fast axonal transport.

## ADVANTAGES OF THE ZEBRAFISH MODEL

### Relevance to Mammalian Models

The genome of *Danio rerio* is fully sequenced and presents at least one ortholog for 70% of human genes (Howe et al., 2013). In particular, kinesin, dynein and myosin molecular motors implicated in neuronal transports are extremely well conserved in eukaryotes and even more in vertebrates (Kim and Endow, 2000; Sittaramane and Chandrasekhar, 2008). These proteins have a higher conservation with the human ortholog in zebrafish compared to *D. melanogaster* for example. Zebrafish and drosophila dynein Dync1h1 show 91% and 72% identity (NCBI Blastp) with the human protein, respectively. Similarly, the processive Myo6 is 85% and 53% identical to the human one in zebrafish and drosophila, respectively. This high degree of conservation provides support for using zebrafish as a model system to investigate the functions of these molecular motors.

### Genetic Manipulations

Compared especially to the mouse, the ease of stable or transient genetic manipulations has positioned the zebrafish as an ideal vertebrate model for live *in vivo* imaging.

Transgenesis in zebrafish is routinely and efficiently performed to express fusion proteins, mutated proteins or the *gal4* transcription factor under a tissue-specific promoter thanks to the use of transposon elements. Ease of genetic manipulations in zebrafish has tremendously increased with the development of the CRISPR/Cas9 technology. The generation of knock-out mutants has become extremely powerful (Hwang et al., 2013) and using a Gal4/UAS-based restriction of Cas9 expression makes it possible to induce tissue-specific mutations and restrict the phenotype to a subset of cells (Di Donato et al., 2016). Recent advances based on the fusion of a mutated Cas9 (nickase) with an acetyl deaminase leading to the precise editing of a single nucleic acid (Komor et al., 2016) was also shown to work in zebrafish (Zhang et al., 2017). This technology makes it possible to target a specific protein domain in order to interfere with protein–protein interaction and opens the possibility of reproducing mutations associated with human diseases to elucidate the underlying pathological mechanism.

To recapitulate endogenous expression of a protein of interest, both in terms of pattern and level, bacterial artificial chromosome (BAC) transgenesis, where very large DNA sequence (up to 300 kb) can be inserted into the genome, is used in zebrafish (Lee et al., 2001; Suster et al., 2011). The CRISPR/Cas9 era has now opened the possibility of direct knock-in at a targeted

locus. This strategy has been successful in zebrafish, based on the error-prone non-homologous end-joining DNA damage repair mechanism (Auer et al., 2014) and by short or long homology arm recombination (Hruscha et al., 2013; Hwang et al., 2013; Irion et al., 2014; Zhang et al., 2016). However, the efficiency of the latter technique is still low and locus-dependant. Its optimization will be an important technical advance in the field (Albadri et al., 2017), for example, to allow endogenous expression of a fusion protein of choice for visualization *in vivo*, without overexpression.

## Pharmacological Manipulations

Zebrafish embryos are amenable to pharmacological treatment by bath application, allowing for treatment of intact, live embryos with compounds known for the modulation of cytoskeletal dynamics, for instance, targeting microtubules: Colchicine (Roche et al., 1994), vinblastine (Keiter et al., 2016; Yao et al., 2017), vincristine (Mizgirev and Revskoy, 2010; Khan et al., 2012; Holloway et al., 2016), nocodazole (Plucińska et al., 2012; Jayachandran et al., 2016) and paclitaxel (Jayachandran et al., 2016). For the actin cytoskeleton: Cytochalasin D (Nukada et al., 2015; Artelt et al., 2018) and latrunculin A (Artelt et al., 2018), jasplakinolide (Artelt et al., 2018), phalloidin oleate (Dutta and Kumar Sinha, 2015), and the inhibitor of actin-myosin interaction BDM (Norden et al., 2009) have been used with success.

Finally, zebrafish embryos are well suited to high-throughput approaches that have made them an excellent tool in drug discovery by small molecule screening (Zon and Peterson, 2005; Mathias et al., 2012; Miscovic et al., 2012; Tamplin et al., 2012; MacRae and Peterson, 2015).

## EXAMPLES AND RECOMMENDATIONS FOR THE ANALYSIS OF *IN VIVO* TRANSPORT IN ZEBRAFISH

To date, a few studies have taken advantage of the zebrafish model to perform *in vivo* axonal transport assays, generating tools to study the movement of mitochondria (Plucińska et al., 2012; Campbell et al., 2014; Paquet et al., 2014; Auer et al., 2015; Drerup et al., 2017), endosomes (Clark et al., 2011; Ponomareva et al., 2014, 2016), autophagosomes (He et al., 2009), lysosomes (Drerup and Nechiporuk, 2013), synaptophysin-containing vesicles (Auer et al., 2015) as well as motor proteins and components of their regulatory complexes (Drerup and Nechiporuk, 2013, 2016). The *in vivo* analysis of myosin-based transport is only starting in zebrafish neuronal development (Liu et al., 2013).

Based on published evidence, it is plain to see that the metrics reported for the same cargo visualized *in vivo* in zebrafish display variation between cell types and developmental stages. Indeed, we have observed metrics for mitochondrial anterograde transport in primary motor neurons (MN; axon) and in retinal ganglion cells (RGC; arbor) and while we did not find differences in average run speed, average run length and duration were significantly different in these two

cell types. Furthermore, the average run speed detected was approximately 0.4  $\mu\text{m/s}$  (Figure 1C), which is consistent with reported data from Campbell et al. (2014) (peripheral sensory neuron arbors, approx. 0.4  $\mu\text{m/s}$ ) but inconsistent with data from Plucińska et al. (2012) (peripheral Rohon-Beard sensory neuron axons, approx. 1.2  $\mu\text{m/s}$  'moving speed' and 0.6  $\mu\text{m/s}$  'average speed') and from Drerup et al. (2017) (peripheral lateral line axon, approx. 1.0  $\mu\text{m/s}$ ). We also found discrepancies between cell types in the transport of recycling endosomes (labeled with Rab11a-GFP), where we observed an average speed of approx. 0.5  $\mu\text{m/s}$  (Figure 2B), whereas Ponomareva et al. (2014) report an average speed of approx. 0.18  $\mu\text{m/s}$ /0.03  $\mu\text{m/s}$  (central/peripheral Rohon-Beard sensory neuron axon).

Based on the evidence above outlining the variability of these processes, we will highlight a few key points to take into consideration when designing experiments to characterize transport in zebrafish neurons.

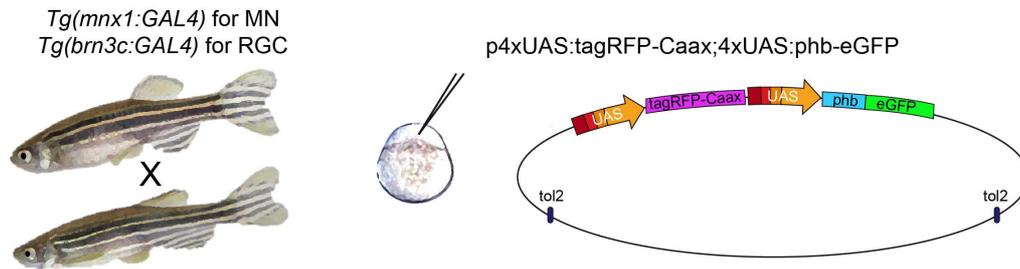
## Regulation of Construct Expression and Imaging

Most approaches discussed here rely on the overexpression of fusion proteins, allowing *in vivo* detection of the bound fluorescent protein. This can be achieved injecting DNA constructs to obtain single-cell labeling of cargoes, as shown here (Figure 1A), or creating stable transgenic lines, where restriction of expression can be achieved by a combination of Gal4- and UAS-expressing lines. While this technique produces a bright signal well suited to time-lapse imaging, overexpression of protein can lead to deleterious effects by interfering with endogenous expression and triggering stress response mechanisms (Cheng and Lee, 2010). It is therefore essential to ensure that the construct does not lead to toxicity by monitoring cell morphology and embryonic development. The acquisition parameters in time-lapse microscopy are optimized to limit bleaching of fluorescent proteins and damage of the target cell, while still observing the target movement (for instance: high frequency sampling but reduced duration). In the case of the examples presented here, time-lapse imaging of labeled cargo in neurons was performed at 2 Hz for endosomes (5 min duration; Figure 2B) and at 1 Hz for mitochondria (10 min duration; Figures 1B,C) on a spinning disk confocal microscope.

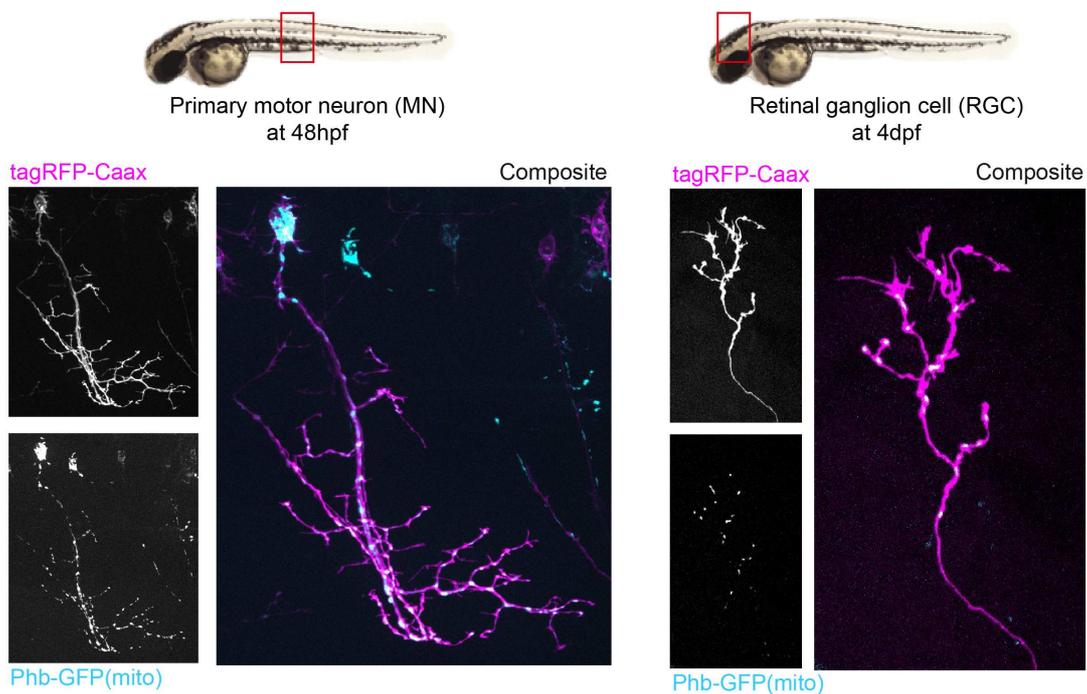
## Single Cell Type

Axonal transport dynamics can be influenced by the varying expression of subunits composing the motor protein complexes or particular cargo adapters, as well as by the axon caliber, due to differences in microtubule density affecting engagement of motors (Yu et al., 2017) and due to activity- and myelination-dependent number of neurofilaments (de Waegh et al., 1992). It is therefore recommended to target one cell type (Figures 1B,C), and in the case of spinal cord neurons, to limit observation to a specific region as cell size can fluctuate along the trunk and tail owing to the rostral-caudal developmental wave.

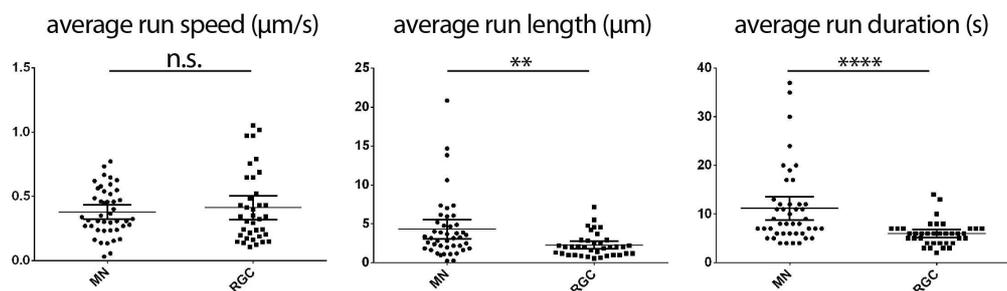
### A Injection of construct in a restrictive genetic background



### B Single-cell selection



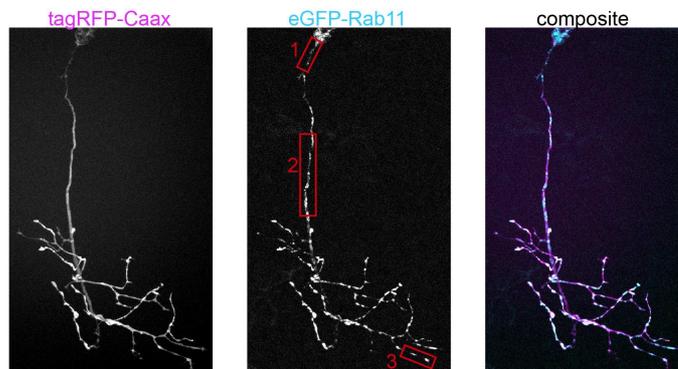
### C Kymogram analysis



**FIGURE 1** | Construct expression and cell type selection. **(A)** Injection of DNA constructs coding for fusion proteins can be restricted to a single cell type by use of the Gal4/UAS system. **(B)** In this example, we injected a UAS construct labeling mitochondria (phb, prohibitin-GFP see schematics in **A**) combined with a membrane reporter (tagRFP-Caax). We obtained labeling of a single primary motor neuron (MN; in the *Tg(mnx1:gal4)* background) and a single retinal ganglion cell (RGC; in the *Tg(brn3c:gal4)* background), respectively, in the embryonic spinal cord (48 hpf) and in the larval optic tectum (4 dpf). **(C)** Time-lapse imaging of mitochondria (1 Hz for 10 min) was performed on these cell types, and transport dynamics were calculated from kymograms. Here, we show example of the disparity in transport metrics that can arise when comparing different cell types for a single cargo (MN  $n = 7$  cells/44 anterograde runs; RGC  $n = 7$  cells/37 anterograde runs). \*\* $p < 0.01$ , \*\*\*\* $p < 0.0001$ .

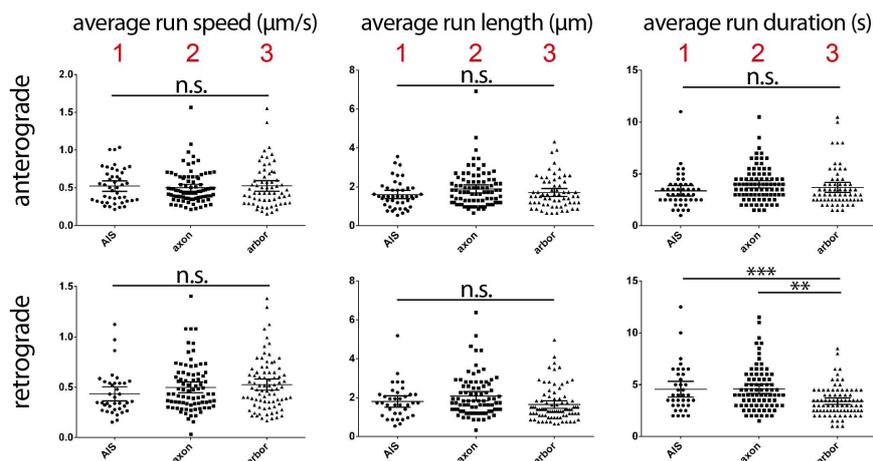
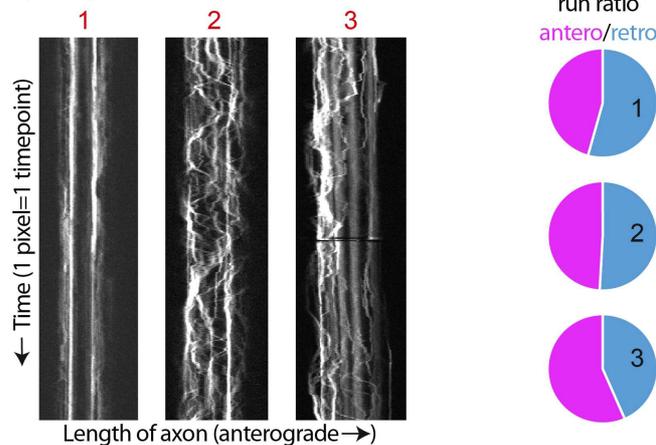
### A Timelapse imaging

Single-cell labeling of a caudal primary motor neuron



*Tg(mnx1:GAL4) + p4xUAS:tagRFP-Caax;4xUAS:eGFP-Rab11*

### B Kymogram analysis



**FIGURE 2 |** Examples of time-lapse analysis. **(A)** As described in **Figure 1A**, single cell labeling of primary motor neurons was obtained for recycling endosomes (Rab11a-eGFP), combined with a membrane reporter (tagRFP-Caax) to identify cell type. Red boxes: Three cell compartments were imaged (2 Hz for 5 min), 1-axon initial segment, 2-mid-axonal segment, 3-axonal arbor segment. **(B)** Kymograms were generated from the time-lapses acquired (Kymograph tool, ImageJ) and a variety of transport metrics can be calculated manually (compiled in Excel, statistics in Graphpad Prism6). In this example, significant differences between neuronal segments are detected for the transport direction ratio (anterograde/retrograde), and retrograde run duration ( $n = 3$  cells; AIS  $n = 47/54$  anterograde/retrograde runs; mid-axon  $n = 85/82$  anterograde/retrograde runs; arbor  $n = 63/86$  anterograde/retrograde runs).  $**p < 0.01$ ,  $***p < 0.001$ .

## Single Cell Compartment

Different cell compartments have different cytoskeletal composition and regulatory mechanisms that can affect the composition and expression of motor proteins and their adaptor proteins, hence influencing the regulation of axonal transport (de Waegh et al., 1992; Yu et al., 2017; Gummy and Hoogenraad, 2018). As shown here (Figure 2), time-lapse imaging in different compartments of the same neuron for the same cargo can yield significant differences in some metrics of transport dynamics, but not others. The selection of a compartment most suited to the research hypothesis and consistency in the segment imaged across individual embryos and larva appears thus crucial, as well as the analysis of a variety of metrics such as run length and duration, average speed, pause frequency, average pause duration, switching behavior, area flux and transport rates of cargoes, both in the anterograde and retrograde directions.

## PERSPECTIVES FOR FUTURE RESEARCH

### Alternative Labeling

Nanoparticles are inorganic semiconductors representing an attractive alternative for fluorescent labeling in live imaging applications because of their high spatial resolution and photostability. In addition, it is possible to tune their emission wavelength by varying their size and chemical composition. Because of this, and their broad absorption profile, it is possible to excite multiple colors at once, which is useful to reduce sample phototoxicity (Gao et al., 2005). In contrast to genetically encoded fluorescent protein tags, however, they need to be efficiently targeted to their biologically relevant endpoint. This has so far relied on surface modifications and solubilization strategies that led to very large particles better suited to high-sensitivity detection of low number of targets, such as single-molecule detection (Pinaud et al., 2006). Of note, this approach has allowed for real-time visualization of single-molecules in living cells (Dahan et al., 2003). Conjugation to biomolecules is, however, an interesting avenue to allow precise targeting and while still requiring the expression of a genetically encoded protein adaptor (Gao et al., 2005; Howarth et al., 2005), would provide the advantageous optical properties of nanoparticles over traditional fluorescent proteins.

### Microscopy Improvements

Advances in imaging technology in the last years have yielded many optimized systems applicable to the study of *in vivo* axonal transport in the zebrafish embryo. Indeed, a great example of this is the swept field confocal microscope, which permits higher frame-rate capture when compared with spinning disk confocal and allows the rapid acquisition of z-stack time-lapses or high speed imaging (upward of 1,000 fps) of movement in single-plane (Castellano-Muñoz et al., 2012). Other systems circumvent classic caveats to an *in vivo* approach, such as photodamage, single-plan and temporal restriction and low signal, for instance: 2-photon microscopy (Renninger and Orger, 2013), light-sheet

microscopy (Huisken et al., 2004; Panier et al., 2013; Park et al., 2015; Tomer et al., 2015; Fu et al., 2016). In the context of *in vivo* imaging of axonal transport, these strategies could allow the tracking of cargo and motors with exquisite temporal resolution, while also permitting 3D tracking in a whole embryo over long periods of time; considerable advantages over *in vitro* and other *in vivo* models.

### Automated Detection and Analysis

The generation of kymograms as a 2D representation of time-lapse imaging is a common tool for the analysis of axonal transport, where the tracked target often moves on a single focal plane, in a linear trajectory. When analyzing movement in more complex environment, however, single particle tracking becomes a necessity, which renders manual analysis an arduous task. In the past years, many options have become available for automated detection and tracking, both commercially (Imaris, Metamorph, Igor Pro, etc.) and *via* open-source programs (MATLAB, ImageJ, etc.). Still, time-lapse videos obtained *in vivo* from intact animals are often noisier by nature than their cell culture counterpart, and since these samples are prone to photodamage, lead to undersampled data. This in turn impedes automatic detection and requires manual check of extracted metrics, while possibly omitting crucial information. Further advances in detection algorithms, based on *in vivo* data estimating how cargoes should behave, will surely be of benefit to researchers facing the tedious task of manual tracking.

## CONCLUSION

The zebrafish embryo has emerged as an excellent model to pursue the characterization of processive transport *in vivo* as it can meet the need for more inclusive models, where the contribution of neuronal activity, glia and the cell cytoskeleton are taken into account. We outlined here some advantages and technical hints to use the zebrafish model for this type of analysis. Considering the recent breakthroughs in genetic manipulations and imaging technologies, this vertebrate is gaining attention in the field of neurodegenerative disease modeling, where axonal transport deficits are common hallmarks. In addition, a new emerging model sharing the same subfamily as zebrafish, *Danio rerio*, which remains transparent throughout its life, will further expand the possibilities of adult neuronal imaging *in vivo* (Schulze et al., 2018). It is thus only a matter of time before axonal transport assays in zebrafish embryos become widespread for the study of physiological and pathological conditions.

## ETHICS STATEMENT

As the EU directive 2010\_63 explicitly states only “independently feeding larval forms” must be classified as animal experiments, therefore only zebrafish larvae past 120 h post fertilization should be subject to the regulations of European animal protection guidelines. For our experiments, we did not use larvae that

have reached an “independent feeding” stage and therefore we did not have to submit an ethical approval to the competent local/national ethical/legal bodies.

## DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this manuscript will be made available by the authors, without undue reservation, to any qualified researcher.

## AUTHOR CONTRIBUTIONS

All authors listed have made a substantial, direct and intellectual contribution to the work, and approved it for publication.

## REFERENCES

- Albadri, S., Del Bene, F., and Revenu, C. (2017). Genome editing using CRISPR/Cas9-based knock-in approaches in zebrafish. *Methods* 121–122, 77–85. doi: 10.1016/j.ymeth.2017.03.005
- Artelt, N., Ludwig, T. A., Rogge, H., Kavvadas, P., Siegerist, F., Blumenthal, A., et al. (2018). The role of palladin in podocytes. *J. Am. Soc. Nephrol.* 29, 1662–1678. doi: 10.1681/ASN.2017091039
- Auer, T. O., Duroure, K., De Cian, A., Concordet, J., and Del Bene, F. (2014). Highly efficient CRISPR / Cas9-mediated knock-in in zebrafish by homology-independent DNA repair. *Genome Res.* 24, 142–153. doi: 10.1101/gr.161638.113
- Auer, T. O., Xiao, T., Bercier, V., Gebhardt, C., Duroure, K., Concordet, J.-P., et al. (2015). Deletion of a kinesin I motor unmasks a mechanism of homeostatic branching control by neurotrophin-3. *eLife* 4:e05061. doi: 10.7554/eLife.05061
- Babin, P. J., Goizet, C., and Raldu, D. (2014). Zebrafish models of human motor neuron diseases: advantages and limitations. *Prog. Neurobiol.* 118, 36–58. doi: 10.1016/j.pneurobio.2014.03.001
- Balasyan, V., and Arnold, D. B. (2014). Actin and myosin-dependent localization of mRNA to dendrites. *PLoS One* 9:e92349. doi: 10.1371/journal.pone.0092349
- Bandmann, O., and Burton, E. A. (2010). Genetic zebrafish models of neurodegenerative diseases. *Neurobiol. Dis.* 40, 58–65. doi: 10.1016/j.nbd.2010.05.017
- Bridgman, P. C. (1999). Myosin VA movements in normal and *Dilute-Lethal* axons provide support for a dual filament motor complex. *J. Cell Biol.* 146, 1045–1060. doi: 10.1083/jcb.146.5.1045
- Calliari, A., Farias, J., Puppo, A., Canclini, L., Mercer, J. A., Munroe, D., et al. (2014). Myosin Va associates with mRNA in ribonucleoprotein particles present in myelinated peripheral axons and in the central nervous system. *Dev. Neurobiol.* 74, 382–396. doi: 10.1002/dneu.22155
- Campbell, P. D., Shen, K., Sapio, M. R., Glenn, T. D., Talbot, W. S., and Marlow, F. L. (2014). Unique function of Kinesin Kif5A in localization of mitochondria in axons. *J. Neurosci.* 34, 14717–14732. doi: 10.1523/JNEUROSCI.2770-14.2014
- Castellano-Muñoz, M., Peng, A. W., Salles, F. T., and Ricci, A. J. (2012). Swept field laser confocal microscopy for enhanced spatial and temporal resolution in live-cell imaging. *Microsc. Microanal.* 18, 753–760. doi: 10.1017/S1431927612000542
- Chen, Y., Tian, L., Zhang, F., Liu, C., Lu, T., Ruan, Y., et al. (2013). Myosin Vb gene is associated with schizophrenia in Chinese Han population. *Psychiatry Res.* 207, 13–18. doi: 10.1016/j.psychres.2013.02.026
- Cheng, C.-H., and Lee, W.-C. (2010). Protein solubility and differential proteomic profiling of recombinant *Escherichia coli* overexpressing double-tagged fusion proteins. *Microb. Cell Fact.* 9:63. doi: 10.1186/1475-2859-9-63
- Chevalier-Larsen, E., and Holzbaur, E. L. F. (2006). Axonal transport and neurodegenerative disease. *Biochim. Biophys. Acta* 1762, 1094–1108. doi: 10.1016/j.bbadis.2006.04.002
- Clark, B., Winter, M., Cohen, A., and Link, B. (2011). Generation of Rab-based transgenic lines for in vivo studies of endosome biology in zebrafish. *Dev. Dyn.* 240, 2452–2465. doi: 10.1002/dvdy.22758
- Correia, S. S., Bassani, S., Brown, T. C., Lisé, M. F., Backos, D. S., El-Husseini, A., et al. (2008). Motor protein-dependent transport of AMPA receptors into spines during long-term potentiation. *Nat. Neurosci.* 11, 457–466. doi: 10.1038/nn2063
- Dahan, M., Le, S., and Luccardini, C. (2003). Diffusion dynamics of glycine receptors revealed by single – quantum dot tracking. *Science* 302, 442–445. doi: 10.1126/science.1088525
- de Waegh, S. M., Lee, V. M. Y., and Brady, S. T. (1992). Local modulation of neurofilament phosphorylation, axonal caliber, and slow axonal transport by myelinating Schwann cells. *Cell* 68, 451–463. doi: 10.1016/0092-8674(92)90183-D
- Di Donato, V., De Santis, F., Auer, T. O., Testa, N., Sánchez-iranzo, H., Mercader, N., et al. (2016). 2C-Cas9: a versatile tool for clonal analysis of gene function. *Genome Res.* 26, 681–692. doi: 10.1101/gr.196170.115
- Drerup, C. M., Herbert, A. L., Monk, K. R., and Nechiporuk, A. V. (2017). Regulation of mitochondria-dynactin interaction and mitochondrial retrograde transport in axons. *eLife* 6:e22234. doi: 10.7554/eLife.22234
- Drerup, C. M., and Nechiporuk, A. V. (2013). JNK-interacting protein 3 mediates the retrograde transport of activated c-Jun N-terminal kinase and lysosomes. *PLoS Genet.* 9:e1003303. doi: 10.1371/journal.pgen.1003303
- Drerup, C. M., and Nechiporuk, A. V. (2016). In vivo analysis of axonal transport in zebrafish. *Methods Cell Biol.* 131, 311–329. doi: 10.1016/bs.mcb.2015.06.007
- Dutta, A., and Kumar Sinha, D. (2015). Turnover of the actomyosin complex in zebrafish embryos directs geometric remodeling and the recruitment of lipid droplets. *Sci. Rep.* 5:13915. doi: 10.1038/srep13915
- Fontana, B. D., Mezzomo, N. J., Kalueff, A. V., and Roseberg, D. B. (2018). The developing utility of zebrafish models of neurological and neuropsychiatric disorders: a critical review. *Exp. Neurol.* 299, 157–171. doi: 10.1016/j.expneurol.2017.10.004
- Fu, Q., Martin, B. L., Matus, D. Q., and Gao, L. (2016). Imaging multicellular specimens with real-time optimized tiling light-sheet selective plane illumination microscopy. *Nat. Commun.* 7:11088. doi: 10.1038/ncomms11088
- Gao, X., Yang, L., Petros, J. A., Marshall, F. F., Simons, J. W., and Nie, S. (2005). In vivo molecular and cellular imaging with quantum dots. *Curr. Opin. Biotechnol.* 16, 63–72. doi: 10.1016/j.copbio.2004.11.003
- Gibbs, K. L., Kalmar, B., Sleight, J. N., Greensmith, L., and Schiavo, G. (2016). In vivo imaging of axonal transport in murine motor and sensory neurons. *J. Neurosci. Methods* 257, 26–33. doi: 10.1016/j.jneumeth.2015.09.018
- Goldstein, L. S. B. (2012). Axonal transport and neurodegenerative disease: can we see the elephant? *Prog. Neurobiol.* 99, 186–190. doi: 10.1016/j.pneurobio.2012.03.006
- Gumy, L. F., and Hoogenraad, C. C. (2018). Local mechanisms regulating selective cargo entry and long-range trafficking in axons. *Curr. Opin. Neurobiol.* 51, 23–28. doi: 10.1016/j.conb.2018.02.007

## FUNDING

The Del Bene laboratory “Neural Circuits Development” is part of the Laboratoire d’Excellence (LABEX) entitled DEEP (ANR-11-LABX-0044), and of the École des Neurosciences de Paris Ile-de-France network. VB was supported by an FRSQ and CIHR Doctoral Award and was enrolled in the ENP Graduate Program. MR was supported by the Fondation pour la Recherche Médicale (FRM grant number ECO20170637481). CR was supported by a EU H2020 Marie Skłodowska-Curie Action fellowship (H2020-MSCA-IF-2014 #661527). This work has been supported by an ATIP/AVENIR program starting grant (FDB), ERC-StG #311159 (FDB), ERA-NET E-rare grant (for Research Programs on Rare Diseases, 0601165051) (FDB).

- He, C., Bartholomew, C., Zhou, W., and Klionsky, D. (2009). Assaying autophagic activity in transgenic GFP-Lc3 and GFP- Gabarap zebrafish embryos. *Autophagy* 5, 520–526. doi: 10.4161/auto.5.4.7768
- Hirokawa, N., Niwa, S., and Tanaka, Y. (2010). Molecular motors in neurons: transport mechanisms and roles in brain function, development, and disease. *Neuron* 68, 610–638. doi: 10.1016/j.neuron.2010.09.039
- Holloway, M. P., Denardo, B. D., Phornphutkul, C., Nguyen, K., Davis, C., Jackson, C., et al. (2016). An asymptomatic mutation complicating severe chemotherapy-induced peripheral neuropathy (CIPN): a case for personalised medicine and a zebrafish model of CIPN. *NPJ Genom. Med.* 1:16016. doi: 10.1038/npgjgenmed.2016.16
- Howarth, M., Takao, K., Hayashi, Y., and Ting, A. Y. (2005). Targeting quantum dots to surface proteins in living cells with biotin ligase. *Proc. Natl. Acad. Sci. U.S.A.* 102, 7583–7588. doi: 10.1073/pnas.0503125102
- Howe, K., Clark, M., Torroja, C., Torrance, J., Berthelot, C., Muffato, M., et al. (2013). The zebrafish reference genome sequence and its relationship to the human genome. *Nature* 496, 498–503. doi: 10.1038/nature12111
- Hruscha, A., Krawitz, P., Rechenberg, A., Heinrich, V., Hecht, J., Haass, C., et al. (2013). Efficient CRISPR/Cas9 genome editing with low off-target effects in zebrafish. *Development* 140, 4982–4987. doi: 10.1242/dev.099085
- Huang, J. D., Brady, S. T., Richards, B. W., Stenoien, D., Resau, J. H., Copeland, N. G., et al. (1999). Direct interaction of microtubule- and actin based transport motors. *Nature* 397, 267–270. doi: 10.1038/16722
- Huisken, J., Swoger, J., Del Bene, F., Wittbrodt, J., and Stelzer, E. H. K. (2004). Optical sectioning deep inside live embryos by selective plane illumination microscopy. *Science* 305, 1007–1009. doi: 10.1126/science.1100035
- Hwang, W. Y., Fu, Y., Reyon, D., Maeder, M. L., Shengdar, Q., Sander, J. D., et al. (2013). Efficient *In Vivo* genome editing using RNA-guided nucleases. *Nat. Biotechnol.* 31, 227–229. doi: 10.1038/nbt.2501
- Irion, U., Krauss, J., and Nusslein-Volhard, C. (2014). Precise and efficient genome editing in zebrafish using the CRISPR/Cas9 system. *Development* 141, 4827–4830. doi: 10.1242/dev.115584
- Jayachandran, P., Olmo, V. N., Sanchez, S. P., McFarland, R. J., Vital, E., Werner, J. M., et al. (2016). Microtubule-associated protein 1b is required for shaping the neural tube. *Neural Dev.* 11, 1–18. doi: 10.1186/s13064-015-0056-4
- Kabashi, E., Champagne, N., Brustein, E., and Drapeau, P. (2010). In the swim of things: recent insights to neurogenetic disorders from zebrafish. *Trends Genet.* 26, 373–381. doi: 10.1016/j.tig.2010.05.004
- Keiter, S., Burkhard-Medicke, K., Wellner, P., Kais, B., Färber, H., Skutlarek, D., et al. (2016). Does perfluorooctane sulfonate (PFOS) act as chemosensitizer in zebrafish embryos? *Sci. Total Environ.* 54, 317–324. doi: 10.1016/j.scitotenv.2015.12.089
- Khan, T. M., Benaich, N., Malone, C. F., Bernardos, R. L., Russell, A. R., Downes, G. B., et al. (2012). Vincristine and bortezomib cause axon outgrowth and behavioral defects in larval zebrafish. *J. Peripher. Nerv. Syst.* 17, 76–89. doi: 10.1111/j.1529-8027.2012.00371.x
- Kim, A. J., and Endow, S. A. (2000). A kinesin family tree. *J. Cell Sci.* 113, 3681–3682.
- Klinman, E., and Holzbaur, E. L. F. (2016). Comparative analysis of axonal transport markers in primary mammalian neurons. *Methods Cell Biol.* 131, 409–424. doi: 10.1016/bs.mcb.2015.06.011
- Knabbe, J., Nassal, J. P., Verhage, M., and Kuner, T. (2018). Secretory vesicle trafficking in awake and anesthetized mice: differential speeds in axons versus synapses. *J. Physiol.* 596, 3759–3773. doi: 10.1111/JP276022
- Komor, A. C., Kim, Y. B., Packer, M. S., Zuris, J. A., and Liu, D. R. (2016). Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage. *Nature* 533, 420–424. doi: 10.1038/nature17946
- Lalli, G., Gschmeissner, S., and Schiavo, G. (2003). Myosin Va and microtubule-based motors are required for fast axonal retrograde transport of tetanus toxin in motor neurons. *J. Cell Sci.* 116, 4639–4650. doi: 10.1242/jcs.00727
- Lasek, R. J., Garner, J. A., and Brady, S. T. (1984). Axonal transport of the cytoplasmic matrix. *J. Cell Biol.* 99, 212s–221s. doi: 10.1083/jcb.99.1.212s
- Lazo, O. M., Gonzalez, A., Ascano, M., Kuruvilla, R., Couve, A., and Bronfman, F. (2013). BDNF regulates Rab11-mediated recycling endosome dynamics to induce dendritic branching. *J. Neurosci.* 33, 6112–6122. doi: 10.1523/JNEUROSCI.4630-12.2013
- Lee, E. C., Yu, D., Martinez De Velasco, J., Tessarollo, L., Swing, D. A., Court, D. L., et al. (2001). A highly efficient *Escherichia coli*-based chromosome engineering system adapted for recombinogenic targeting and subcloning of BAC DNA. *Genomics* 73, 56–65. doi: 10.1006/geno.2000.6451
- Lindsay, A. J., and Mccaffrey, M. W. (2014). Myosin Va is required for the transport of fragile X mental retardation protein (FMRP) granules. *Biol. Cell* 106, 57–71. doi: 10.1111/boc.201200076
- Liu, X., Rizzo, V., and Puthanveetil, S. (2012). Pathologies of axonal transport in neurodegenerative diseases. *Transl. Neurosci.* 3, 355–372. doi: 10.2478/s13380-012-0044-7
- Liu, Y., Xu, X. H., Chen, Q., Wang, T., Deng, C. Y., Song, B. L., et al. (2013). Myosin Vb controls biogenesis of post-Golgi Rab10 carriers during axon development. *Nat. Commun.* 4:2005. doi: 10.1038/ncomms3005
- MacRae, C. A., and Peterson, R. T. (2015). Zebrafish as tools for drug discovery. *Nat. Rev. Drug Discov.* 14, 721–731. doi: 10.1038/nrd4627
- Maday, S., Twelvetrees, A. E., Moughamian, A. J., and Holzbaur, E. L. F. (2014). Axonal transport: cargo-specific mechanisms of motility and regulation. *Neuron* 84, 292–309. doi: 10.1016/j.neuron.2014.10.019
- Mathias, J. R., Saxena, M. T., and Mumm, J. S. (2012). Advances in zebrafish chemical screening technologies. *Future Med. Chem.* 4, 1811–1822. doi: 10.4155/fmc.12.115
- Millicamps, S., and Julien, J.-P. (2013). Axonal transport deficits and neurodegenerative diseases. *Nat. Rev. Neurosci.* 14, 161–176. doi: 10.1038/nrn3380
- Miscevic, F., Rotstein, O., and Wen, X.-Y. (2012). Advances in zebrafish high content and high throughput technologies. *Comb. Chem. High Throughput Screen.* 15, 515–521. doi: 10.2174/138620712801619140
- Mizgirev, I. V., and Revskoy, S. (2010). A new zebrafish model for experimental leukemia therapy. *Cancer Biol. Ther.* 9, 895–903. doi: 10.4161/cbt.9.11.11667
- Nalavadi, V., Griffin, L., Picard-Fraser, P., Swanson, A., Takumi, T., and Bassell, G. (2012). Regulation of ZBP1 transport dynamics in axons by MyosinVa. *J. Neurosci.* 32, 15133–15141. doi: 10.1523/JNEUROSCI.2006-12.2012
- Nash, J. E., Appleby, V. J., Corrêa, S. A. L., Wu, H., Fitzjohn, S. M., Garner, C. C., et al. (2010). Disruption of the interaction between myosin VI and SAP97 is associated with a reduction in the number of AMPARs at hippocampal synapses. *J. Neurochem.* 112, 677–690. doi: 10.1111/j.1471-4159.2009.06480.x
- Norden, C., Young, S., Link, B. A., and Harris, W. A. (2009). Actomyosin is the main driver of interkinetic nuclear migration in the retina. *Cell* 138, 1195–1208. doi: 10.1016/j.cell.2009.06.032
- Nukada, Y., Horie, M., Fukui, A., Kotani, T., and Yamashita, M. (2015). Real-time imaging of actin filaments in the zebrafish oocyte and embryo. *Cytoskeleton* 72, 491–501. doi: 10.1002/cm.21253
- Ohashi, S., Koike, K., Omori, A., Ichinose, S., Ohara, S., Kobayashi, S., et al. (2002). Identification of mRNA/protein (mRNP) complexes containing Pur $\alpha$ , mStaufen, Fragile X Protein, and myosin Va and their association with rough endoplasmic reticulum equipped with a kinesin motor. *J. Biol. Chem.* 277, 37804–37810. doi: 10.1074/jbc.M203608200
- Osterweil, E., Wells, D. G., and Mooseker, M. S. (2005). A role for myosin VI in postsynaptic structure and glutamate receptor endocytosis. *J. Cell Biol.* 168, 329–338. doi: 10.1083/jcb.200410091
- Panier, T., Romano, S. A., Olive, R., Pietri, T., Sumbre, G., Candelier, R., et al. (2013). Fast functional imaging of multiple brain regions in intact zebrafish larvae using selective plane illumination microscopy. *Front. Neural Circuits* 7:65. doi: 10.3389/fncir.2013.00065
- Paquet, D., Plucińska, G., and Misgeld, T. (2014). In vivo imaging of mitochondria in intact zebrafish larvae. *Methods Enzymol.* 547, 151–164. doi: 10.1016/B978-0-12-801415-8.00009-6
- Park, O. K., Kwak, J., Jung, Y. J., Kim, Y. H., Hong, H., and Hwang, B. J. (2015). 3D light-sheet fluorescence microscopy of cranial neurons and vasculature during zebrafish embryogenesis. *Mol. Cells* 38, 975–981. doi: 10.14348/molcells.2015.0160
- Patten, S. A., Armstrong, G. A. B., Lissouba, A., Kabashi, E., Parker, J. A., and Drapeau, P. (2014). Fishing for causes and cures of motor neuron disorders. *Dis. Model. Mech.* 7, 799–809. doi: 10.1242/dmm.015719
- Pinaud, F., Michalet, X., Bentolila, L. A., Tsay, J. M., Doose, S., Li, J., et al. (2006). Advances in fluorescence imaging with quantum dot bioprobes. *Biomaterials* 27, 1679–1687. doi: 10.1016/j.biomaterials.2005.11.018
- Plucińska, G., Paquet, D., Hruscha, A., Godinho, L., Haass, C., Schmid, B., et al. (2012). In vivo imaging of disease-related mitochondrial dynamics

- in a vertebrate model system. *J. Neurosci.* 32, 16203–16212. doi: 10.1523/JNEUROSCI.1327-12.2012
- Ponomareva, O. Y., Eliceiri, K. W., and Halloran, M. C. (2016). Charcot-Marie-Tooth 2b associated Rab7 mutations cause axon growth and guidance defects during vertebrate sensory neuron development. *Neural Dev.* 11:2. doi: 10.1186/s13064-016-0058-x
- Ponomareva, O. Y., Holmen, I. C., Sperry, A. J., Eliceiri, K. W., and Halloran, M. C. (2014). Calsyntenin-1 regulates axon branching and endosomal trafficking during sensory neuron development *In Vivo. J. Neurosci.* 34, 9235–9248. doi: 10.1523/JNEUROSCI.0561-14.2014
- Renninger, S. L., and Orger, M. B. (2013). Two-photon imaging of neural population activity in zebrafish. *Methods* 62, 255–267. doi: 10.1016/j.ymeth.2013.05.016
- Roche, H., Bogé, G., and Pèrès, G. (1994). Acute and chronic toxicities of colchicine in *Brachydanio rerio*. *Bull. Environ. Contam. Toxicol.* 52, 69–73. doi: 10.1007/BF00197359
- Santoriello, C., and Zon, L. I. (2012). Science in medicine Hooked?! modeling human disease in zebrafish. *J. Clin. Invest.* 122, 2337–2343. doi: 10.1172/JCI60434
- Schulze, L., Henninger, J., Faustino, A. I., Chaigne, T., Kadobianskyi, M., Hakiy, N., et al. (2018). Transparent *Danionella translucida* as a genetically tractable vertebrate brain model. *Nat. Methods* 15, 977–983. doi: 10.1038/s41592-018-0144-6
- Sittaramane, V., and Chandrasekhar, A. (2008). Expression of unconventional myosin genes during neuronal development in zebrafish. *Gene Expr. Patterns* 8, 161–170.
- Sleigh, J. N., Vagnoni, A., Twelvetrees, A. E., and Schiavo, G. (2017). Methodological advances in imaging intravital axonal transport. *F1000Res.* 6:200. doi: 10.12688/f1000research.10433.1
- Sui, W.-H., Huang, S.-H., Wang, J., Chen, Q., Liu, T., and Chen, Z.-Y. (2015). Myosin Va mediates BDNF-induced postendocytic recycling of full-length TrkB and its translocation into dendritic spines. *J. Cell Sci.* 128, 1108–1122. doi: 10.1242/jcs.160259
- Suster, M. L., Abe, G., Schouw, A., and Kawakami, K. (2011). Transposon-mediated BAC transgenesis in zebrafish. *Nat. Protoc.* 6, 1998–2021. doi: 10.1038/nprot.2011.416
- Tamplin, O. J., White, R. M., Jing, L., Kaufman, C. K., Lacadie, S. A., Li, P., et al. (2012). Small molecule screening in zebrafish: swimming in potential drug therapies. *Wiley Interdiscip. Rev. Dev. Biol.* 1, 459–468. doi: 10.1002/wdev.37
- Tomer, R., Lovett-barron, M., Kauvar, I., Broxton, M., and Deisseroth, K. (2015). SPED light sheet microscopy: fast mapping of biological system structure and function. *Cell* 163, 1796–1806. doi: 10.1016/j.cell.2015.11.061
- Vagnoni, A., and Bullock, S. L. (2017). A simple method for imaging axonal transport in ageing neurons using the adult *Drosophila* wing. *Nat. Protoc.* 11, 1711–1723. doi: 10.1038/nprot.2016.112
- Vukoja, A., Rey, U., Petzoldt, A. G., Ott, C., Vollweiler, D., Quentin, C., et al. (2018). Presynaptic biogenesis requires axonal transport of lysosome-related vesicles. *Neuron* 99, 1216–1232. doi: 10.1016/j.neuron.2018.08.004
- Wu, H., Nash, J. E., Zamorano, P., and Garner, C. C. (2002). Interaction of SAP97 with minus-end-directed actin motor myosin VI: implications for AMPA receptor trafficking. *J. Biol. Chem.* 277, 30928–30934. doi: 10.1074/jbc.M203735200
- Wu, X., Bowers, B., Rao, K., Wei, Q., and Hammer, J. A. (1998). Visualization of melanosome dynamics within wild-type and dilute melanocytes suggests a paradigm for myosin v function in vivo. *J. Cell Biol.* 143, 1899–1918. doi: 10.1083/jcb.143.7.1899
- Wu, X., Jung, G., and Hammer, J. A. III (2000). Functions of unconventional myosins. *Curr. Opin. Cell Biol.* 12, 42–51. doi: 10.1016/S0955-0674(99)00055-1
- Yao, Y., Sun, S., Fei, F., Wang, J., Wang, Y., Zhang, R., et al. (2017). Screening in larval zebrafish reveals tissue-specific distribution of fifteen fluorescent compounds. *Dis. Model. Mech.* 10, 1155–1164. doi: 10.1242/dmm.028811
- Yoshimura, A., Fujii, R., Watanabe, Y., Okabe, S., Fukui, K., and Takumi, T. (2006). Myosin-Va facilitates the accumulation of mRNA/protein complex in dendritic spines. *Curr. Biol.* 16, 2345–2351. doi: 10.1016/j.cub.2006.10.024
- Yu, C. C., Reddy, B. J. N., Wortman, J. C., and Steven, P. (2017). Axonal transport: a constrained system. *J. Neurol. Neurosurg.* 2, 20–24. doi: 10.29245/2572.942X/2017/3.1118
- Zhang, Y., Huang, H., Zhang, B., and Lin, S. (2016). TALEN- and CRISPR-enhanced DNA homologous recombination for gene editing in zebrafish. *Methods Cell Biol.* 135, 107–120. doi: 10.1016/bs.mcb.2016.03.005
- Zhang, Y., Qin, W., Lu, X., Xu, J., Huang, H., Bai, H., et al. (2017). Programmable base editing of zebrafish genome using a modified CRISPR-Cas9 system. *Nat. Commun.* 8:118. doi: 10.1038/s41467-017-00175-6
- Zhu, X. J., Wang, C. Z., Dai, P. G., Xie, Y., Song, N. N., Liu, Y., et al. (2007). Myosin X regulates netrin receptors and functions in axonal path-finding. *Nat. Cell Biol.* 9, 184–192. doi: 10.1038/ncb1535
- Zon, L. I., and Peterson, R. T. (2005). In vivo drug discovery in the zebrafish. *Nat. Rev. Drug Discov.* 4, 35–44. doi: 10.1038/nrd1606

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2019 Bercier, Rosello, Del Bene and Revenu. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

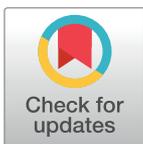
RESEARCH ARTICLE

# Regulation of the apical extension morphogenesis tunes the mechanosensory response of microvilliated neurons

Laura Desban<sup>1</sup>, Andrew Prendergast<sup>1</sup>, Julian Roussel<sup>1</sup>, Marion Rosello<sup>2</sup>, David Geny<sup>3</sup>, Claire Wyart<sup>1\*</sup>, Pierre-Luc Bardet<sup>1\*</sup>

**1** Institut du Cerveau et de la Moelle épinière (ICM), Inserm U 1127, CNRS UMR 7225, Sorbonne Université, Paris, France, **2** Institut Curie, PSL Research University, INSERM U934, CNRS UMR3215, Paris, France, **3** Plateforme d'imagerie Neurlmag de l'Institut de Psychiatrie et Neurosciences de Paris (IPNP), Inserm U 1266, Paris, France

\* [pierre-luc.bardet@icm-institute.org](mailto:pierre-luc.bardet@icm-institute.org) (PB); [claire.wyart@icm-institute.org](mailto:claire.wyart@icm-institute.org) (CW)



 OPEN ACCESS

**Citation:** Desban L, Prendergast A, Roussel J, Rosello M, Geny D, Wyart C, et al. (2019) Regulation of the apical extension morphogenesis tunes the mechanosensory response of microvilliated neurons. *PLoS Biol* 17(4): e3000235. <https://doi.org/10.1371/journal.pbio.3000235>

**Academic Editor:** Andy Groves, Baylor College of Medicine, UNITED STATES

**Received:** August 9, 2018

**Accepted:** April 3, 2019

**Published:** April 19, 2019

**Copyright:** © 2019 Desban et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** All relevant data are within the paper and its Supporting Information files.

**Funding:** CW: European Research Council (ERC) Starting Grant 'Optoloco' (grant no. 311673) <https://erc.europa.eu/>. CW: Human Frontier Science Program (HFSP) Research Grant (grant no. RGP063-2018) <http://www.hfsp.org/>. CW: New York Stem Cell Foundation (NYSCF) Neuroscience Award (grant no. NYSCF-R-NI39) <https://nyscf.org/>. PLB, CW: Agence Nationale pour le Recherche

## Abstract

Multiple types of microvilliated sensory cells exhibit an apical extension thought to be instrumental in the detection of sensory cues. The investigation of the mechanisms underlying morphogenesis of sensory apparatus is critical to understand the biology of sensation. Most of what we currently know comes from the study of the hair bundle of the inner ear sensory cells, but morphogenesis and function of other sensory microvilliated apical extensions remain poorly understood. We focused on spinal sensory neurons that contact the cerebrospinal fluid (CSF) through the projection of a microvilliated apical process in the central canal, referred to as cerebrospinal fluid-contacting neurons (CSF-cNs). CSF-cNs respond to pH and osmolarity changes as well as mechanical stimuli associated with changes of flow and tail bending. In vivo time-lapse imaging in zebrafish embryos revealed that CSF-cNs are atypical neurons that do not lose their apical attachment and form a ring of actin at the apical junctional complexes (AJCs) that they retain during differentiation. We show that the actin-based protrusions constituting the microvilliated apical extension arise and elongate from this ring of actin, and we identify candidate molecular factors underlying every step of CSF-cN morphogenesis. We demonstrate that *Crumbs 1* (*Crb1*), *Myosin 3b* (*Myo3b*), and *Espin* orchestrate the morphogenesis of CSF-cN apical extension. Using calcium imaging in *crb1* and *espin* mutants, we further show that the size of the apical extension modulates the amplitude of CSF-cN sensory response to bending of the spinal cord. Based on our results, we propose that the apical actin ring could be a common site of initiation of actin-based protrusions in microvilliated sensory cells. Furthermore, our work provides a set of actors underlying actin-based protrusion elongation shared by different sensory cell types and highlights the critical role of the apical extension shape in sensory detection.

“Investissements d’avenir” ANR-10-IAIHU-06 (Big Brain Theory ICM Program) and ANR-11-INBS-0011 (NeurATRIS: Translational Research Infrastructure for Biotherapies in Neurosciences). <http://www.agence-nationale-recherche.fr/en/>. Medical Research Foundation (FRM): LD: (grant no. FDT20170437143) and MD: (FRM grant number ECO20170637481. <https://www.frm.org/>). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Competing interests:** The authors have declared that no competing interests exist.

**Abbreviations:** AJC, apical junctional complex; ASIC, Acid-Sensing Ion Channels; Baiap, Brain-specific angiogenesis inhibitor 1-associated protein; Cdh2, Cadherin 2; Crb1, Crumbs 1; CRISPR/Cas9, Clustered Regularly Interspaced Short Palindromic Repeats/CRISPR-Associated Protein 9; CSF-cN, cerebrospinal fluid-contacting neuron; DN, dominant-negative; FISH, fluorescent in situ hybridization; GFP, green fluorescent protein; GLM, general linear model; gRNA, guide RNA; hpf, hours post fertilization; I-BAR, Inverse Bin-Amphiphysin-Rvs167; IHC, immunohistochemistry; IRSp53, Insulin Receptor Substrate protein of 53 kDa; Myo3b, Myosin 3b; Pkd2l1, Polycystic kidney disease-2-like 1; RFP, red fluorescent protein; Shh, Sonic hedgehog; SH3, SRC Homology 3; STED, stimulated emission depletion; zEspin1, zebrafish Espin isoform 1; ZO-1, Zonula Occludens-1.

## Introduction

Cells assume a variety of shapes known to be instrumental in their specific functions. Morphologies of neurons are among the most complex and diverse: they extend protrusions, build dendritic arborizations, and project axons in an intricate and stereotypical manner [1]. Sensory neurons, in particular, display a high degree of specialization, allowing them to detect sensory cues, transduce the information into electrical signals, and propagate action potentials to the central nervous system. Much effort has been invested in the identification of molecular determinants underlying sensory detection and transduction, mostly in hair cells—non-neuronal sensory cells—to decipher the pathways and morphological features of audition [2,3]. In contrast, mechanisms responsible for the morphological differentiation of sensory neurons and the role of their morphology in sensory functions remain elusive. Here, we tackle these questions by studying spinal cerebrospinal fluid-contacting neurons (CSF-cNs), a class of evolutionarily conserved sensory neurons [4]. We investigated the molecular mechanisms underlying morphogenesis of the CSF-cN apical extension and its role in sensory-mediated activity.

Spinal CSF-cNs are located along the central canal, into which they extend an apical extension composed of 1 motile cilium surrounded by actin-based protrusions that are in direct contact with the CSF [5–8]. Over the past decade, multiple studies investigated the hypothesis of Kolmer and Agduhr [9,10] that CSF-cNs are sensory and detect cues via their apical extension reminiscent of the hair bundle in hair cells. Diverse sensory cues can be detected by CSF-cNs, including changes in pH and osmolarity [11,12] and mechanical stimuli associated with CSF flow [11,13] or spine curvature [14]. CSF-cNs express the transient potential receptor channel 3 (Trpp3), further referred to as Pkd2l1 (for Polycystic kidney disease-2-like 1) [8,12,15]. While CSF-cN response to pH changes seems to be carried by Acid-Sensing Ion Channels (ASICs) rather than by Pkd2l1 [11,12], multiple lines of evidence indicate that Pkd2l1 is necessary for responses to mechanosensory inputs, such as changes of osmolarity, CSF flow, and tail bending [12–14]. The Pkd2l1 channel is abundantly enriched at the level of CSF-cN microvilli [13], reinforcing previous suggestions that the CSF-cN apical extension might play a similar role in sensory functions to the hair bundle in hair cells [16]. During embryogenesis, spinal CSF-cNs are born from 2 ventral progenitor domains in mouse and zebrafish [17–19]. This dual origin is associated with 2 functionally and morphologically distinct subpopulations organized in 2 ventral and dorsolateral rows [13,20,21]. We previously showed that ventral and dorsolateral CSF-cNs exhibit differently shaped apical extensions [21]. When considered with the localization of Pkd2l1 at the microvilli [13], this observation suggests that such morphological distinction might underlie distinct sensory functions described in the 2 CSF-cN subtypes. How CSF-cNs form their apical extension and how their shape contributes to sensory function remain unknown.

The CSF-cN neural progenitors are epithelial cells displaying an apicobasal polarity with their apical side facing the CSF. At this side, CSF-cN precursors form apical junctional complexes (AJCs), including adherens junctions and tight junctions, closely associated with actin filaments organized in a circumferential belt [22]. Together, AJCs and actin cytoskeleton control the cohesiveness and permeability of the developing tissue [23,24]. The apical localization and maintenance of AJCs rely on the activity of a family of transmembrane proteins, the Crumbs family [25]. Neural progenitors give birth to differentiating neurons through self-renewing asymmetric divisions [26]. It was recently shown that the daughter cell fated to become a neuron loses its apical domain and adhesiveness [27], a process thought to be key for its subsequent neuronal differentiation [23,24]. During their differentiation, many neurons then exhibit actin-based protrusions in a polarized fashion, which indicates that they re-establish an apicobasal polarity. It is unknown whether CSF-cNs also differentiate through a process of apical domain loss and polarity re-establishment or not.

In the case of sensory neurons, specialized protrusions called microvilli can arise at their apical side to constitute the sensory organelle. Our current understanding of the formation of actin-based protrusions comes from studies of filopodia—short-lived thin processes—and involves 2 critical steps: initiation and elongation [28,29]. The initiation of filopodia is characterized by the recruitment of actin nucleators and organizers to the apical membrane specifically. Insulin Receptor Substrate protein of 53 kDa (IRSp53) family members, including Brain-specific angiogenesis inhibitor 1-associated protein (Baiap) and Baiap-like proteins, are key elements of this step because they are capable of inducing membrane curvature and recruiting actin nucleators essential for the initiation of protrusions [30]. So far, it is not known whether this model translates to long-lived microvilli found in sensory cells [31].

After their initiation, membrane protrusions lengthen and, in the case of microvilli, stabilize. The elongation of sensory microvilli has been thoroughly studied in hair cells of the inner ear that form rigid microvilli derivatives referred to as stereocilia [31,32]. Through the isolation of mutations responsible for deafness in humans and animal models, multiple genes have been linked to morphogenesis of the sophisticated hair cell's sensory organelle [33]. Among these genes, *espin* and *myo3b* encode proteins that have been shown to interact and ensure the proper morphogenesis of stereocilia [34–36]. Loss of Espin function was further demonstrated to lead to stereocilia defects and, eventually, degeneration followed by hair cell death [33,37]. It is not clear whether these interaction and morphogenetic mechanisms translate to other microvilliated sensory cells.

Here, we monitor CSF-cN differentiation using long-term time-lapse imaging of F-actin and report 3 critical steps in their morphogenesis: (i) the presence of a ring of actin at the level of the AJCs, (ii) the initiation, and (iii) the elongation of actin-based protrusions from this actin ring. We show that, in contrast to other newborn neurons, CSF-cNs do not withdraw from the ependymal surface to migrate and differentiate. Instead, they appear to retain their apical cell polarity throughout differentiation, an original element of their morphogenesis. We describe new molecular factors known to be involved in cell polarity and/or morphogenesis, which are specifically enriched in CSF-cNs in the spinal cord. We identified a polarity factor—*crb1*—and several actin cytoskeleton interactors—*baiap2a*, *baiap2l1b*, *myo3b*, and *espin*. Our observations suggest a model in which these factors cooperate to regulate each step of the formation of the CSF-cN apical extension. Investigating the phenotype of a newly generated *crb1* mutant, we describe a critical role of Crumbs 1 in the morphogenesis of CSF-cN apical extensions. Using a dominant-negative (DN) form of Myosin 3b, we demonstrate that this protein is required to address the actin-bundling factor Espin to the apical extension and enable it to reach the proper size. The analysis of a newly generated *espin* mutant indicates that Espin actin-bundling activity is necessary for the correct elongation of the microvilli constituting the CSF-cN apical extensions. Finally, functional analysis demonstrated that shorter apical extensions were associated to reduced mechanosensory responses in CSF-cNs lacking Crb1 or Espin. All together, these results provide insightful elements to build a mechanistic model of the formation of the apical extension by CSF-cNs and how its structure contributes to CSF-cN sensory function.

## Results

### Critical steps to form a specialized apical extension

To investigate the formation of the actin-based apical extension in CSF-cNs, we used the mosaic expression of LifeAct, a marker of F-actin [38], under the control of the *pkd2l1* promoter [39]. We performed long-term time-lapse imaging on single CSF-cNs in live zebrafish from 20 to 22 hours post-fertilization (hpf) (S1 Movie). CSF-cNs first appear as cuboidal, with

a growing axon and a ring of actin at the apical side (Fig 1, Stage 1). Then, while the CSF-cN soma becomes round, short protrusions arise from the actin ring (Fig 1, Stage 2). Finally, CSF-cNs acquire their typical pear-like shape bearing elongating actin-based protrusions (Fig 1, Stage 3). The same sequence of events was observed in several cells from the 2 subpopulations (ventral and dorso-lateral; Fig 1) and confirmed on fixed tissues after immunostaining at 24, 48 and 72 hpf of LifeAct-GFP-positive CSF-cNs (Fig 2).

Our results indicate that the actin cytoskeleton goes through dynamic and extensive remodeling, suggesting an active participation in CSF-cN morphogenesis. F-actin is first organized at the apical side of the cells to form a ring from which, later on, protrusions seem to arise to form the apical extension.

### The ring of actin colocalizes with CSF-cN apical junctional complexes

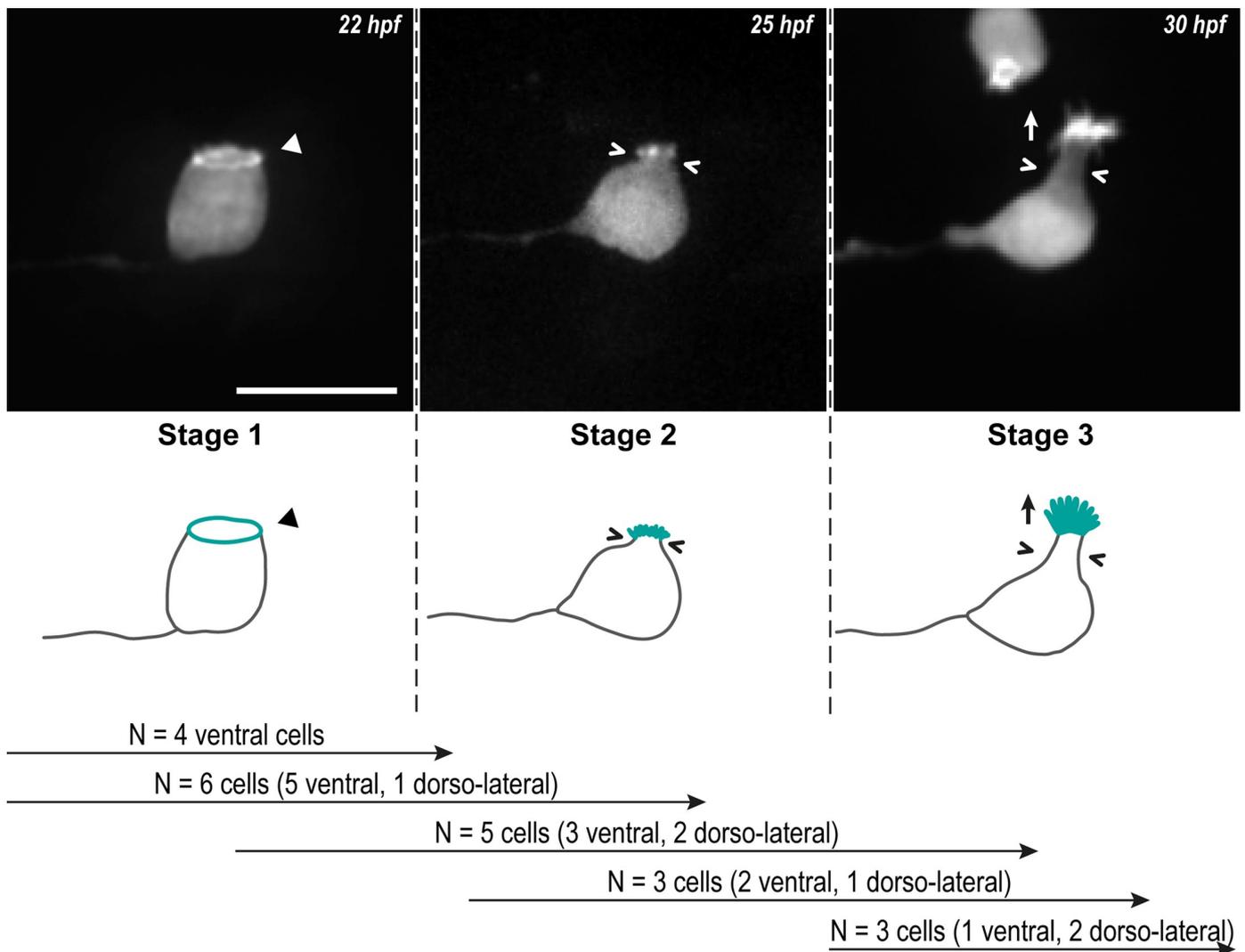
The localization and shape of the ring of actin are reminiscent of the circumferential structure formed by actin and adherens and tight junctions at the AJCs in epithelial cells [22]. These complexes are usually lost in newborn neurons to allow their withdrawal from the ventricular zone in order to migrate and differentiate [23,27]. To investigate the localization of AJCs in CSF-cNs relative to the ring of actin, we chose 3 different markers representative of the apical polarization complex—Crumbs 1 (Crb1) [25], adherens junctions—Cadherin 2 (Cdh2) [40], and tight junctions—Zonula-occludens-1 (ZO-1) [41]. At 24 hpf, we observed LifeAct-labeled CSF-cNs displaying the typical ring of actin at their apical side (Fig 2A–2C, arrowheads). At the vicinity of the ring, we also detected Cdh2 (Fig 2A), ZO-1 (Fig 2B), and Crb1 (Fig 2C), confirming that F-actin is enriched at the level of CSF-cN AJCs at 24 hpf.

In 72-hpf larvae, in contrast to ZO-1 being expressed in all cells lining the central canal (Fig 2D and 2E, arrowheads), the expression of Crb1 (Fig 2E and 2F, arrowheads) was specific to CSF-cNs and showed distinct patterns between ventral and dorsolateral subtypes (Fig 2F). In ventral cells, Crb1 formed a narrow subapical ring, as seen in 24-hpf embryos, whereas it expanded in the entire apical extension of dorsolateral cells (Fig 2E and 2F).

These results show that a ring of actin is formed early on in CSF-cNs at the level of the AJCs and that both structures persist after the apical extension is formed. This observation highlights a peculiarity of CSF-cNs compared with other neurons [27] because they retain their apical polarity during differentiation. The initiation and elongation of the actin-based protrusions from the ring of actin formed at the AJCs suggest that the apical complexes participate in the formation of the CSF-cN apical extension.

### Crb1 is required for the morphogenesis of CSF-cN apical extension

Because its expression was specific to CSF-cNs at 72 hpf and showed different patterns between ventral and dorsolateral cells that exhibit distinct morphologies [21], we reasoned that Crb1 might play a special role in the maturation of the apical extension. We therefore knocked out the *crb1* gene using clustered regularly interspaced short palindromic repeats/CRISPR/Cas9-mediated genome editing (S1 Fig). The newly generated mutated allele *crb1<sup>icm31</sup>*, thereafter referred to as *crb1<sup>-</sup>*, encodes for a truncated protein lacking most functional domains and resulted in the loss of Crb1 immunoreactivity in homozygous mutant larvae (S1 Fig). Analysis of the morphogenesis of CSF-cNs in *crb1<sup>-/-</sup>* animals revealed that mutant cells established a normal apical polarity and presented apical actin-based protrusions (S1 Fig). This result can be explained by the functional redundancy between members of the *crumbs* family allowing the cells lacking Crb1 function to compensate for cell polarization with other Crb proteins [25]. Nonetheless, we noticed that Crb1 loss of function was associated with specific morphological defects of CSF-cN apical extensions at 72 hpf (Fig 3A). Using a

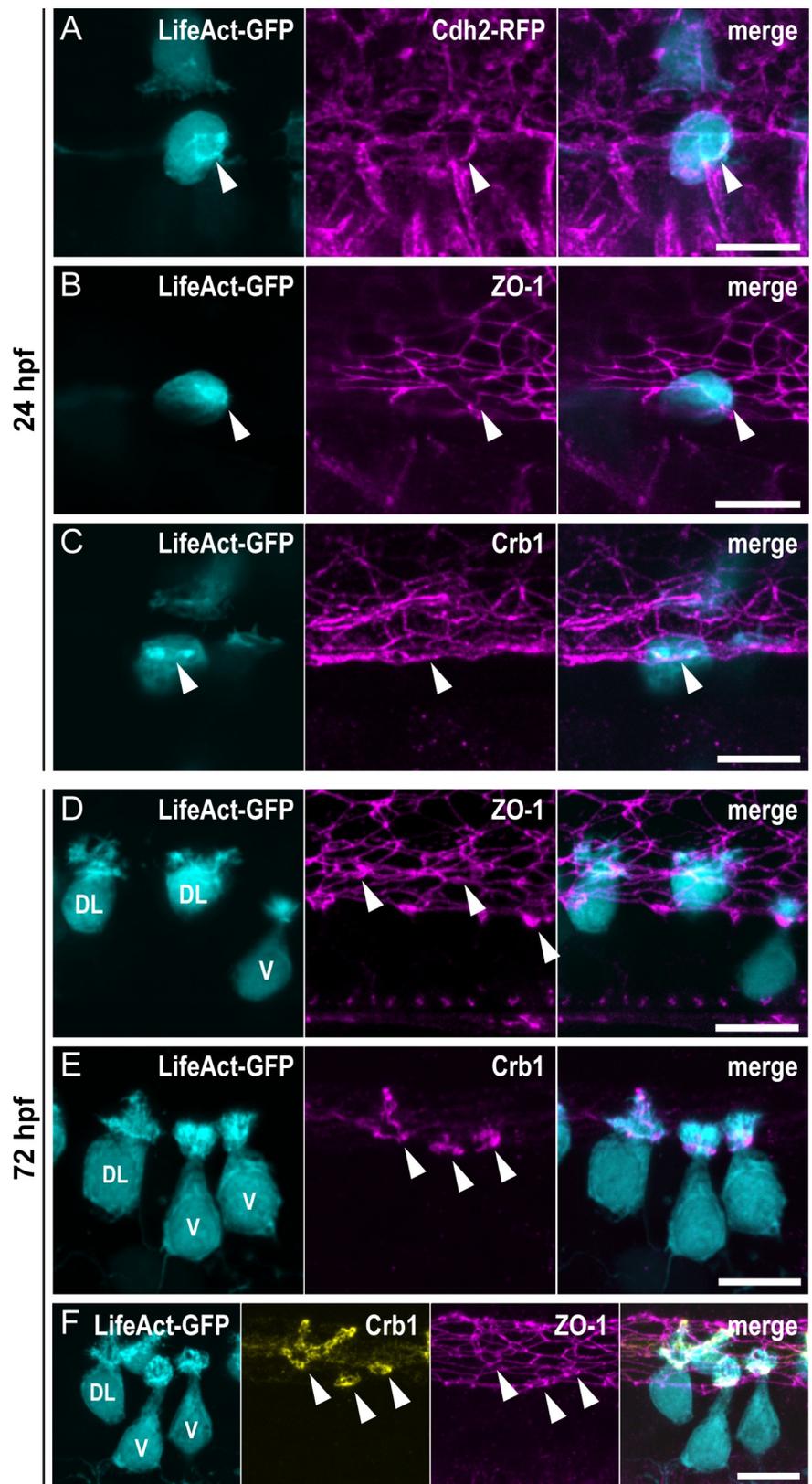


**Fig 1. CSF-cNs go through 3 critical steps to form their apical extension.** Z-projections from time-lapse acquisitions (top panels; lateral views of ventral CSF-cNs with rostral to the left) and schematics (bottom panels) showing the 3 stages CSF-cNs go through during the formation of their AE. The CSF-cN soma becomes round and short actin-based protrusions (Stage 2) arise from the ring of actin (Stage 1, arrowhead) concomitantly with a subapical constriction (Stages 2 and 3; chevrons). Gradually, protrusions lengthen to form the AE (Stage 3; arrow) characteristic of differentiated CSF-cNs. Data collected over 13 time-lapse sessions for 15 ventral and 6 dorsolateral cells. The number of cells imaged transitioning from one stage to the next is indicated in the bottom legend. Scale bar, 10  $\mu$ m. AE, apical extension; CSF-cN, cerebrospinal fluid-contacting neuron. hpf, hours post fertilization.

<https://doi.org/10.1371/journal.pbio.3000235.g001>

previously published method [21], we measured the area covered by the CSF-cN apical extensions in 72-hpf mutant larvae compared with wild-type siblings. We found that mutant dorso-lateral and, to a lesser extent, ventral CSF-cNs formed significantly smaller apical extensions (15% reduction in ventral and 28% reduction in dorso-lateral cells) (Fig 3B). This effect resulted in the disappearance of the difference between ventral versus dorsolateral apical extensions in terms of size (Fig 3B) but not in terms of global shape [21].

This result confirms our hypothesis that, beyond its known function as a polarity factor, Crb1 plays a specific role in CSF-cN morphogenesis, which cannot be compensated by other members of the Crb family.

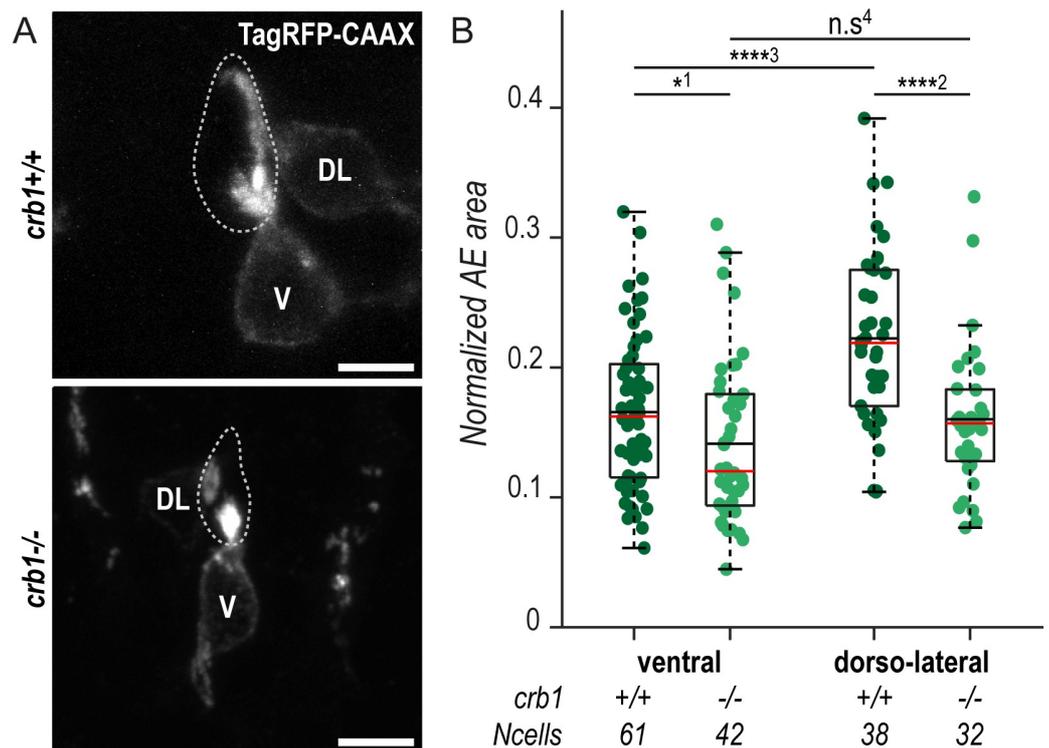


**Fig 2. The ring of actin colocalizes with the CSF-cN apical junctional complexes.** (A–C) Z-projections from lateral views of the spinal cord in 24-hpf embryos showing the colocalization of the ring of actin (LifeAct; arrowheads) with different markers of the AJCs: (A) Cdh2 for adherens junctions, (B) ZO-1 for tight junctions, and (C) Crb1 for the apical domain. (A) Double immunostaining for GFP and RFP in triple transgenic *Tg(pkcd211:Gal4,UAS:LifeAct-GFP; cryaa:V,cdh2:cdh2-RFP)* embryos. (B, C) Double immunostaining for GFP and (B) ZO-1, or (C) Crb1 in *Tg(pkcd211:Gal4,UAS:LifeAct-GFP;cryaa:V)* embryos. (D–F) Z-projections from lateral views of 72-hpf *Tg(pkcd211:Gal4,UAS:LifeAct-GFP;cryaa:V)* larvae immunostained for GFP and (D) ZO-1, (E) Crb1, or (F) ZO-1 and Crb1. The markers ZO-1 and Crb1 are retained at the AJCs (arrowheads) in both V and DL CSF-cNs after differentiation. In DL cells, Crb1 expands to the apical extension. Scale bars, 10  $\mu$ m. AJC, apical junctional complexes; Cdh2, Cadherin 2; Crb1, Crumbs 1; CSF-cN, cerebrospinal fluid-contacting neurons; DL, dorso-lateral; GFP, green fluorescent protein; hpf, hours post fertilization; RFP, red fluorescent protein; V, ventral; ZO-1, zonula-occludens-1.

<https://doi.org/10.1371/journal.pbio.3000235.g002>

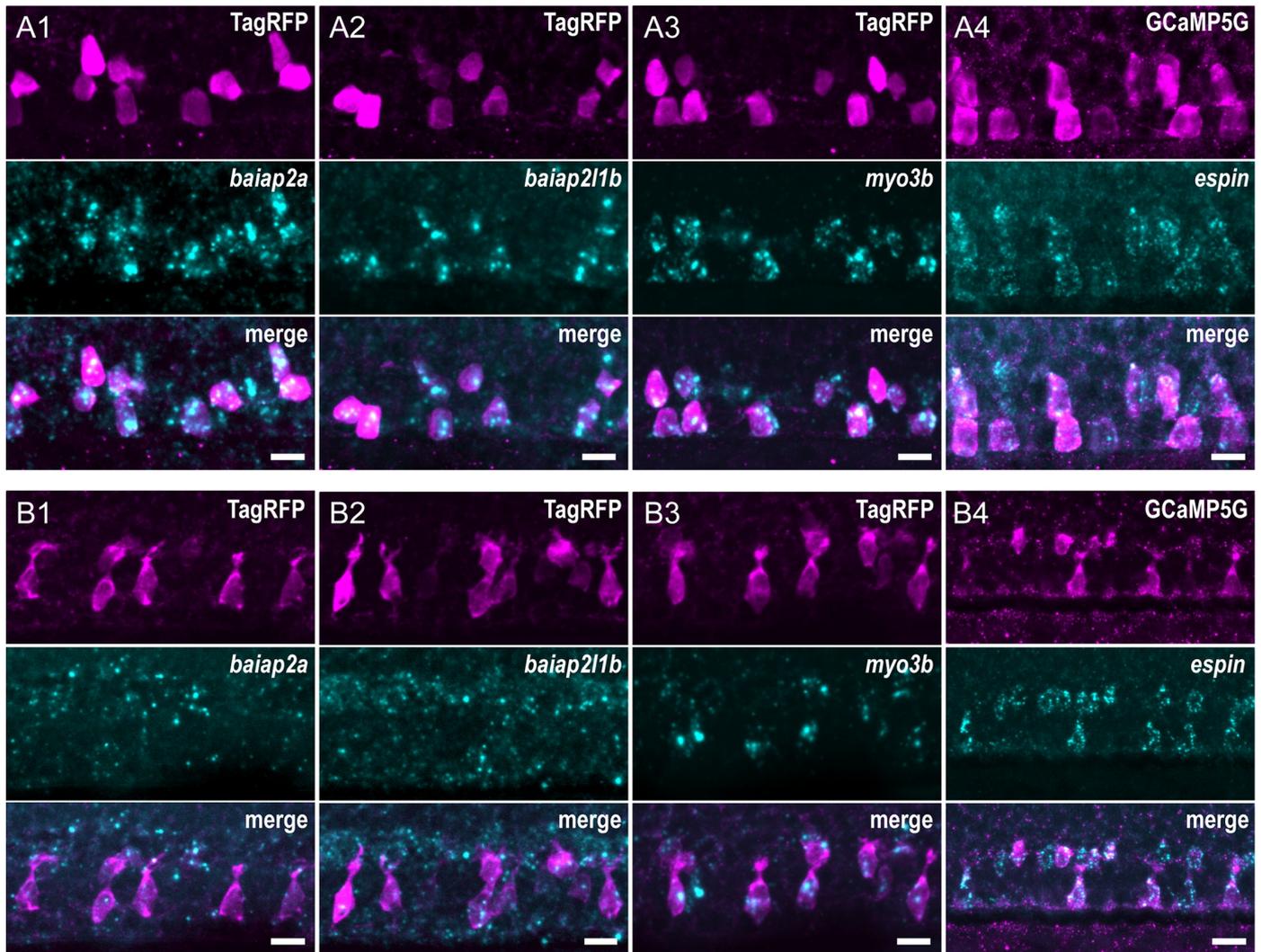
### CSF-cNs specifically express factors involved in the initiation and lengthening of actin-based membrane protrusions

We hypothesized that, once formed, the AJC-located actin ring serves as a platform to recruit specific actin-modifying proteins and monomers of actin to enable the initiation and the elaboration of the apical extension. We searched for actin organizers involved in the formation of actin-based protrusions and investigated their expression in CSF-cNs by combining fluorescent in situ hybridization (FISH) with immunohistochemistry (IHC) for fluorescent proteins



**Fig 3. Crb1 participates in the proper development of the CSF-cN apical extension.** (A) Z-projections from transversal sections showing V and DL TagRFP-CAAX-expressing CSF-cNs in 72-hpf larvae illustrating the smaller AE in *crb1<sup>-/-</sup>* (bottom panel) compared with wild-type siblings (top panel). The central canal is outlined (dotted lines) according to ZO-1 staining. Scale bars, 5  $\mu$ m. (B) Quantification of the normalized area covered by CSF-cN AEs at 72 hpf in V and DL cells in *crb1<sup>-/-</sup>* (light green;  $N = 3$  fish) compared with wild-type siblings (dark green;  $N = 4$  fish). In both CSF-cN subtypes, the AE was significantly smaller in mutant larvae compared with wild-type ( $p_1 = 0.0477$ ,  $p_2 = 9.9019 \times 10^{-5}$ ,  $p_3 = 2.6381 \times 10^{-5}$ ,  $p_4 = 0.1876$ ). Underlying data can be found in [S1 Data](#). AE, apical extension; Crb1, Crumbs 1; CSF-cN, cerebrospinal fluid-contacting neuron; DL, dorsolateral; hpf, hours post fertilization; n.s., not significant; V, ventral; ZO-1, zonula-occludens-1.

<https://doi.org/10.1371/journal.pbio.3000235.g003>



**Fig 4. CSF-cNs express a set of known morphogenetic factors.** Expression of candidate factors known to be involved in the formation of actin-based protrusions—*baiap2a*, *baiap2l1b*, *myo3b*, and *espin*—assessed by FISH at 24 hpf (A1–4) and 72 hpf (B1–4). Expression in CSF-cNs was validated by combining FISH to IHC for RFP or GFP in *Tg(pkcd2l1:TagRFP)* (A1–3, B1–3) or *Tg(pkcd2l1:GCaMP5G)* (A4, B4) transgenic animals, respectively. In 24-hpf embryos, *baiap2a* (A1), *baiap2l1b* (A2), *myo3b* (A3), and *espin* (A4) were enriched in CSF-cNs. In 72-hpf larvae, expression of *baiap2a* (B1) and *baiap2l1b* (B2) was not clearly detected, whereas *myo3b* (B3) and *espin* (B4) remained strongly expressed in CSF-cNs. Scale bars, 10  $\mu$ m. CSF-cN, cerebrospinal fluid-contacting neuron; FISH, fluorescent in situ hybridization; GFP, green fluorescent protein; hpf, hours post fertilization; IHC, immunohistochemistry; RFP, red fluorescent protein.

<https://doi.org/10.1371/journal.pbio.3000235.g004>

under the control of the *pkcd2l1* promoter to label CSF-cNs. In 24-hpf embryos, we found at least 4 factors specifically enriched in both ventral and dorsolateral CSF-cNs (Fig 4) in the spinal cord: *baiap2a* (Fig 4A1), *baiap2l1b* (Fig 4A2), *myo3b* (Fig 4A3), and *espin* (Fig 4A4). In 72-hpf larvae, *baiap2a* (Fig 4B1) and *baiap2l1b* (Fig 4B2) RNA was not clearly detected in CSF-cNs, whereas both *myo3b* (Fig 4B3) and *espin* (Fig 4B4) still showed strong and specific expression. The patterns of expression of Myo3b and Espin were also confirmed by IHC using specific polyclonal antibodies (S2 Fig) [35,42]. Both proteins are enriched at the apical extension of CSF-cNs in the spinal cord. Furthermore, Espin was clearly detected in various sensory microvilliated cells, namely, olfactory receptor neurons in the olfactory pit and hair cells in the inner ear and lateral line (S2 Fig, panel B).

These observations provided candidate molecular factors potentially involved in the elaboration of the mature CSF-cN apical extension.

### Myo3b is necessary for the localization of Espin and the morphogenesis of the CSF-cN apical extension

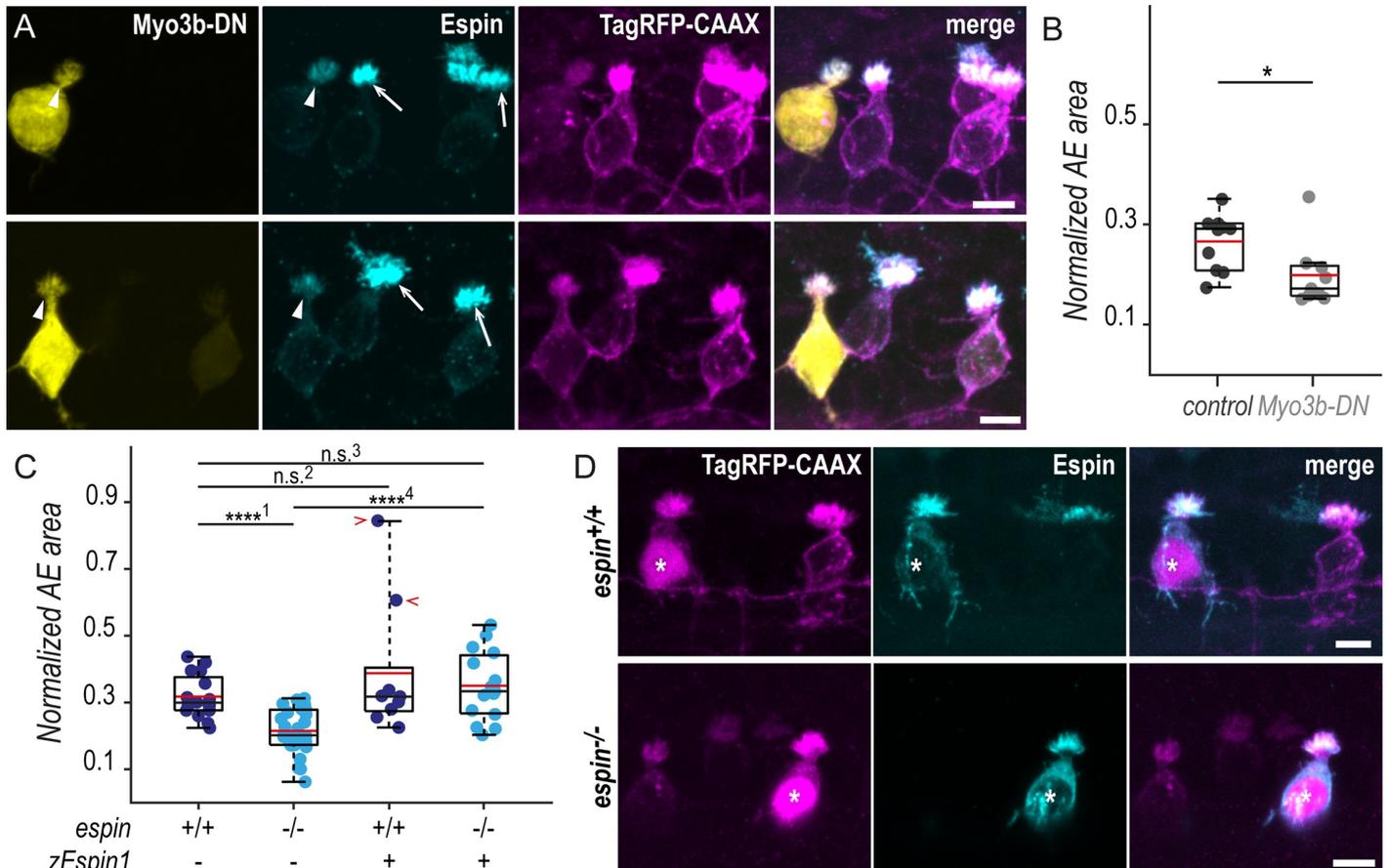
The highly specific expression of *myo3b* and *espin* in CSF-cNs, along with their known interaction to promote filopodia lengthening [34,36] and their role in stereocilia morphogenesis in hair cells [34,35,37], motivated us to verify whether these 2 factors contribute to the morphogenesis of CSF-cN apical extensions. We disrupted Myosin 3b (Myo3b) normal function by expressing specifically in CSF-cNs Myo3b-DN, a dominant-negative (DN) form of the protein lacking the motor domain. Mosaic expression of Myo3b-DN enabled us to compare the morphology of cells with reduced Myo3b function with neighboring control ones within the same animal (Fig 5A). At 72 hpf, the expression of Myo3b-DN in CSF-cNs resulted in the reduction of Espin apical staining and of the size of apical extensions (Fig 5A). We quantified the reduction in size by measuring the area covered by apical extensions of ventral CSF-cNs expressing Myo3b-DN or not. We found that cells lacking Myo3b function exhibited significantly smaller apical extensions compared with control cells (25% reduction; Fig 5B).

Our results in CSF-cNs are consistent with previous observations in hair cells and further demonstrate that Myo3b is required for addressing Espin to the actin-based protrusions and is critical for the correct morphogenesis of the apical extension.

### Espin actin-bundling activity is required for the morphogenesis of CSF-cN apical extensions

Espin was shown to rely on its interaction with Myo3b to reach the tip of actin-based protrusions and ensure their lengthening through its actin-bundling activity [34,36]. To test whether the shortening of the apical extension described with Myo3b-DN is due to Espin loss at the CSF-cN apical extension, we generated a mutant allele of *espin* using CRISPR/Cas9-mediated genome editing (S3 Fig). In zebrafish, there is a single *espin* paralog, whom predicted products share a high sequence identity with the main domains described in the rat Espin protein. Because several Espin isoforms have been described in rodents [42–44], we designed our guide RNA against the sequence encoding for the very first amino acids of the conserved actin-bundling module (S3 Fig, panel B). This actin-bundling module is essential for the protein actin-bundling function and present in all described isoforms [42,44,45]. Our newly generated mutant *espin*<sup>icm26</sup> allele, thereafter referred to as *espin*<sup>-</sup>, disturbed the sequence of the module from the first amino acid (S3 Fig, panels B-C) and resulted in loss of Espin immunoreactivity at embryonic and larval stages, confirming the efficiency of our knock-out strategy (S3 Fig, panel D). We observed a reduction of Espin staining in heterozygous fish, indicating that the loss of one *espin* wild-type allele is sufficient to reduce protein levels (S3 Fig, panel E).

In 72-hpf *espin*<sup>-/-</sup> larvae, we noticed that CSF-cNs displayed smaller apical extensions than in wild-type siblings (Fig 5D). Using the same analysis than previously for Myo3b-DN, we measured a significant reduction of the area covered by apical extensions in ventral CSF-cNs of homozygous compared with wild-type larvae (32% reduction; Fig 5C). To further ascertain that this reduction in the size of CSF-cN apical extensions was specific to the loss of Espin function, we devised a vector to express the zebrafish Espin isoform 1 (zEspin1) in CSF-cNs under the control of the *pkd2l1* promoter. Mosaic expression of zEspin1 allowed us to compare overexpressing versus wild-type CSF-cNs as well as rescued versus mutant cells within *espin*<sup>+/+</sup> and *espin*<sup>-/-</sup> siblings, respectively (Fig 5D). In 72-hpf wild-type larvae, the overexpression of Espin occasionally led to a drastic increase in apical extension area (Fig 5C, red chevrons),



**Fig 5. In the absence of Espin actin-bundling activity, CSF-cNs form shorter apical extensions.** (A) Z-projections of whole-mounted spinal cords at 72 hpf showing mosaic expression of Myo3b-DN under the control of the *pkd211* promoter. Immunostaining reveals Espin (cyan) in TagRFP-CAAX-positive CSF-cNs (magenta) expressing Myo3b-DN (yellow, arrowheads) or not (arrows). In ventral CSF-cNs expressing Myo3b-DN, Espin staining is reduced (observed in 9 cells out of 9), and the AE appears smaller compared with wild-type cells. (B) Quantification of the normalized area covered by the AE of ventral cells expressing Myo3b-DN ( $N = 9$  cells) compared with nonexpressing neighboring cells (control;  $N = 10$  cells). Cells expressing Myo3b-DN form a significantly smaller AE ( $N = 5$  fish;  $p = 0.025$ ). (C) The same quantification was performed in *espin*<sup>-/-</sup> ventral cells at 72 hpf ( $N = 30$  cells in 5 fish) compared with wild-type cells ( $N = 16$  cells in 5 fish). Mutant cells display significantly smaller AEs ( $p_1 = 8.4010 \times 10^{-6}$ ). In wild-type ventral cells, the overexpression of the zEspin1 was sometimes associated with abnormally long microvilli (observed in 2 cells out of 9; red chevrons;  $p_2 = 0.2123$ ). In *espin*<sup>-/-</sup> cells, the mutant phenotype was rescued by zEspin1 ( $N = 15$  cells;  $p_3 = 0.3029$  and  $p_4 = 3.8556 \times 10^{-6}$ ). (D) Z-projections of whole-mounted spinal cords at 72 hpf showing mosaic expression of zEspin1 under the control of the *pkd211* promoter. Immunostaining for Espin (cyan) in TagRFP-CAAX-positive ventral CSF-cNs (magenta) reveals the loss of Espin immunoreactivity in *espin*<sup>-/-</sup> larvae, which is retrieved in mutant cells expressing zEspin1 (nuclear RFP; magenta; stars). Scale bars, 5  $\mu$ m. Underlying data can be found in [S1 Data](#). AE, apical extension; CSF-cN, cerebrospinal fluid-contacting neuron; DN, dominant-negative; hpf, hours post fertilization; Myo3b, myosin 3b; n.s., not significant; RFP, red fluorescent protein; zEspin1, zebrafish Espin isoform 1.

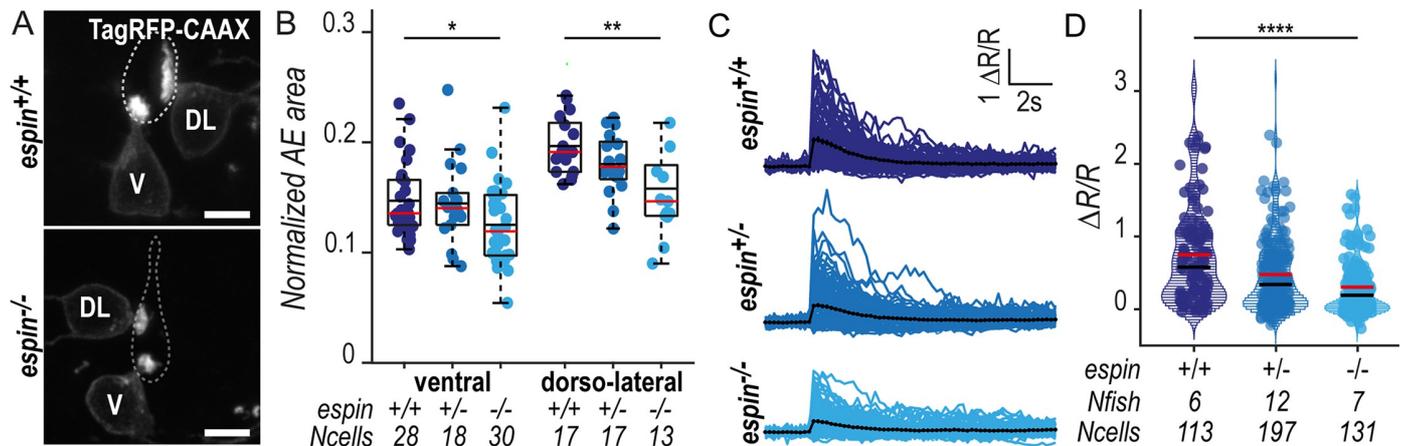
<https://doi.org/10.1371/journal.pbio.3000235.g005>

suggesting that Espin level modulates the final size of the apical extension size. In *espin*<sup>-/-</sup> animals, the expression of zEspin1 in a subset of CSF-cNs restored their apical extension size to wild-type values (Fig 5C and 5D).

These results demonstrate that the reduction in the apical extension size observed in our newly generated *espin*<sup>-/-</sup> mutant is specifically due to the loss of Espin in CSF-cNs.

### The actin-bundling factor Espin promotes the elongation of CSF-cN microvilli

To further analyze the impact of Espin loss on the CSF-cN apical extension, we evaluated the precise morphological features of *espin*<sup>-/-</sup> cells on spinal cross sections (Fig 6A). In 72-hpf larvae, we found that both ventral and dorsolateral CSF-cN subtypes form significantly smaller



**Fig 6. CSF-cNs with shorter apical extensions exhibit reduced mechanosensory response.** (A) Z-projections from transversal sections of spinal cords with V and DL TagRFP-CAAX-positive CSF-cNs at 72 hpf illustrating the smaller AEs in *espin*<sup>-/-</sup> (bottom panel) compared with wild-type siblings (top panel). The central canal was outlined (dotted lines) according to ZO-1 staining. Scale bars, 5 μm. (B) Quantification of the normalized area covered by the CSF-cN AE at 72 hpf in V and DL cells in *espin*<sup>-/-</sup> (light blue, N = 2 fish), *espin*<sup>+/-</sup> (blue, N = 2 fish), and *espin*<sup>+/+</sup> (dark blue, N = 3 fish) siblings (1 representative experiment out of 2). In both CSF-cN subtypes, the AE size gradually decreases when cells miss 1 (*espin*<sup>+/-</sup>) or 2 (*espin*<sup>-/-</sup>) copies of the wild-type allele ( $p_V = 0.0112$ ,  $df = 74$ , and  $t = 2.3311$  in ventral cells;  $p_{DL} = 0.0017$ ,  $df = 45$ , and  $t = 3.0935$  in DL cells). (C) Overlay of calcium transients in ipsilateral DL CSF-cNs in response to passive tail bending induced by a glass probe in paralyzed control wild-type larvae versus *espin*<sup>+/-</sup> and *espin*<sup>-/-</sup> siblings at 5 days (120 hpf). The average across cells is shown in black (pulled data from 4 experiments). (D) The amplitude of CSF-cN calcium transients shown in (C) is represented as the ratio of peak fluorescence over baseline ( $\Delta R/R$ ) and is gradually reduced in *espin*<sup>+/-</sup> and *espin*<sup>-/-</sup> compared with wild-type siblings in a wild-type allele number-dependent manner ( $p = 1.1102 \times 10^{-6}$ ,  $df = 1309$ , and  $t = 8.462$ ). Underlying data can be found in [S1 Data](#). AE, apical extension; CSF-cN, cerebrospinal fluid-contacting neuron; df, degrees of freedom; DL, dorso-lateral; hpf, hours post fertilization; V, ventral; ZO-1, zonula-occludens-1.

<https://doi.org/10.1371/journal.pbio.3000235.g006>

apical extensions following a gradual decrease matching the *espin* gene copy number, with an intermediate phenotype in *espin*<sup>+/-</sup> heterozygotes (Fig 6B). The loss of Espin protein in *espin*<sup>-/-</sup> larvae led to a 15% reduction in apical extension area in ventral CSF-cNs and 20% in dorso-lateral cells (Fig 6B). The same effect was observed at 144 hpf (S4 Fig, panel A), demonstrating that the amount of Espin is critical for the morphogenesis of CSF-cN apical extensions.

Because *espin* and *crb1* mutant phenotypes appeared similar, we tested whether Crb1 function is required for Espin localization in CSF-cNs. 72-hpf *crb1*<sup>-/-</sup> larvae showed normal Espin enrichment at the CSF-cN apical extension. Conversely, we found Crb1 protein normally located at AJCs and apical extensions of CSF-cNs in 72-hpf *espin*<sup>-/-</sup> larvae (S5 Fig). We therefore concluded that, although *crb1* and *espin* mutations led to comparable morphological defects, the 2 proteins act through at least partially distinct regulatory pathways to control the morphogenesis of CSF-cN apical extensions.

We then investigated in more details the structure of CSF-cN apical extension in the *espin* mutant using stimulated emission depletion (STED) microscopy to improve imaging resolution compared with standard confocal microscopy (S4 Fig, panel B1). STED microscopy confirmed our previous observation that CSF-cN apical extensions exhibit high density of microvilli and lack of obvious spatial organization [21]. It also enabled us to clearly delineate the region of densely packed membrane—corresponding to actual protrusions—from the base of the apical region—devoid of intense membrane staining—corresponding to the junctional region labeled by ZO-1 (S4 Fig, panel B1, chevrons). The loss of Espin in *espin*<sup>-/-</sup> mutant larvae caused a strong reduction in the height of the membrane-rich protrusions. We therefore measured the vertical extension of the region above ZO-1-labeled junctions (S4 Fig, panel B2) and confirmed that the previously described reduction in the area covered by the apical

extension (Figs 5 and 6) is associated with a reduction in length of membrane protrusions in *espin* mutant larvae (36% reduction).

These results demonstrate that Espin actin-bundling activity is necessary for the morphogenesis of CSF-cN apical extensions in a dose-dependent manner, through the regulation of microvilli elongation.

### The size of the CSF-cN apical extension tunes the amplitude of the mechanosensory response

We previously reported that CSF-cNs respond to mechanical stimuli associated with passive and active tail bending in zebrafish larvae [14]. To investigate the functional consequences of the shortening of CSF-cN apical extension, we tested CSF-cN mechanosensory function using the assay previously reported by Böhm and colleagues [14]. In paralyzed wild-type 120-hpf larvae, passive deformation of the tail with a glass probe induced large intracellular calcium transients in ipsilateral dorsolateral CSF-cNs (Fig 6C, S2 Movie). In *espin*<sup>+/-</sup> and *espin*<sup>-/-</sup> larvae, however, CSF-cN calcium transients in response to tail bends were smaller, correlated with wild-type *espin* gene copy number (Fig 6C and 6D and S2 Movie) (38% reduction in *espin*<sup>+/-</sup> and 60% in *espin*<sup>-/-</sup>).

Similarly, *crb1*<sup>-/-</sup> CSF-cNs displayed reduced calcium transients compared with cells in wild-type siblings (S6 Fig, panels A and B) (79% reduction of the median).

The CSF-cN mechanoreponse to tail bending was previously found to rely on the activity of the transient receptor potential channel Pkd2l1, enriched at the apical extension [8,13,14]. We therefore checked the localization of the channel in *espin* and *crb1* mutants and found that Pkd2l1 was normally present at the apical extension of CSF-cNs (S6 Fig, panels C and D). It is therefore unlikely that the reduced mechanoreponse observed in mutant cells are explained by defects of Pkd2l1 addressing to the apical extension.

These observations in *crb1* and *espin* mutant larvae demonstrate that the size of CSF-cN apical extension sets the amplitude of the cell mechanosensory response to spine curvature.

## Discussion

### CSF-cNs retain their apical junctional complexes and the accompanying actin ring

To investigate how CSF-cNs develop their apical extension, we undertook a descriptive approach based on long-term time-lapse imaging to monitor F-actin in single cells. We show that early during their differentiation, CSF-cNs display an apical ring of actin beneath the AJCs that they retain throughout differentiation. It is generally thought that newborn cells fated to become neurons withdraw their apical domain—including the primary cilium—while they differentiate [23,27]. Although we cannot exclude an early delamination event followed by the re-establishment of an apical domain before the expression driven by the *pkd2l1* promoter, our results strongly suggest that CSF-cNs are not subject to apical abscission or retraction but rather differentiate while retaining their apical cell polarity and AJCs. This is consistent with previous findings that differentiated CSF-cNs still possess a cilium at their apical extension [5,6,14]. The loss of apical domain has been proposed to allow neuronal differentiation by preventing Sonic Hedgehog (Shh) cilium-dependent signaling [27]. Zebrafish spinal CSF-cNs remain in contact with the source of the Shh ligand, the CSF, throughout differentiation but attenuate their responsiveness to Shh through a decrease in Notch signaling [46]. Whether this constant exposure to the CSF allows the reception of other signals than Shh important for their differentiation remains to be determined.

### Actin-based protrusions originate from the actin ring

Our time-lapse movies suggest that the ring of actin is the site of emergence of actin-based protrusions to form the apical extension. This actin ring could therefore represent a site for the recruitment of the machinery necessary for protrusion initiation, including the IRSp53 family of proteins that we found specifically expressed in CSF-cNs. This idea is further supported by the observation of actin rings at the apical side of the mechanosensory hair cells of the zebrafish lateral line prior to the sensory microvilli differentiation [47]. The initiation of the apical extension directly from the apical actin belt would ensure the correct positioning of the sensory apparatus and its development directly into the extracellular compartment in contact with the medium from which it senses. The establishment of a strong apicobasal polarity and the specific organization of AJCs and F-actin in circumferential belts are landmarks of microvilliated cell types [48]. Therefore, our hypothesis of an involvement of the F-actin circumferential belt in the arising of protrusions offers a new avenue to solve the long-standing problem of apical microvilli initiation in different cell types, like intestinal brush border epithelial cells or microvilliated sensory cells [31,49].

### I-BAR containing factors as initiators of the actin-based protrusions in CSF-cNs

We show that CSF-cNs specifically express transiently at least 2 factors containing an I-BAR domain (for Inverse Bin-Amphiphysin-Rvs167) at 24 hpf—*baiap2a* and *baiap211b*. I-BAR domain proteins are capable of binding both actin cytoskeleton and membrane lipids, which enables them to simultaneously sense and induce membrane curvature [50,51]. We know from the literature that I-BAR family members are capable of generating negative membrane curvature and modulating actin dynamics to induce the formation of filopodia [30]. However, the role of these factors in the establishment of more stable microvilli in sensory cell types has never been investigated. Loss of function approaches are challenging in zebrafish due to a high redundancy of IRSp53 factors, and we failed to see an effect with a DN form. To our knowledge, this study is the first evidence for their possible involvement in the initiation of the actin-based protrusions forming a sensory apical extension.

### Special role of Crb1 in CSF-cN apical extension morphogenesis

Our results suggest that polarization factors involved in the establishment of the apical domain and located at the vicinity of AJCs, near the ring of actin, are critical for the formation of the apical extension. One of these factors, Crb1, particularly drew our attention. First, Crb1 is specifically retained in spinal CSF-cNs at 72 hpf, while being down-regulated in the other cells lining the central canal. Second, Crb1 shows distinct patterns of expression in ventral CSF-cNs, where it forms a ring of actin at the vicinity of the AJCs, and dorso-lateral cells, where it covers the entire apical domain and protrusions. Our dominant-negative and knock-out strategies to induce Crb1 loss of function failed to cause early polarization defects in CSF-cNs, a result that can be explained by the high functional redundancy between intracellular domains of the different Crb proteins in zebrafish [52].

However, we find that the spatial restriction of Crb1 to CSF-cN apical domain during differentiation participates in the proper morphogenesis of apical extensions. Our results suggest that the differential localization of Crb1 in the 2 CSF-cN subtypes contributes to the observed difference in size between ventral and dorsolateral cells but does not fully explain the difference in terms of shape [21]. Further investigation will be required to address the mechanisms underlying the formation of distinct apical extension morphologies among CSF-cNs.

Crb proteins have been demonstrated to regulate morphogenesis of sensory protrusions in zebrafish and fly photoreceptors [53,54] as well as wing hairs in *Drosophila* [55]. In humans, mutations in the *crb1* gene, mostly in the region encoding the extracellular domain, causes inherited retinal dystrophies [52]. It was shown that the unique *Drosophila* Crb plays a critical role in the morphogenesis of microvilli-based structures in the photoreceptor cells, rhabdomeres [54]. This role involves the extracellular domain of the protein, whose function is less known and is unique to Crb1 in vertebrates [52,53]. Together with these results, our work suggests that Crb1 function in actin-based protrusion formation is conserved in vertebrates and calls for further studies on the role of the unique Crb1 extracellular domain in the morphogenesis of sensory apical extensions.

### The actin-bundling protein Espin and its interactor Myo3b are key players in the formation of CSF-cN actin-based protrusions

We also show that CSF-cNs express *myo3b* and *espin* during their differentiation. Myo3b is a class III unconventional myosin, whereas Espin is an actin-bundling factor. Actin-bundling factors are able to cross-link actin filaments together and form higher-magnitude actin structures, such as microvilli [31,56,57]. In particular, Espin is commonly found in mechano- and chemosensory microvilliated cells in which this factor is a critical actin organizer in the development and maintenance of the sensory apparatus [58,59]. ESPIN was also found to be a hallmark of the CSF-cN population in the mouse central nervous system in a large-scale single-cell sequencing study [60]. Previous work demonstrated that Espin and Myo3b are interdependent to be transported to the tip of actin-based protrusions, where Espin ensures the lengthening of actin filaments through its actin-bundling activity [34,36]. Our results indicate that CSF-cNs use Myo3b and Espin in a similar fashion to ensure the lengthening of actin-based protrusions forming the apical extension.

Our work further shows that suppressing one actin-bundling factor, Espin, shortens but does not suppress the actin-based CSF-cN apical extension. It has been shown that the development of bundles of actin filaments involves a combination of at least 2 different factors, usually in a sequential manner [29]. The activity of several actin cross-linkers is required to subsequently initiate loose bundling and then tightly pack the bundles into an organized functional structure [61]. Therefore, our work suggests that other actin-bundling factors, yet to be identified, are necessary during the earlier phases of microvilli initiation and elongation of the CSF-cN apical extension.

It is interesting to note here that Baiap factors contain an SRC Homology 3 (SH3) domain, which was shown to strongly interact with the proline-rich regions present in the longest isoforms of Espin [43]. Our results indicate that members of the Baiap family could potentially act together with Espin and Myo3b to coordinate the initiation and the elongation of actin-based protrusions in CSF-cNs.

We describe comparable morphological defects resulting from both *crb1* and *espin* mutations, namely, a reduction of the apical extension size in mutant cells. Because we observe no mislocalization of Espin in *crb1* mutant cells and, conversely, no mislocalization of Crb1 in *espin* mutant cells, it is unlikely that the 2 proteins regulate CSF-cN apical extension morphogenesis through the same pathway. Consequently, the correct morphogenesis of the CSF-cN apical extension appears to rely on at least 2 distinct mechanisms, whose interactions remain unclear. Further investigation of the putative interplay between Crb1, Espin, and their respective interactors, as well as how the 2 pathways converge to support the formation of CSF-cN apical extensions, will bring essential elements to decipher molecular mechanisms underlying the morphogenesis of sensory microvilli organelles.

## The length of microvilli tunes mechanosensory response in CSF-cNs

The functional relevance of Espin has been largely documented in hair cells of the inner ear in which Espin loss is associated with incorrect lengthening of stereocilia, which leads to degeneration of hair cells and deafness [37,62,63]. In this study, we show that loss of Espin in CSF-cNs does not induce degeneration of these sensory cells but impacts, in a dose-dependent manner, microvilli length within the apical extension, a result consistent with the current understanding of Espin actin-bundling functions [45].

The reduction of microvilli length without CSF-cN degeneration in *espin* and *crb1* mutants provided a unique opportunity to investigate the contribution of microvilli to the process of sensory transduction. We show that the reduction of microvilli length is associated with a reduction in amplitude of CSF-cN mechanosensory response to passive bending of the spinal cord, indicating that the CSF-cN apical extension acts as a sensory antenna where mechanotransduction occurs. This observation is consistent with previous studies from our group and others demonstrating that CSF-cN-mediated detection of spinal bending [14], CSF flow and pressure applied against the membrane [11,13] requires the Pkd2l1 channel [13,14], which is confined to the apical extension [13].

The fact that the length of microvilli tunes the amplitude of CSF-cN sensory response while cells retain Pkd2l1 expression suggests that the mechanism underlying mechanotransduction in CSF-cNs is constrained by microvilli length. One possible explanation for such an effect could be that the length of microvilli determines the number of Pkd2l1 channels located at the membrane, which sets the amplitude of CSF-cN response to spinal cord bending. Longer microvilli may, furthermore, enable CSF-cNs to sample changes of CSF flow associated with spinal bending over larger volumes in the central canal and thereby amplify the mechanosensory response. Alternatively, proper lengthening of microvilli may enable the establishment of physical links between microvilli within the apical extension allowing for cohesion and/or mechanotransduction gating in a similar manner to hair cells [64,65]. The investigation of dynamic mechanisms underlying CSF-cN mechanotransduction in relation to the shape of microvilli will be the subject of future studies.

In hair cells of the inner ear, a large body of work has demonstrated the critical role of the staircase organization and patterned elongation of stereocilia for hearing [31]. Our results in CSF-cNs, where microvilli show no obvious organization, suggest that the lengthening of actin-based protrusions nonetheless finely modulates sensory function. Further physiological investigation in other microvilliated sensory cells such as vomeronasal olfactory receptor neurons or solitary chemosensory cells, will reveal whether our observation may be generalized to more microvilliated cells.

## Materials and methods

### Ethics statement

Animal handling and protocols were carried out with the validation of the Institut du Cerveau et de la Moelle épinière in agreement with the French National Ethics Committee (Comité National de Réflexion Ethique sur l'Expérimentation Animale, Ce5/2011/056) and European Communities Council Directive (2010/63/EU). Because experimentation on zebrafish larvae prior to 5 days old does not require approval of a protocol by the ethics committee, our project received the approval from the local ICM health and ethics committee.

### Animal care

Zebrafish adults and embryos and/or larvae were reared and maintained in a 14/10-hour light cycle. Embryos and larvae were raised at 28.5°C until the start of experiments conducted at

**Table 1. Transgenic lines used in our study.**

Allele name	Transgenic	Labeling	Reference
s1020tEt	<i>Tg(1020:Gal4)</i>	Motorneurons, CSF-cNs	[74]
Zf518Tg	<i>Tg(cdh2:cdh2-RFP, crybb1:eCFP)</i>	Cdh2	[40]
icm07	<i>Tg(pkcd2l1:GCAMP5G)</i>	CSF-cNs	[14]
icm10	<i>Tg(pkcd2l1:Gal4)</i>	CSF-cNs	[39]
icm17	<i>Tg(pkcd2l1:TagRFP)</i>	CSF-cNs	[14]
icm22	<i>Tg(UAS:TagRFP-CAAX;cmlc2:eGFP)</i>	Nonapplicable	[21]
icm28	<i>Tg(UAS:LifeAct-GFP;cryaa:V)</i>	Nonapplicable	[21]

**Abbreviations:** Cdh2, Cadherin 2; CSF-cN, cerebrospinal fluid-contacting neuron.

<https://doi.org/10.1371/journal.pbio.3000235.t001>

room temperature. When performed below 48 hpf, staging was assessed by counting somites (30 at 24 hpf) according to Kimmel and colleagues [66].

### Generation and use of transgenic lines

Transgenic lines used in this study are listed in Table 1. Transgenic lines generated using the Tol2 system are already described in [21].

### Generation of mutant lines and genotyping

We used the CRISPR/Cas9-mediated genome editing system to generate our mutant lines as listed in Table 2. crRNAs were designed using the online CRISPOR program (crispor.tefor.net/) [67] and selected to target the sequence of interest containing a restriction site (S1 and S3 Figs, panels A and B; see also Table 2). crRNA and universal 67mer tracrRNA were ordered from Integrated DNA Technologies. crRNA and tracrRNA were annealed in Duplex buffer and complexed with the Cas9 protein (provided by Jean-Paul Concordet lab) prior to injections into 1 cell-stage wild-type eggs. The efficacy of the crRNA was assessed directly after injection at 48 hpf by performing genotyping on usually 8 pools of 5 injected embryos. The genotyping workflow involved (1) isolating genomic DNA by lysing embryos in proteinase K solution (10 mM Tris [pH 8], 2 mM EDTA, 0.2% Triton X-100, 200 µg/mL proteinase K) after euthanasia in 0.2% MS-222 (Sigma-Aldrich), (2) performing PCR to amplify the target region (FastStart mix, Roche) using forward and reverse primers (as described in Table 2), and (3) digesting the PCR product with the corresponding restriction enzyme. In the presence of the icm26 mutation, the loss of the BstXI site on 1 (+/-) or 2 (-/-) alleles results in the conservation of the intact 350-bp PCR band instead of digested 172- and 182-bp bands in the wild type. In the presence of the icm31 mutation, the loss of the RsaI site on 1 (+/-) or 2 (-/-) alleles results in the conservation of the 25- and 325-bp PCR bands instead of 25-, 132-, and 195-bp bands in the wild type. When efficient editing was observed, siblings of the genotyped embryos

**Table 2. Mutant lines generated for this study.**

Allele name	Mutation	Gene	Target sequence (exon)	Restriction enzyme	PCR primers (5' to 3')
icm26	5-bp deletion	<i>espin</i>	CCAGAACAAAGACCAGCGTGG (exon 11)	BstXI	fw: CAAAACCCAACGACACCC rev: ACTTCAACTCATGTTGGCA
icm31	10-bp deletion	<i>crb1</i>	GTACGCTCTGCCAACTGTCC (exon 2)	RsaI	fw: TAACCCTCCTGACAGATGCA rev: AAGACCTCAACACTCTGCCT

**Abbreviations:** crb1, Crumbs 1; fw, forward; rev, reverse.

<https://doi.org/10.1371/journal.pbio.3000235.t002>

were raised to sexual maturity. This F0 generation was screened to find actual transmitters of mutations by crossing them to AB wild-type fish and genotyping the offspring at 48 hpf. When mutations were transmitted, siblings of the genotyped offspring were raised to F1, and adults were genotyped: a piece of fin tissue was sectioned under anesthesia (fin clipping) using 0.02% MS-222 (Sigma-Aldrich), and the PCR product was sequenced (GATC technology) using either the forward or the reverse primer described above to determine the type of mutation. For our study, we selected a 5-bp deletion in exon 11 of *espin* resulting in truncated Espin proteins lacking the conserved ABM (mutated allele named *espin<sup>icm26</sup>*) and a 10-bp deletion in exon 2 of *crb1*, resulting in a severely truncated Crb1 protein lacking all domains but one EGF-like extracellular (mutated allele named *crb1<sup>icm31</sup>*) (S1 and S3 Figs, panels B-C).

### Time-lapse imaging

To monitor the formation of the CSF-cN apical extension, we performed live imaging of single CSF-cNs labeled with LifeAct-GFP, a marker of F-actin [38], during early development. To obtain mosaic labeling, we injected approximately 1 ng of the DNA construct *UAS:LifeAct-GFP;cryaa:Venus* [21] with approximately 3.5 ng Tol2 transposase RNA into *Tg(pkd2l1:Gal4)* or *Tg(1020:Gal4)* eggs at the 1-cell stage. Imaging was initiated as soon as LifeAct-GFP-positive cells were detected in the spinal cord, usually between 20 and 22 hpf. LifeAct-positive embryos were first anaesthetized in 0.02% MS-222 (Sigma-Aldrich), then mounted laterally in 1.5% low-melting agarose in 50-mm glass-bottom dishes (MatTek #P50G-1.5-14-F). Because the imaging took place at early development of the fish, we made sure to set areas to image, containing at least 1 positive cell for LifeAct, large enough to not lose the cell over time because of developmental growth. Z-stacks spanning the entire cell depth were taken every 5 minutes during several consecutive 1-hour-long sessions using Slidebook software (Intelligent Imaging Innovations, Denver, CO). Imaging was performed using Zeiss ×40 or ×63 water-immersion objectives on a custom spinning disk microscope (Intelligent Imaging Innovations, Denver, CO) with a 488-nm laser.

### Fluorescent IHC and imaging

We followed a similar protocol in fluorescent-positive embryos and larvae after screening. First, samples were fixed in 4% PFA at +4°C for 4 hours, then washed in PBS Tween 0.01%, blocked for 2–5 hours at room temperature in 1% DMSO, 10% NGS, 0.5% Triton X-100 PBS 1×, and incubated with the following primary antibodies: mouse IgG1 anti-ZO-1 (1/200, Invitrogen #339100), chicken anti-GFP (1/500, Abcam #ab13970), mouse IgG1 anti-RFP (1/500, Thermo Fisher #ma5615257) or rabbit anti-RFP (1/500, Life Technologies #R10367), rabbit anti-Crb1 (1/500; kind gift from Wei lab), rabbit anti-Espin (50 ng/μL, kind gift from Bartles lab), and rabbit anti-Myo3b (1/500; kind gift from Petit lab) in 1% DMSO, 1% NGS, 0.5% Triton X-100 PBS 1X at +4°C overnight. After washing 3 times for at least 10 minutes in PBS 0.1X Tween 0.01% (PBST), samples were incubated with the secondary antibody (usually diluted to 1/500) and DAPI (1/4,000) in the dark, overnight, at +4°C. The next day, embryos or larvae were washed again in PBST for at least 3 times for 30 minutes, and proper staining was assessed by checking the pattern of fluorescence of DAPI control staining before proceeding to mounting. Pkd2l1 staining was performed following an adapted protocol using a polyclonal rabbit antibody as described by Sternberg and colleagues [13]. Samples were mounted using a mounting medium (Vectashield Antifade Mounting) and imaged using an Olympus FV1000 confocal microscope equipped with a ×63 oil immersion objective using the 405, 473, 543 or 567, and 633 or 647 nm laser lines.

### Detailed analysis of the CSF-cN apical extension morphology

Confocal images of the CSF-cN apical extension on spinal cord microtome sections were acquired following the procedure presented in our previous work [21]. We used the stable *Tg(UAS:TagRFP-CAAX;cmlc2:eGFP)* transgenic line crossed to *Tg(pkd2l1:Gal4)* to access the CSF-cN apical extension morphology. To assess the area covered by the CSF-cN apical extension in wild-type or *espin* and *crb1* mutant siblings, we drew polygons outlining the apical extension and the soma on Z-projections using the polygon tool in Fiji.sc/ [68] and used the “area” measurement. For each cell, we normalized the obtained area value by soma area.

### FISH and imaging

All the probes were generated by amplifying coding fragments from zebrafish 72-hpf larva total cDNA. The resulting PCR product was extracted from gel using the QIAquick Gel extraction kit (QIAGEN, Germany) and cloned into the pCRII-Blunt-TOPO vector (Life Technologies, Carlsbad, CA). Clones and orientation were verified by sequencing. For *espin*, the selected plasmid was linearized using NotI, and the resulting product was purified and measured for DNA concentration. For *myo3b*, *baipa2a*, and *baiap2l1b*, the product was generated by a second high-fidelity PCR using T7 forward primer, the corresponding reverse primer (Table 3) and Phu polymerase (Phusion High-Fidelity DNA polymerase, New England Bio-Labs). Digoxigenin-labeled antisense probes were synthesized from these clones using the T7 RNA polymerase with the RNA labeling kit (Sigma-Aldrich). All probes were then purified using the mini Quick Spin RNA columns (Roche, Basel, Switzerland). In the case of *myo3b*, *baipa2a*, and *baiap2l1b*, probes went through an additional step of hydrolysis by adding 40 mM of NaHCO<sub>3</sub> and 60 mM Na<sub>2</sub>CO<sub>3</sub> for 10 minutes at 60°C. The reaction was stopped and precipitated by adding 4 volumes of H<sub>2</sub>O, 0.33 volumes of 3 M NaOA, 0.025 volumes of glacial acetic acid, and 11 volumes of cold ethanol. The primers and restriction enzymes used for every probe are listed in Table 3. Whole-mount FISH was performed on 24-hpf embryos and 72-hpf larvae of the transgenic line *Tg(pkd2l1:GCaMP5G)* or *Tg(pkd2l1:TagRFP)* prior to fluorescent IHC to reveal GFP or RFP, respectively, to specifically label CSF-cNs. For FISH, samples were prepared by fixation in 4% PFA in PBS overnight at +4°C. At 72 hpf, larvae were treated with Proteinase K in PBST for 20 minutes at 100 µg/mL to ensure the penetration of the probes. Probes were diluted to 1/200 for hybridization overnight at +70°C. To reveal the expression of candidate genes, primary antidigoxigenin-POD coupled antibodies (Sigma-Aldrich) were added to 1/300. Staining was realized using either TAMRA (1/100) in *Tg(pkd2l1:GCaMP5G)* samples or FITC (1/200) in *Tg(pkd2l1:TagRFP)* tyramide-fluorophore solution for 30 minutes. Imaging was performed on an Olympus FV1000 confocal microscope equipped with a ×20 water immersion objective.

**Table 3. Primers and restriction enzymes used to generate the FISH probes in this study.**

Gene	Primers (from 5' to 3')	Product size (bp)	Restriction enzyme	Riboprobe hydrolysis?
<i>espin</i>	fw: AAGATCTACCGCTTCCTCCACC rev: CCAGCTTGGTTTCATCATAGCCT	825	NotI	No
<i>myo3b</i>	fw: TGTTTGATAAAGGATTTTGAGACTCGTC rev: ATATTATATGTACAGCACATTCGGTAGC	3,155	No (use rev primer)	Yes
<i>baipa2a</i>	fw: GATAACAAACCACGAGCAGCAAAACAC rev: AGATAATACATGAGGGATTACAAAAACG	2,265	No (use rev primer)	Yes
<i>baiap2l1b</i>	fw: GAGTTATTGTGGAGTAAATCGACTGAAC rev: ACAACCTTAAAACAAGCTGAATAATCCG	1,937	No (use rev primer)	Yes

**Abbreviations:** FISH, fluorescent in situ hybridization; fw, forward; rev, reverse.

<https://doi.org/10.1371/journal.pbio.3000235.t003>

### Morphological analysis of cells lacking Myo3b

To interfere with Myo3b activity, we built a truncated version of the protein devoid of the motor domain, reasoning that it would act as a dominant-negative form by titrating the interactors of the functional Myo3b [34]. The part of the cDNA encoding the 235 C-terminal amino acids was amplified by PCR (Phusion High-Fidelity DNA polymerase, New England BioLabs) from zebrafish 72-hpf larva total cDNA, using primers containing attB sites for Gateway cloning (fw: 5'-GGGGACAAGTTTGTACAAAAAAGCAGGCTGAATTCACCATGAA TCTGATGTTGCGGGAAGTGATCG-3', and rev: 5'-GGGGACCACTTTGTACAAGAAAGCTGGGTCTCCGATCCTCTGGAAGAGATAGTGGGCGG-3'). The stop codon was removed, and a GSG sequence was added to allow efficient coexpression of T2A-eGFP [69]. The PCR product was then recombined into the pDONR221 and assembled into the final expression vector in a 3-fragment gateway reaction (ThermoFisher Scientific) using p5E-10XUAS, pME-Myo3b-DN-GSG, p3E-T2A-eGFP, and pDest-cryaa:Venus. p3E-T2A-EGFP was subcloned from pME-HA-UPRT-T2A-EGFP (gift of Teresa Nicolson) [70] using att-site tagged primers (fw: 5'-GGGGACAGCTTTCTTGTACAAAGTGGGGTTAACGGCAGTG GAGAGGG-3', rev: 5'-GGGGACAACCTTTGTATAATAAAGTTGGTTACTTGTACAGC TCGTCCATG-3'). The resulting fragment was inserted into pDONR p2rp3 in a BP recombination reaction (Thermo Fisher). The resulting plasmid was injected into 1-cell eggs carrying a combination of *Tg(UAS:TagRFP-CAAX;cmlc2:eGFP)* and *Tg(pkd2l1:Gal4)*. Confocal images of laterally mounted spinal cords were acquired, and we performed a Z-projection of individual ventral CSF-cNs either expressing Myo3b-DN (GFP and RFP positive) or control (RFP-only positive). The area of the apical extension and the cell body was assessed by drawing a polygon around the structure and measuring the area covered using Fiji.ct/ [68] and normalized by soma area.

### Morphological analysis of *espin<sup>icm26/icm26</sup>* mutant cells and rescue of the phenotype

To assess morphological consequences of the loss of function of Espin in *espin<sup>icm26/icm26</sup>* cells and to compare them with previous observations in cells lacking Myo3b function, we proceeded to the same analysis on confocal images of laterally mounted spinal cords at 72 hpf. To further confirm that the loss of Espin function is responsible for the observed phenotype in both conditions (*espin<sup>icm26/icm26</sup>* and Myo3b-DN), we generated a plasmid to rescue the sole expression of Espin. The cDNA of the Espin isoform 1 was amplified by PCR (Phusion High-Fidelity DNA polymerase, New England BioLabs) from zebrafish 72-hpf larva total cDNA using primers containing attB sites for Gateway cloning (forward: 5'-CACCATGGTGGTTGA AAGGACACTTCTCGC-3' and reverse: 5'-TCCGGATCCCTGTTTCGCTATGTCTCC-3'). The stop codon was removed, and a GSG sequence was added to allow efficient coexpression of T2A-RFP-NLS [69]. The resulting PCR product was recombined into the pENTR TOPO (pENTR/D-TOPO Cloning Kit, ThermoFisher Scientific) to generate pME-zEspin1-GSG. The final rescue vector was assembled in a 3-way Gateway reaction (ThermoFisher Scientific) using p5E-10XUAS, pME-zEspin1-GSG, p3E-T2A-tagRFP-NLS, and pDest-cryaa:mCherry. p3E-T2A-TagRFP-NLS was generated by inserting the TagRFP-NLS cDNA into a p3E-T2A plasmid [71] linearized with AvrII (NEBuilder HiFi DNA Assembly Cloning Kit, New England BioLabs). The TagRFP cDNA was amplified from pME-TagRFP-CAAX [72] by PCR using forward 5'-CGAGGAGAATCCTGGCCCACTTGTGTCTAAGGGCGAAGAG-3' and the nuclear targeting sequence NLS-tagged reverse 5'-GACCGAAATTAATTAATAAACTTAGAC TTTCTTCTTCTTGGGATTAAGTTTGTGCC-3' primers. The resulting rescue plasmid was injected into 1-cell eggs obtained from *Tg(pkd2l1:Gal4,UAS:TagRFP-CAAX;cmlc2:*

*eGFP)espin<sup>icm26</sup>* incross allowing the sparse expression of the rescue zEspin1 in *espin<sup>+/+</sup>*, *espin<sup>+/icm26</sup>*, and *espin<sup>icm26/icm26</sup>* embryos and larvae.

### STED microscopy imaging of the CSF-cN apical extension

To gain in confocal imaging resolution, we undertook STED microscopy imaging followed by deconvolution treatment on spinal cord microtome sections from a combination of *Tg(UAS:TagRFP-CAAX;cmlc2:eGFP)* and *Tg(pk211:Gal4)* transgenic lines carrying the *espin<sup>icm26</sup>* mutation. The original IHC protocol [21] was adapted using the goat antirabbit IgG Alexa Fluor 594 secondary antibody (#A-11012; ThermoFisher) diluted to 1/200 to reveal the membrane-tagged TagRFP labeling in CSF-cNs in a STED-stable manner. The ZO-1 staining to outline the central canal was kept unmodified. Imaging of both ZO-1 and TagRFP-CAAX was carried out on a Leica TCS SP8 STED 3× confocal microscope using a 93× glycerol objective. ZO-1 labeling was excited with a white laser at 488 nm (3.5%) and detected on a HyD detector from 500 to 550 nm. TagRFP-CAAX labeling was excited with a white light laser at 590 nm (8.5%), depleted with a 775 nm depletion laser (80%), and detected on a HyD detector from 600 to 650 nm, gated from 0.6 to 5.37 ns. Images were then processed by deconvolution using Huygens software (SVI, the Netherlands). To assess the vertical extension of CSF-cN apical extensions in *espin* wild-type and mutant fish, we drew polygons outlining the apical extension and fitted an ellipse to extract the “minor” measurement using the polygon tool in Fiji.sc/ [68].

### Functional analysis of the mutants

To assess the sensory function of the CSF-cNs in mutant versus wild-type siblings, we performed the same passive tail bending assay as described in the work by Böhm and colleagues [14]. *Tg(pk211:GCaMP5G,pkd211:TagRFP)espin<sup>icm26</sup>* adult fish were incrossed, and offspring were screened for GCaMP5G and TagRFP fluorescence at 72 hpf after light anesthesia in 0.02% MS-222 (Sigma-Aldrich). At 120 hpf, double positive larvae were embedded in 1.5% low-melting-point agarose and paralyzed by injecting 0.5 nL of  $\alpha$ -Bungarotoxin (0.5 mM; Tocris #11032-79-4) in the musculature above the swimming bladder. Half of the tail was freed unilaterally to give access to the glass probe. Calcium imaging was performed using a 2-photon microscope (Intelligent Imaging Innovations) equipped with a ×20 water objective. Both GCaMP and RFP signals were recorded simultaneously at 5 Hz for 770 frames, allowing for 11 probe deflections to bend the tail every 70 frames, starting at 10 frames. Each stimulation therefore lasted 14 seconds. A new version of the Matlab script used in the work by Böhm and colleagues [14] was generated to perform the analysis. Regions of interest were manually selected in Matlab by drawing the CSF-cN cell bodies in the region of deflection. In order to correct for artefactual and/or experimental movements, we registered all image time series based on a TagRFP reference frame built as the average over 50 frames in a quiet period after the first stimulation. Background noise, calculated in a region with no signal, was subtracted to every frame. Raw signal was extracted for both GCaMP and TagRFP channels, and, to correct for motion artefacts, the 3 frames following the stimulation were removed, and the ratio  $\Delta R/R$  was calculated as  $(F_{\text{GCaMP}}(t) \times F_{0-\text{TagRFP}}) \div (F_{\text{TagRFP}}(t) \times F_{0-\text{GCaMP}})$  at every time point left with  $F_0$  the average fluorescence of the 10 frames preceding the stimulation. The calcium transient amplitude was calculated as the difference  $(\Delta R/R)_{\text{post}} - (\Delta R/R)_{\text{pre}}$ , with  $(\Delta R/R)_{\text{post}}$  the average of the 3 frames right after stimulation and  $(\Delta R/R)_{\text{pre}}$  the average of the 10 frames before stimulation.

### Statistics

All statistics were run on Matlab (Mathworks, Massachusetts). In most cases, student *t* tests were used for hypothesis testing. For obviously skewed distributions, we used the two-sample

Kolmogorov-Smirnov test. Values are mean (red line) or median (black line). When displayed, the bottom and top edges of box plots indicate the 25th and the 75th percentiles, respectively, and whiskers extend to the most extreme data points not considered as outliers (outliers are plotted individually). To test the impact of wild-type *espin* allele numbers on apical extension size and calcium responses, we used multiple regression. More precisely, we relied on the general linear model (GLM) [73]: the design matrix included 3 regressors of interest (encoding the number of alleles: +/+, +/-, and -/-) and, when applicable, 1 confounding variable (day of the experiment). Statistical significance of the effect of interest (above and beyond confounding factors) was tested using a *t* test. In all figures, \* means that  $p < 0.05$ , \*\* means that  $p < 0.01$ , \*\*\* means that  $p < 0.001$ , and \*\*\*\* means that  $p < 0.0001$ .

## Supporting information

**S1 Data. Excel spreadsheet containing, in separate sheets, the underlying numerical data for Figs 3B, 5B, 5C, 6B, 6C, 6D, S4A, S4B2, S6A and S6B.**

(XLSX)

**S1 Movie. Time-lapse imaging showing 3 CSF-cNs going from Stage 1 to Stage 3.** Typical time-lapse recording used to obtain the results presented in Fig 1. A Z-stack spanning the entire cells was taken every 5 minutes in a live embryo mounted laterally in 1.5% low-melting agarose from 22 hpf and over 8.5 hours. Three cells (1 ventral and 2 dorsolateral cells) were individually marked by LifeAct-GFP under the control of the *pkd2l1* promoter. CSF-cN, cerebrospinal fluid-contacting neuron; hpf, hours post fertilization.

(AVI)

**S2 Movie. Neuronal activity in response to passive tail bending reported by calcium imaging in 120-hpf (5-day-old) wild-type and *espin*<sup>-/-</sup> mutant larvae.** Typical recordings supporting the results in Fig 6C and 6D. TagRFP and GCAMP5G, a green fluorescent calcium indicator, were coexpressed in CSF-cNs under the control of the *pkd2l1* promoter in both wild-type (“+/+,” left panel) or mutant *espin*<sup>-/-</sup> (“*icm26/icm26*,” right panel) 120-hpf larvae. The 2 signals were recorded at 5 Hz with a 2-photon laser-scanning microscope while the spinal cord of paralyzed animals was deflected with a glass probe. TagRFP signal (upper panels) is used as a reference to correct for motion artefact in all 3 dimensions. GCAMP5G (middle panels) fluorescence varies with calcium concentration and sensory activity. The change of ratio ( $\Delta R/R$ ) between the 2 signals was used to quantify neuronal activity (lower panels, representing traces of different ROIs tracked on the movies above). CSF-cN, cerebrospinal fluid-contacting neuron; hpf, hours post fertilization; RFP, red fluorescent protein; ROI, region of interest.

(AVI)

**S1 Fig. *Crb1* locus organization and generation of the *crb1*<sup>*icm31*</sup> mutant.** (A) Localization and genomic structure of the unique *crb1* locus in zebrafish on Chromosome 22. (B, Top) Genomic region targeted by the sgRNA in exon 2 (sequence in bold), the earliest compatible target region containing a restriction site, here for *RsaI*, which is lost when editing occurs and enables a 2-step genotyping with a PCR followed by *RsaI* digestion. (Bottom) Sequence of the *crb1*<sup>*icm31*</sup> allele generated showing the 10-bp deletion generated by the CRISPR-Cas9 genome editing technique. The early frameshift results in an amino acid sequence disturbed from early on (green) leading to a premature stop codon. (C) Schematics showing the predicted mutant truncated *Crb1* protein obtained with the *crb1*<sup>*icm31*</sup> 10-bp deletion. Green boxes, EGF-like domains; violet boxes, laminin G-like domains. (D) IHC for *Crb1* (cyan) showing the loss of immunoreactivity in TagRFP-CAAX-positive CSF-cNs (magenta) in 72-hpf *crb1*<sup>-/-</sup> larvae

compared with wild-type siblings. Scale bars, 10  $\mu\text{m}$ . Crb1, Crumbs 1; CSF-cN, cerebrospinal fluid-contacting neuron; EGF, epidermal growth factor; IHC, immunohistochemistry; PAM, protospacer adjacent motif; sgRNA, single guide RNA.

**S2 Fig. Myo3b and Espin are enriched at the AE of microvilliated sensory cells.** (A) IHC for Myo3b shows the enrichment of the protein (cyan) at the level of AEs of TagRFP-CAAX-positive CSF-cNs (magenta) in 72-hpf larvae. Scale bar, 10  $\mu\text{m}$ . (B) IHC for Espin was performed on whole-mount zebrafish 72-hpf larvae. Scale bar, 100  $\mu\text{m}$ . Espin is enriched at the apical extension of various microvilliated sensory cell types: olfactory neurons in the olfactory pit (B1), hair cells of the inner ear (B2), lateral line hair cells (B3), and CSF-cNs (B4). Scale bars, 10  $\mu\text{m}$ . AE, apical extension; CSF-cN, cerebrospinal fluid-contacting neuron; hpf, hours post fertilization; IHC, immunohistochemistry.

**S3 Fig. Espin locus organization and generation of the *espin<sup>icm26</sup>* mutant.** (A) Localization and genomic structure of the unique *espin* locus in zebrafish on Chromosome 8. The conserved actin-bundling module is encoded by exons 11 to 13. (B, Top) Genomic region targeted by the sgRNA in exon 11 (sequence in bold), right upstream of the coding sequence for the actin-bundling module (amino acid sequence indicated in blue). The target region contains a BstXI digestion site, upstream of the PAM, which is disabled when editing occurs. (Bottom) Sequence of the *espin<sup>icm26</sup>* allele generated showing the 5-bp deletion generated by the CRISPR-Cas9 genome editing technique. In *espin<sup>icm26</sup>*, the coded amino acid sequence of the actin-bundling module is disturbed from the first codon (green). (C) Schematics showing the predicted mutant truncated Espin protein obtained with the *espin<sup>icm26</sup>* 5-bp deletion. The actin-bundling module is entirely disabled (white box). Green boxes, ankyrin-like repeats; violet boxes, proline-rich regions; red box, WH2 domain; blue box, actin-bundling module. (D) IHC for Espin (cyan) showing the loss of immunoreactivity in TagRFP-CAAX-positive CSF-cNs (magenta) in 72-hpf *espin<sup>-/-</sup>* larvae compared with wild-type siblings. (E) IHC for Espin showing the gradual loss of immunoreactivity in CSF-cNs of *espin<sup>+/-</sup>* and *espin<sup>-/-</sup>* compared with *espin<sup>+/+</sup>* 72-hpf larvae. Samples were analyzed simultaneously, and images were acquired and treated with the same parameters. Scale bars, 10  $\mu\text{m}$ . CSF-cN, cerebrospinal fluid-contacting neuron; hpf, hours post fertilization; IHC, immunohistochemistry; PAM, protospacer adjacent motif; sgRNA, single guide RNA; WH2, WASP (for Wiskott-Aldrich Syndrome protein) homology 2.

**S4 Fig. Espin is required for the proper lengthening of CSF-cN microvilli.** (A) Quantification of the area covered by the CSF-cN apical extension at 144 hpf (6 days) in ventral and dorso-lateral cells in *espin<sup>-/-</sup>* mutant larvae (light blue;  $N = 8$  fish) compared with wild-type siblings (dark blue;  $N = 4$  fish). Both CSF-cN subtypes lacking Espin show a significant reduction of the area covered by their apical extension as observed at 72 hpf ( $p_{\text{ventral}} = 0.0019$  and  $p_{\text{dorso-lateral}} = 0.0164$ ). (B1) STED confocal images from spinal cross sections of 72-hpf *espin<sup>+/+</sup>* or *espin<sup>-/-</sup>* larvae showing apical extensions of ventral TagRFP-CAAX-positive CSF-cNs. The junctional region is highlighted by ZO-1 staining (green). Scale bars, 1  $\mu\text{m}$ . (B2) Quantification of the vertical extension of ventral CSF-cN apical extensions at 72 hpf in *espin<sup>-/-</sup>* mutant larvae (light blue) versus wild-type larvae (dark blue) from STED images obtained as in (B1). Mutant cells formed significantly shorter apical extensions ( $p = 2.5571 \times 10^{-4}$ ), suggesting the critical role of Espin actin-bundling activity for the proper lengthening of CSF-cN microvilli. Underlying data can be found in [S1 Data](#). AE, apical extension; CSF-cN, cerebrospinal fluid-contacting

neuron; hpf, hours post fertilization; STED, stimulated emission depletion; ZO-1, zonula-occludens-1.  
(TIF)

**S5 Fig. *espin* and *crb1* mutant CSF-cNs retain the expression of Crb1 and Espin proteins, respectively.** (A) IHC for Crb1 (cyan) in TagRFP-CAAX-positive CSF-cNs (magenta) of 72-hpf larvae shows that Crb1 is similarly expressed and located in mutant CSF-cNs compared with wild-type cells. (B) IHC for Espin (cyan) and ZO-1 (magenta) to highlight the junctional region in *crb1*<sup>-/-</sup> larvae shows that Espin is similarly expressed and enriched at the apical extension of mutant V and DL CSF-cNs compared with wild-type cells. Scale bars, 10 μm. Crb1, Crumbs 1; CSF-cN, cerebrospinal fluid-contacting neuron; DL, dorso-lateral; hpf, hours post fertilization; IHC, immunohistochemistry; V, ventral; ZO-1, zonula-occludens-1.  
(TIF)

**S6 Fig. CSF-cNs with shorter apical extensions exhibit reduced mechanoreponse in a Pkd2l1-independent manner.** (A) Overlay of calcium transients in ipsilateral dorsolateral CSF-cNs in response to tail bending induced by a glass probe in paralyzed wild-type versus *crb1*<sup>-/-</sup> 120-hpf animals (data pooled from 3 experiments). (B) The amplitude of CSF-cN calcium transients shown in (A) is represented as the ratio of peak fluorescence over baseline ( $\Delta R/R$ ) and is significantly reduced in *crb1*<sup>-/-</sup> mutant compared with wild-type siblings ( $p = 1.012 \times 10^{-5}$ ). (C and D) IHC for Pkd2l1 channel (cyan) in *espin*<sup>-/-</sup> (C) or *crb1*<sup>-/-</sup> (D) shows that TagRFP-CAAX-positive CSF-cNs (magenta) retain the expression of the channel at their apical extension in mutant larvae similarly to wild-type siblings. Scale bars, 10 μm. Underlying data can be found in [S1 Data](#). AE, apical extension; CSF-cN, cerebrospinal fluid-contacting neuron; hpf, hours post fertilization; IHC, immunohistochemistry.  
(TIF)

## Acknowledgments

We thank Dr. Céline Revenu for the *Tg(cdh2:cdh2-RFP, crybb1:eCFP)Zf518Tg* line; Professor James R. Bartles, Professor Christine Petit, and Professor Xiangyun Wei for rabbit polyclonal antibodies against Espin, Myo3b, and Crb1, respectively; Dr. Jean-Paul Concordet for Cas9 protein; and Professor Katie Kindt for sharing her movie showing rings of actin during neuro-mast hair cell development and feedbacks on the manuscript. We thank the ICM Quant imaging facility for extensive instrument use. We thank the NeurImag facility for STED microscopy. We thank Dr. Jean Daunizeau for advice on GLM statistical analysis. We thank Sophie Nunes-Figueiredo, Natalia Maties, Bogdan Buzurin, Monica Dicu, and Antoine Arneau for fish care; Dr. Yasmine Cantaut-Belarif for useful tips for performing IHC; Dr. Lydia Djenoune for guidance in carrying out FISH; Dr. Urs Böhm for training on calcium imaging analysis; and all members of the Wyart lab (<https://wyartlab.org>) for critical feedback on the project. We thank Dr. Céline Revenu and Professor Corinne Houart for their insightful feedback throughout the project as thesis advisory committee members.

## Author Contributions

**Conceptualization:** Laura Desban, Claire Wyart, Pierre-Luc Bardet.

**Funding acquisition:** Claire Wyart, Pierre-Luc Bardet.

**Investigation:** Laura Desban, Andrew Prendergast, Julian Roussel, David Geny, Claire Wyart, Pierre-Luc Bardet.

**Methodology:** Laura Desban, Andrew Prendergast, Julian Roussel, David Geny, Claire Wyart, Pierre-Luc Bardet.

**Resources:** Marion Rosello.

**Supervision:** Claire Wyart, Pierre-Luc Bardet.

**Visualization:** Laura Desban, David Geny.

**Writing – original draft:** Laura Desban, Claire Wyart, Pierre-Luc Bardet.

**Writing – review & editing:** Laura Desban, Claire Wyart, Pierre-Luc Bardet.

## References

1. Kessels MM, Schwintzer L, Schlobinski D, Qualmann B. Controlling actin cytoskeletal organization and dynamics during neuronal morphogenesis. *Eur J Cell Biol*. Elsevier GmbH.; 2011; 90: 926–933. <https://doi.org/10.1016/j.ejcb.2010.08.011> PMID: 20965607
2. Hudspeth AJ. How the ear's works work. *Nature*. 1989; 341: 397–404. <https://doi.org/10.1038/341397a0> PMID: 2677742
3. Zhao B, Müller U. The elusive mechanotransduction machinery of hair cells. *Curr Opin Neurobiol*. 2015; 34: 172–179. <https://doi.org/10.1016/j.conb.2015.08.006> PMID: 26342686
4. Djenoune L, Wyart C. Light on a sensory interface linking the cerebrospinal fluid to motor circuits in vertebrates. *J Neurogenet*. Informa UK Limited, trading as Taylor & Francis Group; 2017; 31: 113–127. <https://doi.org/10.1080/01677063.2017.1359833> PMID: 28789587
5. Vigh B, Vigh-Teichmann I. Comparative Ultrastructure of the Cerebrospinal fluid-contacting neurons. *Int Rev Cytol*. 1973; 35: 189–251. PMID: 4584250
6. Dale N, Roberts A, Ottersen OP, Storm-Mathisen J. The morphology and distribution of "Kolmer-Agduhr cells", a class of cerebrospinal-fluid-contacting neurons revealed in the frog embryo spinal cord by GABA immunocytochemistry. *Proc R Soc Lond*. 1987; B232: 193–203.
7. Orts-Del'Immagine A, Kastner A, Tillement V, Tardivel C, Trouslard J, Wanaverbecq N. Morphology, Distribution and Phenotype of Polycystin Kidney Disease 2-like 1-Positive Cerebrospinal Fluid Contacting Neurons in The Brainstem of Adult Mice. *PLoS ONE*. 2014; 9(2): e87748. <https://doi.org/10.1371/journal.pone.0087748> PMID: 24504595
8. Djenoune L, Khabou H, Joubert F, Quan FB, Nunes Figueiredo S, Bodineau L, et al. Investigation of spinal cerebrospinal fluid-contacting neurons expressing PKD2L1: evidence for a conserved system from fish to primates. *Front Neuroanat*. 2014; 8. <https://doi.org/10.3389/fnana.2014.00026> PMID: 24834029
9. Das Kolmer W., "Sagittalarorgan" der Wirbeltiere. *Anat Embryol (Berl)*. 1921; 652–717. <https://doi.org/10.1007/BF02593657>
10. Agduhr E. Über ein zentrales Sinnesorgan (?) bei den Vertebraten. *Z Anat Entwicklungsgesch*. 1922; 66: 223–360. <https://doi.org/10.1007/BF02593586>
11. Jalalvand E, Robertson B, Wallén P, Grillner S. Ciliated neurons lining the central canal sense both fluid movement and pH through ASIC3. *Nat Commun*. 2016; <https://doi.org/10.1038/ncomms10002> PMID: 26743691
12. Orts-Del'Immagine A, Wanaverbecq N, Tardivel C, Tillement V, Dallaporta M, Trouslard J. Properties of subependymal cerebrospinal fluid contacting neurones in the dorsal vagal complex of the mouse brainstem. *J Physiol*. 2012; 590: 3719–3741. <https://doi.org/10.1113/jphysiol.2012.227959> PMID: 22570378
13. Sternberg JR, Prendergast AE, Brosse L, Cantaut-Belarif Y, Thouvenin O, Orts-Del'Immagine A, et al. Pkd2l1 is required for mechanosensation in cerebrospinal fluid-contacting neurons and maintenance of spine curvature. *Nat Commun*. 2018; 1–10. <https://doi.org/10.1038/s41467-017-02088-w>
14. Böhm UL, Prendergast A, Djenoune L, Nunes Figueiredo S, Gomez J, Stokes C, et al. CSF-contacting neurons regulate locomotion by relaying mechanical stimuli to spinal circuits. *Nat Commun*. 2016; 7:10866: 1–8. <https://doi.org/10.1038/ncomms10866> PMID: 26946992
15. Huang AL, Chen X, Hoon MA, Chandrashekar J, Guo W, Tränkner D, et al. The cells and logic for mammalian sour taste detection. *Nature*. 2006; 442: 934–938. <https://doi.org/10.1038/nature05084> PMID: 16929298
16. Vigh B, Vigh-Teichmann I, Manzano e Silva MJ, van den Pol AN. Cerebrospinal fluid-contacting neurons of the central canal and terminal ventricle in various vertebrates. *Cell Tissue Res*. 1983; 231: 615–621. <https://doi.org/10.1007/BF00218119> PMID: 6871973

17. Petracca YL, Sartoretti MM, Di Bella DJ, Marin-Burgin A, Carcagno AL, Schinder AF, et al. The late and dual origin of cerebrospinal fluid-contacting neurons in the mouse spinal cord. *Development*. 2016; 143: 880–891. <https://doi.org/10.1242/dev.129254> PMID: [26839365](https://pubmed.ncbi.nlm.nih.gov/26839365/)
18. Yang L, Rastegar S, Strähle U. Regulatory interactions specifying Kolmer-Agduhr interneurons. *Development*. 2010; 137: 2713–2722. <https://doi.org/10.1242/dev.048470> PMID: [20610488](https://pubmed.ncbi.nlm.nih.gov/20610488/)
19. Park H-C, Shin J, Appel B. Spatial and temporal regulation of ventral spinal cord precursor specification by Hedgehog signaling. *Development*. 2004; 131: 5959–5969. <https://doi.org/10.1242/dev.01456> PMID: [15539490](https://pubmed.ncbi.nlm.nih.gov/15539490/)
20. Hubbard JM, Böhm UL, Prendergast A, Tseng P-EB, Newman M, Stokes C, et al. Intraspinal Sensory Neurons Provide Powerful Inhibition to Motor Circuits Ensuring Postural Control during Locomotion. *Curr Biol*. 2016; 26: 2841–2853. <https://doi.org/10.1016/j.cub.2016.08.026> PMID: [27720623](https://pubmed.ncbi.nlm.nih.gov/27720623/)
21. Djenoune L, Desban L, Gomez J, Sternberg JR, Prendergast A, Langui D, et al. The dual developmental origin of spinal cerebrospinal fluid-contacting neurons gives rise to distinct functional subtypes. *Sci Rep*. Springer US; 2017; 7. <https://doi.org/10.1038/s41598-017-00350-1> PMID: [28389647](https://pubmed.ncbi.nlm.nih.gov/28389647/)
22. Schwayer C, Sikora M, Slováková J, Kardos R, Heisenberg C-P. Actin Rings of Power. *Dev Cell*. 2016; 37: 493–506. <https://doi.org/10.1016/j.devcel.2016.05.024> PMID: [27326928](https://pubmed.ncbi.nlm.nih.gov/27326928/)
23. Paridaen JT, Huttner WB. Neurogenesis during development of the vertebrate central nervous system. *EMBO Rep*. 2014; 15: 351–364. <https://doi.org/10.1002/embr.201438447> PMID: [24639559](https://pubmed.ncbi.nlm.nih.gov/24639559/)
24. Singh S, Solecki DJ. Polarity transitions during neurogenesis and germinal zone exit in the developing central nervous system. *Front Cell Neurosci*. 2015;9. <https://doi.org/10.3389/fncel.2015.00009>
25. Zou J, Wen Y, Yang X, Wei X. Spatial-temporal expressions of Crumbs and Nagie oko and their interdependence in zebrafish central nervous system during early development. *Int J Dev Neurosci*. International Society for Developmental Neuroscience; 2013; 31: 770–782. <https://doi.org/10.1016/j.ijdevneu.2013.09.005> PMID: [24071007](https://pubmed.ncbi.nlm.nih.gov/24071007/)
26. Alexandre P, Reugels AM, Barker D, Blanc E, Clarke JDW. Neurons derive from the more apical daughter in asymmetric divisions in the zebrafish neural tube. *Nat Neurosci*. Nature Publishing Group; 2010; 13: 673–679. <https://doi.org/10.1038/nn.2547> PMID: [20453852](https://pubmed.ncbi.nlm.nih.gov/20453852/)
27. Das RM, Storey KG. Apical Abscission Alters Cell Polarity and Dismantles the Primary Cilium During Neurogenesis. *Science (80-)*. 2014; 343: 200–204. <https://doi.org/10.1126/science.1247521> PMID: [24408437](https://pubmed.ncbi.nlm.nih.gov/24408437/)
28. Faix J, Rottner K. The making of filopodia. *Curr Opin Cell Biol*. 2006; 18: 18–25. <https://doi.org/10.1016/j.cub.2005.11.002> PMID: [16337369](https://pubmed.ncbi.nlm.nih.gov/16337369/)
29. Revenu C, Athman R, Robine S, Louvard D. The co-workers of actin filaments: from cell structures to signals. *Nat Rev Mol Cell Biol*. 2004; 5: 635–646. <https://doi.org/10.1038/nrm1437> PMID: [15366707](https://pubmed.ncbi.nlm.nih.gov/15366707/)
30. Ahmed S, Goh WI, Bu W. I-BAR domains, IRSp53 and filopodium formation. *Semin Cell Dev Biol*. Elsevier Ltd; 2010; 21: 350–356. <https://doi.org/10.1016/j.semcdb.2009.11.008> PMID: [19913105](https://pubmed.ncbi.nlm.nih.gov/19913105/)
31. Barr-Gillespie P-G. Assembly of hair bundles, an amazing problem for cell biology. *Mol Biol Cell*. 2015; 26: 2727–2732. <https://doi.org/10.1091/mbc.E14-04-0940> PMID: [26229154](https://pubmed.ncbi.nlm.nih.gov/26229154/)
32. Nayak GD, Rathnayaka HSK, Goodyear RJ, Richardson GP. Development of the hair bundle and mechanotransduction. *Int J Dev Biol*. 2007; 51: 597–608. <https://doi.org/10.1387/ijdb.072392gn> PMID: [17891720](https://pubmed.ncbi.nlm.nih.gov/17891720/)
33. Michalski N, Petit C. Genetics of auditory mechano-electrical transduction. *Pflügers Arch—Eur J Physiol*. 2015; 467: 49–72. <https://doi.org/10.1007/s00424-014-1552-9> PMID: [24957570](https://pubmed.ncbi.nlm.nih.gov/24957570/)
34. Merritt RC, Manor U, Salles FT, Grati M, Dose AC, Unrath WC, et al. Myosin IIIB Uses an Actin-Binding Motif in Its Espin-1 Cargo to Reach the Tips of Actin Protrusions. *Curr Biol*. 2012; 22: 320–325. <https://doi.org/10.1016/j.cub.2011.12.053> PMID: [22264607](https://pubmed.ncbi.nlm.nih.gov/22264607/)
35. Lelli A, Michel V, Boutet De Monvel J, Cortese M, Bosch-Grau M, Aghaie A, et al. Class III myosins shape the auditory hair bundles by limiting microvilli and stereocilia growth. *J Cell Biol*. 2016; 212: 231–244. <https://doi.org/10.1083/jcb.201509017> PMID: [26754646](https://pubmed.ncbi.nlm.nih.gov/26754646/)
36. Ebrahim S, Avenarius MR, Grati M, Krey JF, Windsor AM, Sousa AD, et al. Stereocilia-staircase spacing is influenced by myosin III motors and their cargos espin-1 and espin-like. *Nat Commun*. 2016; 7: 10833. <https://doi.org/10.1038/ncomms10833> PMID: [26926603](https://pubmed.ncbi.nlm.nih.gov/26926603/)
37. Sekerková G, Richter C-P, Bartles JR. Roles of the Espin Actin-Bundling Proteins in the Morphogenesis and Stabilization of Hair Cell Stereocilia Revealed in CBA/CaJ Congenic Jerker Mice. *PLoS Genet*. 2011; 7(3): e1002032 <https://doi.org/10.1371/journal.pgen.1002032> PMID: [21455486](https://pubmed.ncbi.nlm.nih.gov/21455486/)
38. Riedl J, Crevenna AH, Kessenbrock K, Yu JH, Neurkirchen D, Bista M, et al. LifeAct: a versatile marker to visualize F-actin. *Nat Methods*. 2008; 5. <https://doi.org/10.1038/nmeth.1220> PMID: [18200000](https://pubmed.ncbi.nlm.nih.gov/18200000/)

39. Fidelin K, Djenoune L, Stokes C, Prendergast A, Gomez J, Baradel A, et al. State-Dependent Modulation of Locomotion by GABAergic Spinal Sensory Neurons. *Curr Biol*. Elsevier; 2015; 25: 3035–3047. <https://doi.org/10.1016/j.cub.2015.09.070> PMID: [26752076](#)
40. Revenu C, Streichan S, Dona E, Lecaudey V, Hufnagel L, Gilmour D. Quantitative cell polarity imaging defines leader-to-follower transitions during collective migration and the key role of microtubule-dependent adherens junction formation. *Development*. 2014; 141: 1282–1291. <https://doi.org/10.1242/dev.101675> PMID: [24595289](#)
41. Fanning AS, Jameson BJ, Jesaitis LA, Anderson JM. The Tight Junction Protein ZO-1 Establishes a Link between the Transmembrane Protein Occludin and the Actin Cytoskeleton. *J Biol Chem*. 1998; 273: 29745–29753. <https://doi.org/10.1074/jbc.273.45.29745> PMID: [9792688](#)
42. Bartles JR, Wierda A, Zheng L. Identification and characterization of espin, an actin-binding protein localized to the F-actin-rich junctional plaques of Sertoli cell ectoplasmic specializations. *J Cell Sci*. 1996; 109: 1229–1239. PMID: [8799813](#)
43. Sekerková G, Loomis PA, Changyaleket B, Zheng L, Eytan R, Chen B, et al. Novel Espin Actin-bundling Proteins Are Localized to Purkinje Cell Dendritic Spines and Bind the SH3 Adapter Protein Insulin Receptor Substrate p53. *J Neurosci*. 2003; 23: 1310–1319. 23/4/1310 [pii] PMID: [12598619](#)
44. Bartles JR, Zheng L, Li A, Wierda A, Chen B. Small Espin: A third Actin-bundling Protein and Potential Forked Protein Ortholog in Brush Border Microvilli. *J Cell Biol*. 1998; 143: 107–119. <https://doi.org/10.1083/jcb.143.1.107> PMID: [9763424](#)
45. Loomis PA, Zheng L, Sekerková G, Changyaleket B, Mugnaini E, Bartles JR. Espin cross-links cause the elongation of microvillus-type parallel actin bundles in vivo. *J Cell Biol*. 2003; 163: 1045–1055. <https://doi.org/10.1083/jcb.200309093> PMID: [14657236](#)
46. Huang P, Xiong F, Megason SG, Schier AF. Attenuation of Notch and hedgehog signaling is required for fate specification in the spinal cord. *PLoS Genet*. 2012; 8(6): e1002762 <https://doi.org/10.1371/journal.pgen.1002762> PMID: [22685423](#)
47. Kindt KS, Finch G, Nicolson T. Kinocilia Mediate Mechanosensitivity in Developing Zebrafish Hair Cells. *Dev Cell*. 2012; 23: 329–341. <https://doi.org/10.1016/j.devcel.2012.05.022> PMID: [22898777](#)
48. Apodaca G. Role of Polarity Proteins in the Generation and Organization of Apical Surface Protrusions. *Cold Spring Harb Perspect Biol*. 2018; 10: a027813. <https://doi.org/10.1101/cshperspect.a027813> PMID: [28264821](#)
49. Crawley SW, Mooseker MS, Tyska MJ. Shaping the intestinal brush border. *J Cell Biol*. 2014; 207: 441–451. <https://doi.org/10.1083/jcb.201407015> PMID: [25422372](#)
50. Scita G, Confalonieri S, Lappalainen P, Suetsugu S. IRSp53: crossing the road of membrane and actin dynamics in the formation of membrane protrusions. *Trends Cell Biol*. 2008; 18: 52–60. <https://doi.org/10.1016/j.tcb.2007.12.002> PMID: [18215522](#)
51. Zhao H, Pykäläinen A, Lappalainen P. I-BAR domain proteins: linking actin and plasma membrane dynamics. *Curr Opin Cell Biol*. 2011; 23: 14–21. <https://doi.org/10.1016/j.ceb.2010.10.005> PMID: [21093245](#)
52. Pocha SM, Knust E. Complexities of Crumbs Function and Regulation in Tissue Morphogenesis. *Curr Biol*. 2013; 23: R289–R293. <https://doi.org/10.1016/j.cub.2013.03.001> PMID: [23578880](#)
53. Zou J, Wang X, Wei X. Crb Apical Polarity Proteins Maintain Zebrafish Retinal Cone Mosaics via Intercellular Binding of Their Extracellular Domains. *Dev Cell*. Elsevier Inc.; 2012; 22: 1261–1274. <https://doi.org/10.1016/j.devcel.2012.03.007> PMID: [22579223](#)
54. Pellikka M, Tanentzapf G, Pinto M, Smith C, McGlade CJ, Ready DF, et al. Crumbs, the Drosophila homologue of human CRB1/RP12, is essential for photoreceptor morphogenesis. *Nature*. 2002; 416: 143–149. <https://doi.org/10.1038/nature721> PMID: [11850625](#)
55. Salis P, Payre F, Valenti P, Bazellieres E, Le Bivic A, Mottola G. Crumbs, Moesin and Yurt regulate junctional stability and dynamics for a proper morphogenesis of the Drosophila pupal wing epithelium. *Sci Rep*. 2017; 7. <https://doi.org/10.1038/s41598-017-15272-1> PMID: [29196707](#)
56. DeRosier DJ, Tilney LG. F-Actin Bundles Are Derivatives of Microvilli: What Does This Tell Us about How Bundles Might Form? *J Cell Biol*. 2000; 148: 1–6. <https://doi.org/10.1083/jcb.148.1.1> PMID: [10629213](#)
57. Bartles JR. Parallel actin bundles and their multiple actin-bundling proteins. *Curr Opin Cell Biol*. 2000; 12: 72–78. [https://doi.org/10.1016/S0955-0674\(99\)00059-9](https://doi.org/10.1016/S0955-0674(99)00059-9) PMID: [10679353](#)
58. Sekerková G, Zheng L, Loomis PA, Changyaleket B, Whitlon DS, Mugnaini E, et al. Espins Are Multifunctional Actin Cytoskeletal Regulatory Proteins in the Microvilli of Chemosensory and Mechanosensory Cells. *J Neurosci*. 2004; 24: 5445–5456. <https://doi.org/10.1523/JNEUROSCI.1279-04.2004> PMID: [15190118](#)

59. Sekerková G, Zheng L, Loomis PA, Mugnaini E, Bartles JR. Espins and the actin cytoskeleton of hair cell stereocilia and sensory cell microvilli. *Cell Mol Life Sci.* 2006; 63: 2329–2341. <https://doi.org/10.1007/s00018-006-6148-x> PMID: 16909209
60. Zeisel A, Hochgerner H, Lönnerberg P, Johnsson A, Memic F, van der Zwan J, et al. Molecular Architecture of the Mouse Nervous System. *Cell.* 2018; 174: 999–1014.e22. <https://doi.org/10.1016/j.cell.2018.06.021> PMID: 30096314
61. Tilney LG, Connelly PS, Vranich KA, Shaw MK, Guild GM. Why Are Two Different Cross-linkers Necessary for Actin Bundle Formation In Vivo and What Does Each Cross-link Contribute? *J Cell Biol.* 1998; 143: 121–133. <https://doi.org/10.1083/jcb.143.1.121> PMID: 9763425
62. Zheng L, Sekerková G, Vranich K, Tilney LG, Mugnaini E, Bartles JR. The Deaf Jerker Mouse Has a Mutation in the Gene Encoding the Espin Actin-Bundling Proteins of Hair Cell Stereocilia and Fails to Accumulate Espins. *Cell.* 2000; 102: 377–385. [https://doi.org/10.1016/S0092-8674\(00\)00042-8](https://doi.org/10.1016/S0092-8674(00)00042-8) PMID: 10975527
63. Donaudy F, Zheng L, Ficarella R, Ballana E, Carella M, Melchionda S, et al. Espin gene (ESPN) mutations associated with autosomal dominant hearing loss cause defects in microvillar elongation or organization. *J Med Genet.* 2006; 43: 157–161. <https://doi.org/10.1136/jmg.2005.032086> PMID: 15930085
64. Pickles JO, Comis SD, Osborne MP. Cross-links between stereocilia in the guinea pig organ of Corti, and their possible relation to sensory transduction. *Hear Res.* 1984; 15: 103–112. [https://doi.org/10.1016/0378-5955\(84\)90041-8](https://doi.org/10.1016/0378-5955(84)90041-8) PMID: 6436216
65. Petit C, Richardson GP. Linking genes underlying deafness to hair-bundle development and function. *Nat Neurosci.* 2009; 12: 703–710. <https://doi.org/10.1038/nn.2330> PMID: 19471269
66. Kimmel CB, Ballard WW, Kimmel SR, Ullmann B, Schilling TF. Stages of Embryonic Development of the Zebrafish. *Dev Dyn an Off public.* 1995; 203: 253–310. <https://doi.org/10.1002/aja.1002030302> PMID: 8589427
67. Haeussler M, Schöning K, Eckert H, Eschstruth A, Mianné J, Renaud J-B, et al. Evaluation of off-target and on-target scoring algorithms and integration into the guide RNA selection tool CRISPOR. *Genome Biol.* 2016; 17:148. <https://doi.org/10.1186/s13059-016-1012-2> PMID: 27380939
68. Schindelin J, Arganda-Carreras I, Frise E, Kaynig V, Longair M, Pietzsch T, et al. Fiji—an Open Source platform for biological image analysis. *Nat Methods.* 2012; 9: 676–682. <https://doi.org/10.1038/nmeth.2019> PMID: 22743772
69. Wang Y, Wang F, Wang R, Zhao P, Xia Q. 2A self-cleaving peptide-based multi-gene expression system in the silkworm *Bombyx mori*. *Sci Rep. Nature Publishing Group;* 2015; 5:16273. <https://doi.org/10.1038/srep16273> PMID: 26537835
70. Erickson T, Nicolson T. Identification of sensory hair-cell transcripts by thiouracil-tagging in zebrafish. *BMC Genomics.* 2015; 16. <https://doi.org/10.1186/s12864-015-2072-5> PMID: 26494580
71. Donà E, Barry JD, Valentin G, Quirin C, Khmelinskii A, Kunze A, et al. Directional tissue migration through a self-generated chemokine gradient. *Nature.* 2013; 503: 285–289. <https://doi.org/10.1038/nature12635> PMID: 24067609
72. Auer TO, Xiao T, Bercier V, Gebhardt C, Durooure K, Concordet J-P, et al. Deletion of a kinesin I motor unmasks a mechanism of homeostatic branching control by neurotrophin-3. *Elife.* 2015; 4: 1–26. <https://doi.org/10.7554/eLife.05061> PMID: 26076409
73. Nelder J. A. Wedderburn RWM. Generalized linear models. *J R Stat Soc.* 1972; 135: 370–384.
74. Scott EK, Mason L, Arrenberg AB, Ziv L, Gosse NJ, Xiao T, et al. Targeting neural circuitry in zebrafish using GAL4 enhancer trapping. *Nat Methods.* 2007; 4: 323–326. <https://doi.org/10.1038/nmeth1033> PMID: 17369834

1                   **Myosin-1b interacts with UNC45A and controls intestinal epithelial**  
2                   **morphogenesis**

3  
4  
5   **Running Title: Myo1b and gut development**

6  
7  
8   Céline Revenu<sup>1</sup>, Marianna Parlato<sup>2\*</sup>, Marion Rosello<sup>1\*</sup>, Karine Duroure<sup>1</sup>, Rémi Duclaux-Loras<sup>2</sup>,  
9   Ophélie Nicolle<sup>4</sup>, Marie-Thérèse Prospéri<sup>3</sup>, Julie Stoufflet<sup>1</sup>, Juliette Vouigny<sup>1</sup>, Corinne  
10   Lebreton<sup>2</sup>, Priscilla Lépine<sup>3</sup>, Grégoire Michaux<sup>4</sup>, Nadine Cerf-Benssusan<sup>2</sup>, Evelyne Coudrier<sup>3</sup>,  
11   Filippo Del Bene<sup>1</sup>

12  
13  
14   <sup>1</sup>Institut Curie, PSL Research University, INSERM U934, CNRS UMR3215, 75248 Paris Cedex  
15   05, France.

16   <sup>2</sup>INSERM, UMR1163, Laboratory of Intestinal Immunity and Institut Imagine, 75015 Paris,  
17   France.

18   <sup>3</sup>Institut Curie, PSL Research University, CNRS UMR144, 75248 Paris Cedex 05, France.

19   <sup>4</sup>CNRS, UMR6290, Institut de Génétique et Développement de Rennes, F-35043 15 Rennes,  
20   France.

21   \* equal contributions

22   ‡Author for correspondence: [filippo.del-bene@inserm.fr](mailto:filippo.del-bene@inserm.fr)

23  
24  
25  
26  
27  
28  
29   **Keywords:**

30   Myosin 1b, intestine, zebrafish, UNC45A, villous atrophy

32 **Summary statement (max 30 words)**

33

34

35 Myosin-1b is important for intestinal epithelium folding during zebrafish development and  
36 participates in the villous atrophy clinical manifestation downstream UNC45A loss of function.

37

38

39 **Abstract**

40

41

42 Vesicle trafficking and the establishment of apico-basal polarity are essential processes in  
43 epithelium morphogenesis. Myosin-1b, an actin-motor able to bind membranes, regulates  
44 membrane shaping and vesicle trafficking. Here, we investigate Myosin-1b function in gut  
45 morphogenesis and congenital disorders using cell line and zebrafish larvae as well as patient  
46 biopsies. In a 3D Caco-2 cyst model, lumen formation is impaired in absence of Myosin-1b. In  
47 zebrafish, both Morpholino knock-down and genetic mutation of *myo1b* result in intestinal bulb  
48 epithelium folding defects associated with vesicle accumulation, reminiscent of a villous atrophy  
49 phenotype. We show that Myosin-1b interacts with the chaperone UNC45A, genetic deletion of  
50 which also results in gut folding defects in zebrafish. Loss of function mutations in *UNC45A* have  
51 been reported in complex hereditary syndromes, notably exhibiting intestinal disorders associated  
52 with villous atrophy. In *UNC45A*-depleted cells and in patient biopsies, Myosin-1b protein level  
53 is strikingly decreased. The appearance of Myosin-1b aggregates upon proteasome inhibition in  
54 cells points at a degradation mechanism of misfolded Myosin-1b in the absence of its chaperone.  
55 In conclusion, Myosin-1b plays an unexpected role in the development of the intestinal epithelium  
56 folds or villi downstream *UNC45A*, establishing its role in the gut defects reported in *UNC45A*  
57 patients.

58

59

## 60 **Introduction**

61

62 The establishment of apico-basal polarity and lumen formation are two fundamental steps during  
63 vertebrate intestinal epithelial morphogenesis (Chin et al., 2017). The actin cytoskeleton and the  
64 vectorial vesicle trafficking play a major role in the initiation and maintenance of this process,  
65 leading to a stable single layer of cells with distinct apical and basolateral domains (Lubarsky and  
66 Krasnow, 2003, Martin-Belmonte and Mostov, 2008). The apical membrane of the enterocyte is  
67 further organized in microvilli, plasma membrane protrusions, which are supported by bundles of  
68 parallel actin filaments and interacting proteins interconnected at the basis through a network of  
69 actin, spectrin, and myosins known as terminal web (Revenu et al., 2004). The interaction between  
70 neighboring polarized cells is further strengthened by the formation of cadherin-based adherens  
71 junctions and claudin-based tight junctions. A proper polarization of the intestinal epithelium is  
72 essential to achieve its main physiological roles, such as fluids and nutrient absorption and  
73 secretion. Indeed, defects in intestinal epithelial cell polarity and apical lumen formation result in  
74 early onset intestinal disorders, usually appearing in the first days of life (Kwon et al., 2020).  
75 Recently, loss of function (LOF) mutations in the chaperone *UNC45A* were reported in families  
76 presenting complex phenotypes including congenital diarrhea and several degrees of villous  
77 atrophy (Esteve et al., 2018). *UNC45A* belongs to the conserved UCS protein family (*UNC-*  
78 *45/CRO1/She4p*) of myosin co-chaperones.

79 Myosins 1 are single-headed actin motors targeted to membranes. Myosin1b (*Myo1b*) was detected  
80 in mouse enterocyte brush borders in a mass spectrometry analysis (Revenu et al., 2012). Studies  
81 in cell cultures reported that *Myo1b* associates with organelles and regulates membrane trafficking  
82 by controlling their morphology (Almeida et al., 2011). *Myo1b* can pull out membrane tubes along  
83 actin bundles immobilized on a solid substrate (Yamada et al., 2014) and it controls the formation  
84 of repulsive filopodia, the redistribution of actomyosin fibres driving cell repulsion (Prosperi et al.,  
85 2015) as well as the formation of axons in cultured neurons by controlling actin waves (Iuliano et  
86 al., 2018). Despite this progress in understanding *Myo1b* function *in vitro* and in cellular systems,  
87 its function in tissue biology, especially in the intestinal epithelium where it is expressed, remains  
88 unexplored. This work investigates this question in the context of gut epithelia development and  
89 morphogenesis.

90

91 Here we show that *Myo1b* is one of *UNC-45A* interactors, suggesting a role for myosin1b in the  
92 pathogenesis of *UNC45A* deficiency. *Myo1b* localizes at the apical brush border of intestinal  
93 epithelial cells in humans and loss of *Myo1b* in enterocyte like Caco-2 cells impairs epithelial

94 morphogenesis. In zebrafish, genetic inactivation of Myo1b affects intestinal bulb fold formation  
95 revealing its conserved function during normal intestinal epithelia development.

96

## 97 **Results**

98

### 99 **Myo1b is expressed in the gut epithelium and concentrates apically**

100 Myo1b gene expression and protein localization were analyzed in intestinal epithelial cells. Myo1b  
101 was detected by Western Blot and immunofluorescence in the human epithelial colorectal Caco-2  
102 cells (Fig. 1A-B). It accumulated apically in polarised Caco-2 cysts, as demonstrated by its  
103 colocalisation with the F-actin marker phalloidin demonstrating a localisation in actin-rich area,  
104 microvilli and/or the subjacent terminal web (Fig. 1B). As this model is adenocarcinoma cells, this  
105 expression and localization patterns could be the result of the tumoral state. To investigate Myo1b  
106 distribution *in vivo*, we looked for the homologue of *myo1b* in the zebrafish *Danio rerio*. There is  
107 one single *myo1b* gene with several splicing isoforms in the current zebrafish genome assembly.  
108 The corresponding Myo1b protein shares 80% identity with the *Homo sapiens* and *Mus musculus*  
109 homologues (Fig. 1C). In order to determine the expression pattern of *myo1b* during zebrafish  
110 development, we performed whole mount *in situ* hybridization labelling with specific antisense  
111 probes. *Myo1b* transcripts were unambiguously detected at 3dpf in the digestive tract of zebrafish  
112 (Fig. 1D) coinciding with the onset of gut morphogenesis (Ng et al., 2005, Wallace et al., 2005).  
113 *Myo1b* transcripts were also observed at 5dpf (Fig. 1D) when the intestine becomes functional and  
114 compartmentalised in bulb, mid and posterior intestines. At this stage, the transcripts were  
115 restricted to the intestinal bulb, the anterior part of the gut that forms large folds. Due to the lack  
116 of a zebrafish specific antibody, endogenous Myo1b sub-cellular localisation could not be assessed  
117 in zebrafish larvae. However, expressing eGFP-tagged Myo1b revealed apical accumulation of the  
118 protein, as previously seen in Caco-2 cysts by immunofluorescence (Fig. 1E).

119 Myo1b is thus expressed in human intestinal epithelial cells and in the developing zebrafish  
120 intestinal bulb epithelium, and it preferentially localises apically in the brush border.

121

### 122 **Myo1b loss of function Caco-2 cysts show normal apico-basal polarization but altered** 123 **luminal development.**

124 To address the function of Myo1b in enterocyte polarisation, *myo1b* was knocked-out using  
125 CRISPR/Cas9 in Caco-2 cells (*myo1b* KO, Fig. 1 A-B). The global apico-basal polarity of Caco-

126 2 cysts was not affected in *myo1b* KO cells compared to controls as shown by the correct apical  
127 concentration of F-actin, pERM and villin (Fig. 2 A-B).

128 Despite the absence of major polarisation defects, *myo1b* KO Caco-2 cells were more prone than  
129 controls to the formation of cysts with multiple lumen (Fig. 2). Indeed, *myo1b* KO cells showed a  
130 50% drop in the percentage of well-formed cysts with single central lumen compared to controls  
131 (Fig. 2C).

132

### 133 ***Myo1b* loss of function has no major impact on epithelial cell differentiation of zebrafish** 134 **intestinal bulb**

135 To investigate the implication of Myo1b in intestinal epithelium morphogenesis *in vivo*, we turned  
136 to zebrafish as a good model for gut development. The zebrafish intestinal epithelium differentiates  
137 from 3 days post fertilisation (dpf) where it is essentially a flat monolayered tube. At 5dpf,  
138 epithelial folds are present, especially in the anterior most part of the gut, the intestinal bulb  
139 (Wallace et al., 2005). These folds are equivalent to the mammalian villi, and although no crypts  
140 are present in zebrafish, the region between folds will have a crypt-like role (Crosnier et al., 2005).  
141 First, we designed a splice blocking Morpholino (*Myo1b*-MO) that is efficiently preventing proper  
142 splicing of *myo1b* as determined by RT-PCR (supplemental Fig. 1A-B). At the concentration used,  
143 *Myo1b* MO displayed no overt phenotype, despite occasionally a slight heart oedema  
144 (supplemental Fig. 1C). To extend these results with a genetic loss of function model, we also  
145 generated a mutated allele at the *myo1b* locus using the CRISPR/Cas9 system, resulting in the  
146 insertion of a single base at the beginning of the open reading frame, as confirmed by sequencing  
147 (supplemental Fig. 1D). This leads to a premature stop codon and to the lack of detection of the  
148 protein by Western blot in gut lysates from adult homozygote mutants (*myo1b*<sup>-/-</sup>, supplemental  
149 Fig. 1E). As *myo1b* mRNA is maternally provided (supplemental Fig. 1F), maternal contribution  
150 was suppressed by crossing *myo1b*<sup>-/-</sup> mothers. As for the MO injections, the resulting maternal-  
151 zygotic homozygous mutant larvae displayed no overt phenotype (supplemental Fig. 1C). In cross-  
152 sections (Fig. 3A), the intestinal bulbs of *Myo1b* MO and *myo1b*<sup>-/-</sup> larvae appeared smaller  
153 compared to controls. A significant reduction of the number of cells per cross-section was observed  
154 for both *Myo1b* MO and *myo1b*<sup>-/-</sup> intestinal bulbs at 3 and 5dpf compared to their respective  
155 controls (Fig. 3B). A reduction in the total cell number in the intestinal bulb could be the  
156 consequence of increased apoptosis or reduced cell proliferation. No significant difference with  
157 controls in the proportion of proliferative cells could be detected at 3 and 5dpf (supplemental Fig.  
158 2A). A slight increase in the proportion of apoptotic cells could be detected at 5dpf but not at 3dpf  
159 (supplemental Fig. 2B). This later increase in apoptosis can however not account for the reduced

160 cell number per section reported from 3dpf on and could more be a readout of increased cellular  
161 stress level upon prolonged absence of Myo1b, as reported after the KO of other Myosins 1 in  
162 mouse and drosophila (Hegan et al., 2007, Tyska et al., 2005). Using specific markers for secreting  
163 and absorptive cell lineages, defects in enterocyte differentiation could also be excluded  
164 (supplemental Fig. 2 C-D). Finally, the microvilli marker Villin appeared properly localised  
165 apically, lining the lumen together with F-actin (Fig. 3C), suggesting that apical polarity is not  
166 affected in Myo1b MO and *myo1b* *-/-* intestinal epithelium *in vivo*, as already shown in 3D Caco-  
167 2 cell cultures (Fig. 2B).

168

### 169 ***Myo1b* loss of function zebrafish display MVID-like features in the intestinal bulb epithelium**

170 To analyse in 3D intestinal bulb epithelium morphogenesis in zebrafish, we used the BAC line  
171 *cldn15la:cldn15la-GFP* that specifically labels the gut epithelium (Alvers et al., 2014). Both MO  
172 and KO intestinal bulbs revealed single continuous lumen suggesting that early steps of lumen  
173 fusion events were not affected (Alvers et al., 2014, Horne-Badovinac et al., 2001). However, in  
174 Myo1b MO larvae, the intestinal bulb epithelium appeared most of the time flat at 5dpf, not  
175 developing the expected folds observed in controls (Fig. 3D). Consistently with this phenotype, in  
176 the KO model, we detected a significant reduction in fold length in KO versus control samples  
177 (Fig. 3E). In an attempt to understand this milder phenotype in the mutant compared to the MO  
178 condition, we analysed potential compensation mechanisms by other myosins 1 performing RT-  
179 QPCR. On the 4 myosins 1 tested (*myo1ca*, *myo1cb*, *myo1d* and *myo1eb*), *myo1eb* showed a  
180 reproducible increase of on average 60% of the WT expression level in the mutant using EF1a  
181 (supplemental Fig. 1G) and RPL13a (not shown) as reference genes. *Myo1eb* is broadly expressed  
182 in early developmental stages but has restricted expression patterns after 2dpf (mostly branchial  
183 arches and pharynx) (Thisse and Thisse, 2004). A partial compensation of the loss of *myo1b* by an  
184 upregulation of *myo1eb* could thus potentially explain the subtler and more restricted phenotype of  
185 the mutant larvae compared to the knock-downs.

186 To further characterize the architecture of Myo1b-deficient intestinal bulb epithelium, a  
187 histological analysis by transmission electron microscopy (TEM) was performed on 5dpf larvae.  
188 It confirmed the affected folding of the intestinal bulb epithelium in MO and KO samples, and the  
189 preserved apico-basal polarity of enterocytes (Fig. 4A-B). Quantifying microvilli length and  
190 density did not reveal any significant defect resulting from Myo1b downregulation or absence,  
191 although packing looked less regular (Fig. 4C). In contrast, a darker sub-apical band was visible  
192 in Myo1b affected samples compared to controls, likely corresponding to modifications of the  
193 terminal web (Fig. 4C). Moreover, this ultrastructural analysis showed an important accumulation

194 of vesicles in MO and KO samples compared to controls (Fig. 4B insets) suggesting defects in  
195 membrane trafficking. These TEM observations (epithelium folding impaired, modifications of  
196 the apical pole ultrastructure and trafficking defects) overall indicate that myo1b-deficient  
197 enterocytes display some microvillus inclusion disease-like features.

198

### 199 **Myo1b is destabilized when Unc45A is affected**

200 Loss of function (LOF) mutations in the chaperone *UNC45 homolog A* have recently been  
201 associated with rare human genetic syndromes notably presenting intestinal disorders, including  
202 chronic diarrhea and villous atrophy of variable penetrance (Esteve et al., 2018). A pull-down  
203 assay performed to detect potential Myo1b partners identified UNC45A as the most abundant  
204 protein interacting with Myo1b in a mouse neuronal cell model (Supplementary table 1). This  
205 interaction was confirmed by the reverse experiment using UNC45A as bait in the colorectal Caco-  
206 2 cells (Duclaux-Loras *et al.*, submitted).

207 Considering this interaction, we investigated the impact of *UNC45A* depletion on Myo1b  
208 expression. *UNC45A* depleted Caco-2 cells showed reduced Myo1b levels by  
209 immunofluorescence (Fig. 5A). To detect aggregation-prone proteins normally sent to  
210 degradation, the proteasome machinery was blocked using the proteasome inhibitor MG132.  
211 MG132 induced the appearance of protein aggregates both in control and UNC45A KO Caco-2  
212 cells (Fig. 5B). Myo1b staining partly co-localised with the aggresomes in the UNC45A KO  
213 condition (Fig. 5B). This result suggests that Myosin1b proper folding requires the chaperone  
214 UNC45A.

215 We finally looked at Myo1b expression in duodenal biopsies from control and *UNC45A* loss of  
216 function patients. In control biopsies, the microvilli marker Villin was expressed apically all along  
217 the epithelium and Myo1b was detected apically at the base of the villi and in crypts, partially  
218 colocalising with Villin (Fig. 5C). In *UNC45A* patients, Villin was still localised apically whereas  
219 Myo1b was barely detectable (Fig. 5D). In conclusion, Myo1b protein level is decreased in  
220 *UNC45A*-depleted cells and in duodenal biopsies from an *UNC45A*-mutated patient.

221

### 222 **Discussion**

223

224 This work identifies Myo1b, an actin motor, as an unexpected player in the regulation of the  
225 morphogenesis of the intestinal epithelium during gut development. In zebrafish, we report defects  
226 in epithelial folding and villous atrophy when Myo1b is impaired. This phenotype is similar to the

227 ones reported in zebrafish and in human patients with loss of function mutations in *UNC45A*. We  
228 show that Myo1b interacts with UNC45A and that Myo1b is destabilised in absence of UNC45A.

229

230 For this study, we analysed both *myo1b* mutant- and Morpholino-induced phenotypes. It is now  
231 well established that Morpholino knock-downs often result in more severe overt phenotypes than  
232 the corresponding knock-outs, at least partially due to the induction of genetic compensation  
233 mechanisms in the mutants (Kok et al., 2015, Rossi et al., 2015). In the *myo1b* null case, we  
234 observed a more subtle outcome than the Morpholino, which could be due to compensation  
235 mechanisms (El-Brolosy and Stainier, 2017) as supported by the RT-QPCR of *myo1eb*. Here, the  
236 intestinal bulb phenotypes observed with both approaches converged on reduced cell number of  
237 intestinal bulb sections and impaired epithelial folding.

238

239 Up to now, myosins 1a, c, d and e had been identified in intestinal brush borders constituting the  
240 apical pole of differentiated enterocytes (Benesh et al., 2010). Several class I myosins have been  
241 implicated in the maintenance of intestinal epithelial differentiated state. Myo1a, which is  
242 associated with the highly organised actin network of differentiated enterocytes in mammals  
243 (Revenu et al., 2012, Tyska et al., 2005), but seems to lack in the zebrafish and *Drosophila*  
244 genomes, is important for enterocyte polarity and participates in the structure and composition of  
245 the brush border (Mazzolini et al., 2012, Tyska et al., 2005). The phenotype of the *myo1a* KO mice  
246 is however mild, with reports of clear compensations by other class I myosins (Benesh et al., 2010,  
247 Tyska et al., 2005). Likewise, two of the known class I myosins in *Drosophila* are also localised  
248 in the apical pole of differentiated enterocytes and *Drosophila* Myo61F is necessary for the  
249 stability of enterocyte apical organisation (Hegan et al., 2007). We report the expression of another  
250 myosin 1 in the gut epithelium, Myo1b, and its apical localisation in enterocytes of Human  
251 biopsies, of zebrafish intestinal bulb and in Caco-2 cells. Myo1b localisation in microvilli had  
252 previously been reported in kidney epithelial cells (Komaba and Coluccio, 2015). In Human  
253 biopsies, Myo1b is expressed at the base of the villi and in crypts suggesting a specific role in the  
254 proliferative compartment and not in enterocyte differentiated state. We did not detect global  
255 polarisation defects at the cell level or impaired differentiation in absence of Myo1b, whereas both  
256 the Caco-2 3D model and the zebrafish model demonstrate morphogenetic defects at the tissue  
257 level, respectively single lumen formation and folding.

258

259 Proliferation, apoptosis or differentiation defects do not account for the reduced cell number  
260 observed on transverse sections of the intestinal bulb from 3dpf. As the sections give a 2D

261 overview of a 3D organ, this reduced cell number is thus likely the readout of the different  
262 organisation in space of the epithelium. A specificity of the zebrafish intestinal bulb is the early  
263 folding of the epithelium that remains pronounced to adulthood (Ng et al., 2005). The reduced  
264 folding and villi formation in the mutants and morphants are clear signs of a different architecture  
265 of the tissue. Also the mechanisms underlying intestinal epithelium folding are not yet fully  
266 understood, the impact of tension and forces at the cell and tissue level driving compression, cell  
267 intercalation and invagination through apical constriction have been investigated in other tissues  
268 during development (Mammoto and Ingber, 2010). Myosins are central in the control of actin  
269 cytoskeleton dynamics and in force generation (Reymann et al., 2012). In particular, Myo1b  
270 deforms membranes and participates in organelle formation and trafficking (Almeida et al., 2011,  
271 Coudrier and Almeida, 2011). It also remodels the actin cytoskeleton (Iuliano et al., 2018, Pernier  
272 et al., 2019, Prosperi et al., 2015). Its roles in membrane traffic and in the dynamic organisation of  
273 actin structures make it a plausible actor in the morphogenesis of the gut epithelium. The electron  
274 microscopy data show a strong accumulation of intra-cellular vesicles in Myo1b mutant and  
275 Morpholino tissues suggesting impaired trafficking, in agreement with its role in the formation of  
276 post Golgi carriers and protein transport at the level of multivesicular endosomes (Almeida et al.,  
277 2011, Salas-Cortes et al., 2005). Electron microscopy also reveals modifications of the terminal  
278 web area, the apical actin belt linking adherens junctions in the epithelium, in agreement with its  
279 role in actin dynamics. Myosin 1b function on actin dynamics and consequently on membrane  
280 remodelling and membrane trafficking must impact cell and tissue mechanics (Buske et al., 2012),  
281 and this way contributes to impaired intestinal epithelial folding in the absence or down-regulation  
282 of *myo1b*. Myo1b restricted localisation at the base of the villi in human biopsies may indicate a  
283 specific mechanical role in crypts morphogenesis.

284

285 Villous atrophy is a phenotype associated with various intestinal disorders including some rare  
286 hereditary syndroms presenting congenital diarrhea like microvillous inclusion disease (MVID).  
287 Mutations in MyosinVb are the main cause of MVID (Muller et al., 2008) and have notably been  
288 associated with defective trafficking. A zebrafish mutant of *myoVb* develops a flat intestinal  
289 epithelium (Sidhaye et al., 2016). Recently, loss of function (LOF) mutations in the chaperone  
290 *UNC45A* were reported in families presenting complex phenotypes including congenital diarrhea  
291 and several degrees of villous atrophy (Esteve et al., 2018). A zebrafish mutant for *UNC45A* also  
292 exhibit loss of intestinal epithelium folding (Esteve et al., 2018). *UNC45A* is a chaperone  
293 participating in the conformational maturation of, among others, some Myosins (Barral et al.,  
294 2002, Lee et al., 2014, Lehtimaki et al., 2017). Our results in human cell lines and in patient

295 samples demonstrate a strong reduction in the protein level of Myo1b in absence of a functional  
296 UNC45A variant, probably due to the degradation of misfolded and destabilised Myo1b. Myo1b  
297 proper conformational folding would thus require the chaperone UNC45A. The intestinal  
298 phenotypes associated with LOF mutations in *UNC45A* could hence partly be the consequence of  
299 the reduced protein level of Myo1b. In conclusion, Myo1b contributes to gut morphogenesis and  
300 appears as a potential player in the complex intestinal phenotype of the UNC45A LOF syndrome.

301

302

### 303 **Materials and Methods**

304

305 **CRISPR- Cas9 genome editing of MYO1B in Caco2 cells and 3D culture.** The lentiCRISPRv2  
306 plasmid was a gift from F. Zhang (Massachusetts Institute of Technology, Boston, MA; plasmid  
307 no. 98290, Addgene). The single-guide RNAs (sgRNAs) were designed using the CRISPR Design  
308 Tool (Massachusetts Institute of Technology) and cloned into the BsmBI site. sgRNA sequences:  
309 for-5- CACCGATCCCTACGAGATCAAGATA -3, rev-5- AAAC TA TCT TGA TCT CGT  
310 AGG GAT C -3. Following production of lentiviral particles, the lentiCRISPRv2 plasmids were  
311 transduced in Caco2 cells. Positively transduced cells were selected by puromycin (10µg/ml).  
312 For 3D culture, CaCo2 cells were resuspended at a concentration of 10<sup>4</sup> cells/mL in DMEM  
313 (Gibco) with 20% FCS containing 4% Matrigel (BD Biosciences) and 2.5 10<sup>4</sup> cells/well were  
314 plated in 8-wells chamber slide IBIDI (Biovalley), previously precoated with 100 µL of Matrigel.  
315 Cells were grown for 5 days to obtain cysts. To detect aggregation-prone proteins, the proteasome  
316 inhibitor MG132 (Sigma-Aldrich) was added overnight in the culture medium at a concentration  
317 of 10 µM.

318

319 **Western blot.** Caco-2 cells were lysed in RIPA buffer (Sigma) supplemented with 1X proteinase  
320 inhibitor cocktail mix (Roche, Sigma). Adult zebrafish guts were dissected on ice and mechanically  
321 lysed in 200µL lysis buffer (10 mM HEPES + 300 mM KCl + 5 mM MgCl<sub>2</sub> + 0,45% triton X100  
322 + 0,05% Tween20, pH7) with 10 mM ATP and Complete protease inhibitor (Roche). 40 µg of  
323 extracts in Laemmli buffer were loaded on a 4-12% polyacrylamide gradient concentration gel  
324 (ThermoFisher). Primary antibodies used were mouse anti-tubulin (1:12000, Sigma), rabbit anti-  
325 ratMyo1b (1,8 µg/µL, 1:500, (Salas-Cortes et al., 2005), anti GAPDH (1:1000, Cell Signaling).

326

327 **Phylogenetic analysis.** The Myo1b and Myo1a homologues in *Danio rerio*, *Homo sapiens*, *Mus*  
328 *Musculus*, *Gallus gallus* and *Drosophila melanogaster* were obtained from NCBI HomoloGene.  
329 Protein sequences were aligned and a phylogenetic tree was assembled using the online ‘One  
330 Click’ mode at Phylogeny.fr (Dereeper et al., 2008).

331

332 **Molecular Cloning.** The *βactinhsp70:KalT4;cmcl2:eGFP* construct was generated by combining  
333 four plasmids using the Multisite Gateway system (Invitrogen): p5E-*βactinhsp70*, pME-KalT4,  
334 p3E-polyA and pDEST-cmcl2:eGFP containing Tol2 sites (Kwan et al., 2007). The *βactin*  
335 promoter was cloned into the pCR-bluntII-TOPO vector (Invitrogen) and then inserted in the p5E-  
336 MCS using KpnI and XhoI restriction sites. The 3’ 638bp of the zebrafish *hsp70* promoter (Dalgin  
337 et al., 2011) was inserted into this p5E-*βactin* vector linearized with XhoI using the Gibson  
338 Assembly Cloning Kit (New England Biolabs). The optimised Gal4 KalT4 (Distel et al., 2009)  
339 was amplified and inserted in a pDONR221 using the Multisite Gateway system (Invitrogen).

340 To generate the *14UAS:ubc-eGFP-Myo1b* vector, eGFP-Rat Myo1b cDNA was amplified from a  
341 previously published plasmid (Prosperi et al., 2015). It was inserted into the pT1UciMP plasmid  
342 containing 14xUAS-ubc and Tol1 sites (Horstick et al., 2015) linearized with XhoI using the  
343 Gibson Assembly Cloning Kit (New England Biolabs).

344

345 **Zebrafish (*Danio rerio*) husbandry.** Wild-type Tupfel long fin zebrafish strains were used and  
346 raised according to standard protocols. Stable transgenic lines were generated by injection of the  
347 plasmids with *tol2* or *tol1 transposase* mRNA at 25ng/μL in one-cell stage zebrafish embryos. The  
348 transgenic BAC line claudin15-like-a fused to GFP (*cldn15la:cldn15la-GFP*) was kindly provided  
349 by Michel Bagnat (Alvers et al., 2014).

350 For live-imaging, larvae were anaesthetised in 0.02% MS-222 and immobilised in 1% low melting  
351 point agarose. Imaging was performed on a Zeiss LSM 780 confocal. All animal procedures were  
352 performed in accordance with French and European Union animal welfare guidelines.

353

354 **In situ hybridization.** *In situ* hybridizations (ISH) were performed on larvae treated with 1-  
355 phenyl-2-thiourea (PTU, Sigma-Aldrich) and fixed in freshly made 4% paraformaldehyde (PFA)  
356 2-4h at RT and stored in 100% methanol at -20°C. After rehydration, larvae were treated with  
357 proteinase K (20 μg/ml; Roche diagnostics) at RT for 1h (3dpf) or 2h (5dpf) and fixed again in 4%  
358 PFA at RT for 20min. Digoxigenin-labelled antisense and sense RNA probes were synthesized by  
359 *in vitro* transcription using DIG-labelled UTP according to the manufacturer's instructions (DIG  
360 RNA labelling kit, Roche). Primers used were as follow: *myo1b* sense: CAA TAT GAT AGG

361 GGT AGG GGA CAT G ; antisense: TGG TTT GAA CTC AAT ATT TCC CAG C. Anti-DIG  
362 antibody conjugated to alkaline phosphatase allowed detection of hybridized riboprobes according  
363 to the manufacturer's instructions (Roche).

364

365 ***Myo1b* zebrafish mutant generation with CRISPR/Cas9.** The sgRNA sequence (sgB:  
366 CTTCTGACAAGGGCTCTAGG) was cloned into the BsaI digested pDR274 vector (Addgene).  
367 The sgRNAs were synthesized by *in vitro* transcription (Megascript T7 transcription kit, Ambion).  
368 After transcription, sgRNAs were purified using the RNAeasy Mini Kit (Quiagen). For injections  
369 at one cell stage, the synthesized sgB was injected at 300ng/ $\mu$ L after 5-minute incubation at RT  
370 with Cas9 protein (NEB) at 2 $\mu$ M final in 20mM Hepes-NaOH pH 7.5, 150mM KCl (Albadri et al.,  
371 2017). Injected embryos were grown to adulthood and crossed with wild-types to identify founders.  
372 Pools of 20 embryos per clutch were lysed in NaOH 50mM at 95°C for at least 30min. PCR was  
373 performed on lysates to amplify the genomic region targeted by the sgB with primers forward  
374 5'GGGTGTTGTTTCAGCGATGGA and reverse 5'ATAGATCTCATTGTGATCGA using  
375 Phusion High-Fidelity DNA polymerase (Thermo Scientific). The amplicons were cloned in pCR-  
376 bluntII-TOPO vector (Zero Blunt Topo PCR cloning kit, Invitrogen) and sequenced (GATC  
377 Biotech) to identify indels and the corresponding founder fish. Sequences were analysed using  
378 Geneious. After selection of the founder, genotyping of the line was performed by PCR on fin clips  
379 with primers 5'AGATGAATGCAAGCAAGCCATT and  
380 5'ATACGATCTGATTGTGATCGAATCGCT. The resulting product was digested with  
381 restriction enzyme FspBI, the site of which is lost in the mutant, resulting in 2 fragments (208 and  
382 66bp) for the WT allele and only one (275bp) for the mutated allele.

383

384 **Morpholino oligonucleotide design and injections.** *Myo1b* splice blocking morpholino was  
385 designed to target the splice donor site downstream of exon 22 (*Myo1b*-MO, 5'-  
386 ATGAGAAACTGTGTTTCATTACCTGG). Experiments were performed in parallel with a  
387 standard control-MO (5'-CCTCCTACCTCAGTTACAATTTATA) with no target in the zebrafish  
388 genome. Morpholinos (Gene Tools) diluted at 1mM in water, were injected in 1-cell stage embryos  
389 at a final concentration of 0.2mM. To validate *Myo1b*-MO knock-down, RT-PCR was performed  
390 on 3dpf larvae. Total RNA of 50 larvae was prepared with TRIzol and TURBO DNase-free  
391 reagents (Invitrogen). mRNA (1 $\mu$ g) was retro-transcribed using oligo(dT) primers and the  
392 SuperScript III First-Strand Synthesis System (Invitrogen). To amplify the region targeted by the  
393 MO (supplemental Fig. 1A), PCR was performed on the cDNA with two different forward primers

394 (primer 1 in exon21 was: 5'GGCTGCGATATTCTTGCCTCC, primer 2 at the edge of exon22 and  
395 the targeted intron was: 5'TCTTTCATTCGTGGATGGAAGGCC) and the reverse primer  
396 5'AACCCAGGTAATGAACACAGTTTCTAT. PCR products were run on a gel (supplemental  
397 Fig. 2B) and bands were gel purified (Macherey-Nagel), inserted in a pCR-bluntII-Topo vector  
398 (Invitrogen) and sent for sequencing (GATC) to assess intron retention.

399

400 **Quantitative RT-PCR.** For each experiment, total RNA was prepared from 3 pools of 50 embryos  
401 per phenotype with TRIzol reagent and the TURBO DNA-free kit (ThermoFisher Scientific). RNA  
402 (1µg) was retro-transcribed using random primers and the SuperScript III First-Strand Synthesis  
403 system (ThermoFisher Scientific). For RT-QPCR, the SYBR Green PCR Master Mix  
404 (ThermoFisher Scientific) was used according to the manufacturer's instructions and PCR were  
405 performed on an ABI PRISM 7900HT instrument. Experimental triplicates of each sample were  
406 averaged and the relative expression level quantified with the log<sub>2</sub>ΔCT method using EF1a and  
407 RPL13a reference genes. Shown are values normalised on the wild-type samples.

408

409 **Immunohistochemistry.** Caco-2 cells and cysts were fixed in 4% PFA 30min at 37°C and washed  
410 with PBS. They were permeabilised in PBS, 0.2%TX100, 1%BSA for 5min at RT and blocked in  
411 PBS 3%BSA for 1-2 hours at RT. Primary antibodies used were rabbit anti-ratMyo1b (1,8 µg/µL,  
412 1:100, (Salas-Cortes et al., 2005), mouse anti-Villin clone ID2C3 (1:300, (Robine et al., 1985),  
413 rabbit anti-pERM (1:100, abcam ab47293). After washes, they were incubated with Alexa Fluor  
414 488 secondary antibody (Molecular probes), phalloidin-Alexa Fluor 568 and Dapi. To assess  
415 protein aggregation, the Proteostat Aggresome Detection Kit (ENZO, ENZ-51035) was used.  
416 Briefly, after the primary antibody rabbit anti-ratMyo1b, cells were incubated for 30min at RT with  
417 goat anti-rabbit-Alexa Fluor 635 antibody (1:400, Molecular probes), Proteostat 1:400 and Hoechst  
418 1:800 (ENZO) in PBS 3%BSA.

419 Zebrafish larvae were fixed for 2h at room temperature in 4% PFA and incubated in 30%  
420 sucrose/0.1% PBST overnight at 4°C. They were then frozen in Tissue-Tek OCT (Sakura) at -80°C  
421 and sectioned using a Cryostat (Leica). Zebrafish larvae sections were incubated in blocking buffer  
422 (10% serum in PBST, PBS 0.1%Tween20) and with mouse anti-Villin clone ID2C3 (1:300), mouse  
423 2F11 antibody (1:100, Abcam ab71286) or 4E8 (1:100, Abcam ab73643) overnight at 4°C. After  
424 washes with PBST, they were incubated with Alexa Fluor 488 secondary antibody (Molecular  
425 probes), phalloidin-Alexa Fluor 568 and Dapi. To assess apoptosis, TUNEL assay was performed  
426 with reaction solutions from ApopTag Red In situ Apoptosis detection kit (Millipore) according to

427 the manufacturer recommendations. To assess proliferation, larvae were injected in the yolk with  
428 10mM 5-ethynyl-2'-deoxyuridine (EDU) in 1% DMSO and incubated in 100 $\mu$ M EdU, 0.4%DMSO  
429 for 20 hours after injection. Animals were fixed at indicated time and processed according to the  
430 Click-iT EdU Imaging Kit (Invitrogen).

431 Paraffin embedded sections of intestinal tissues from a UNC45A deficient patient and biopsies  
432 from controls were obtained for diagnosis or therapeutic purposes. Duodenal biopsies were  
433 routinely fixed in 4% buffered formalin for 24 hours and paraffin-embedded. Sections were heated  
434 for 1hr at 65°C and paraffin was removed by two 5-min washes in xylene. Sections were then  
435 hydrated with ethanol solutions of decreasing concentrations. Unmasking of the epitopes was  
436 performed at 100°C for 20 min in Citrate-based Antigen Unmasking Solution (Vector  
437 Laboratories). Sections were incubated for 30 min at room temperature in blocking buffer (3%  
438 BSA in PBS) and then overnight at 4°C with anti Myo1b antibody (1:200, Novus Biologicals  
439 NBP1-87739) in blocking. After washes with PBST, sections were incubated with goat anti rabbit  
440 Alexa Fluor 488 antibody (Molecular probes), phalloidin-Alexa Fluor 568 and Dapi for 2hrs at  
441 RT.

442 After extensive washes and mounting in Vectashield (Vector Lab), all stainings were imaged on a  
443 LSM780 confocal microscope (Zeiss). Images were processed and numbers of cells quantified  
444 using ImageJ.

445

446 **TEM analysis on zebrafish larvae.** 5dpf larvae were collected and stored at 4°C in Trump's  
447 fixative. Enhanced chemical fixation was performed in a mix of 4% PFA with 2.5% glutaraldehyde  
448 in 0.1 mol/L cacodylate buffer overnight at 4°C. A 1.5-hour incubation in 1% OsO<sub>4</sub> was followed  
449 by a 1.5-hour incubation with 2% uranyl acetate at ambient temperature. Larvae were then  
450 dehydrated through graded ethanol solutions, cleared in acetone, infiltrated, and embedded in  
451 Epon-Araldite mix (EMS hard formula). We used adhesive frames (11560294 GENE-FRAME 65  
452  $\mu$ L; Thermo Fisher Scientific) for flat-embedding, as previously described (Kolotuev et al., 2012),  
453 to gain better anteroposterior orientation and sectioning. Ultrathin sections were cut on an  
454 ultramicrotome (UC7; Leica Microsystems) and collected on formvar-coated slot grids (FCF2010-  
455 CU, EMS). Each larva was sectioned transversally in five different places in intestinal bulb with  
456  $\geq 20 \mu$ m between each grid to examine the sample over a large region. Each grid contained at least  
457 4-6 consecutive sections of 70 nm. TEM grids were observed using a JEM-1400 transmission  
458 electron microscope (JEOL) operated at 120 kV, equipped with a Gatan Orius SC1000 camera  
459 (Gatan) and piloted by the Digital Micrograph program. Microvilli length and density were

460 quantified using Fiji on TEM pictures of at least 50 MV from 25 enterocytes of 3 larvae per  
461 condition.

462

463 **Pull Down assay.**  $10^6$  N1E115 cells were transfected with pEGFP Myo1b (Salas-Cortes et al.,  
464 2005) and lysed in TRIS 150mM, NaCl 150mM, EDTA 1mM, EGTA 1mM, ATP 10 mM, 10%  
465 glycerol, 1mM DTT, 0,5% triton and protease inhibitor 24 hours after transfection. The lysate was  
466 then incubated with 15 ml of GFP trap Beads (Chromotek) overnight. After washing the beads  
467 were resuspended in water and treated for mass spectrometry analysis.

468

469 **Statistical analysis.** The numbers of cells reported are coming from manual counting. No  
470 sample was excluded from the analysis, except for the total cell number per section where  
471 we made sure to analyse samples displaying single cell layers through the whole gut cross-  
472 sections and not the side of some villi. The sample size (n=) is defined as the number of larvae  
473 analysed (one section per larva). For statistical analysis, we applied the non-parametric  
474 Wilcoxon-Mann Whitney test.

475

476

477

#### 478 **Acknowledgements**

479 The authors thank the Del Bene team for fruitful suggestions and discussions and members  
480 of the Institut Curie zebrafish facility. The authors acknowledge all members from the PICT-  
481 IBiSA Lhomond Imaging Platform (UMR144) and the Cell and Tissue Imaging Platform of the  
482 Genetics and Developmental Biology Department (UMR3215/U934) of Institut  
483 Curie, member of France-Bioimaging (ANR-10-INSB-04), for help with light microscopy and  
484 the electron microscopy unit of the MRic facility (Rennes, France).

485

#### 486 **Competing Interests**

487 No competing interests declared.

488

#### 489 **Contributions**

490 MP generated the KO cells. JS and CR generated the *Myo1b* null allele in zebrafish. MR  
491 generated the stable zebrafish transgenic lines. CR performed immunofluorescence  
492 stainings. KD, JV and MR performed ISH. MP, MR, CR and RDL performed WB. KD and CR did

493 the RT-qPCR analysis. JS and JV genotyped the mutants. ON and GM did the TEM and analysis.  
494 MP, RDL and CL prepared cell culture samples. PL generated preliminary data with  
495 Morpholinos. MTP and EC performed the pull-down assay. CR analysed the results and wrote  
496 the paper. FDB, EC and NCB supervised the work. MR, MP, GM, EC and FDB edited the  
497 manuscript.

498

## 499 **Funding**

500 This work has been supported by Institut Curie, CNRS, INSERM and grants from the ANR  
501 (ANR-14-CE11-0005-03), ANR/e-RARE (ANR-12-RARE-0003-03) and the ARC foundation  
502 (grant n°SFI2012205571). E.C. group belongs to the CNRS consortium CellTiss and to the  
503 Laboratoire d'Excellence (LABEX) CelTisPhyBio 11-LBX-0038. FDB group is part of the  
504 LABEX DEEP 11-LABX-0044, and of the École des Neurosciences de Paris Ile-de-France  
505 network. CR was supported by a EU H2020 Marie Skłodowska-Curie Action fellowship  
506 (H2020-MSCA-IF-2014 #661527). MR was supported by the Fondation pour la Recherche  
507 Médicale (FRM grant number ECO20170637481).

508

509

## 510 **REFERENCES**

511

- 512 Albadri, S., Del Bene, F. & Revenu, C. 2017. Genome editing using CRISPR/Cas9-based knock-  
513 in approaches in zebrafish. *Methods*, 121-122, 77-85.
- 514 Almeida, C. G., Yamada, A., Tenza, D., Louvard, D., Raposo, G. & Coudrier, E. 2011. Myosin 1b  
515 promotes the formation of post-Golgi carriers by regulating actin assembly and  
516 membrane remodelling at the trans-Golgi network. *Nat Cell Biol*, 13, 779-89.
- 517 Alvers, A. L., Ryan, S., Scherz, P. J., Huisken, J. & Bagnat, M. 2014. Single continuous lumen  
518 formation in the zebrafish gut is mediated by smoothed-dependent tissue  
519 remodeling. *Development*, 141, 1110-9.
- 520 Barral, J. M., Hutagalung, A. H., Brinker, A., Hartl, F. U. & Epstein, H. F. 2002. Role of the myosin  
521 assembly protein UNC-45 as a molecular chaperone for myosin. *Science*, 295, 669-71.
- 522 Benesh, A. E., Nambiar, R., Mcconnell, R. E., Mao, S., Tabb, D. L. & Tyska, M. J. 2010. Differential  
523 localization and dynamics of class I myosins in the enterocyte microvillus. *Mol Biol*  
524 *Cell*, 21, 970-8.
- 525 Buske, P., Przybilla, J., Loeffler, M., Sachs, N., Sato, T., Clevers, H. & Galle, J. 2012. On the  
526 biomechanics of stem cell niche formation in the gut--modelling growing organoids.  
527 *FEBS J*, 279, 3475-87.
- 528 Chin, A. M., Hill, D. R., Aurora, M. & Spence, J. R. 2017. Morphogenesis and maturation of the  
529 embryonic and postnatal intestine. *Semin Cell Dev Biol*, 66, 81-93.
- 530 Coudrier, E. & Almeida, C. G. 2011. Myosin 1 controls membrane shape by coupling F-Actin  
531 to membrane. *Bioarchitecture*, 1, 230-235.

532 Crosnier, C., Vargesson, N., Gschmeissner, S., Ariza-Mcnaughton, L., Morrison, A. & Lewis, J.  
533 2005. Delta-Notch signalling controls commitment to a secretory fate in the zebrafish  
534 intestine. *Development*, 132, 1093-104.

535 Dalgin, G., Ward, A. B., Hao Le, T., Beattie, C. E., Nechiporuk, A. & Prince, V. E. 2011. Zebrafish  
536 *mnx1* controls cell fate choice in the developing endocrine pancreas. *Development*,  
537 138, 4597-608.

538 Dereeper, A., Guignon, V., Blanc, G., Audic, S., Buffet, S., Chevenet, F., Dufayard, J. F., Guindon,  
539 S., Lefort, V., Lescot, M., et al. 2008. Phylogeny.fr: robust phylogenetic analysis for the  
540 non-specialist. *Nucleic Acids Res*, 36, W465-9.

541 Distel, M., Wullimann, M. F. & Koster, R. W. 2009. Optimized Gal4 genetics for permanent  
542 gene expression mapping in zebrafish. *Proc Natl Acad Sci U S A*, 106, 13365-70.

543 El-Brolosy, M. A. & Stainier, D. Y. R. 2017. Genetic compensation: A phenomenon in search of  
544 mechanisms. *PLoS Genet*, 13, e1006780.

545 Esteve, C., Francescato, L., Tan, P. L., Bouchany, A., De Leusse, C., Marinier, E., Blanchard, A.,  
546 Bourgeois, P., Brochier-Armanet, C., Bruel, A. L., et al. 2018. Loss-of-Function  
547 Mutations in *UNC45A* Cause a Syndrome Associating Cholestasis, Diarrhea, Impaired  
548 Hearing, and Bone Fragility. *Am J Hum Genet*, 102, 364-374.

549 Hegan, P. S., Mermall, V., Tilney, L. G. & Mooseker, M. S. 2007. Roles for *Drosophila*  
550 *melanogaster* myosin IB in maintenance of enterocyte brush-border structure and  
551 resistance to the bacterial pathogen *Pseudomonas entomophila*. *Mol Biol Cell*, 18,  
552 4625-36.

553 Horne-Badovinac, S., Lin, D., Waldron, S., Schwarz, M., Mbamalu, G., Pawson, T., Jan, Y.,  
554 Stainier, D. Y. & Abdelilah-Seyfried, S. 2001. Positional cloning of heart and soul  
555 reveals multiple roles for PKC lambda in zebrafish organogenesis. *Curr Biol*, 11, 1492-  
556 502.

557 Horstick, E. J., Jordan, D. C., Bergeron, S. A., Tabor, K. M., Serpe, M., Feldman, B. & Burgess, H.  
558 A. 2015. Increased functional protein expression using nucleotide sequence features  
559 enriched in highly expressed genes in zebrafish. *Nucleic Acids Res*, 43, e48.

560 Iuliano, O., Yoshimura, A., Prospero, M. T., Martin, R., Knolker, H. J. & Coudrier, E. 2018. Myosin  
561 1b promotes axon formation by regulating actin wave propagation and growth cone  
562 dynamics. *J Cell Biol*.

563 Kok, F. O., Shin, M., Ni, C. W., Gupta, A., Grosse, A. S., Van Impel, A., Kirchmaier, B. C., Peterson-  
564 Maduro, J., Kourkoulis, G., Male, I., et al. 2015. Reverse genetic screening reveals poor  
565 correlation between morpholino-induced and mutant phenotypes in zebrafish. *Dev*  
566 *Cell*, 32, 97-108.

567 Kolotuev, I., Bumbarger, D. J., Labouesse, M. & Schwab, Y. 2012. Targeted ultramicrotomy: a  
568 valuable tool for correlated light and electron microscopy of small model organisms.  
569 *Methods Cell Biol*, 111, 203-22.

570 Komaba, S. & Coluccio, L. M. 2015. Myosin 1b Regulates Amino Acid Transport by Associating  
571 Transporters with the Apical Plasma Membrane of Kidney Cells. *PLoS One*, 10,  
572 e0138012.

573 Kwan, K. M., Fujimoto, E., Grabher, C., Mangum, B. D., Hardy, M. E., Campbell, D. S., Parant, J.  
574 M., Yost, H. J., Kanki, J. P. & Chien, C. B. 2007. The Tol2kit: a multisite gateway-based  
575 construction kit for Tol2 transposon transgenesis constructs. *Dev Dyn*, 236, 3088-99.

576 Kwon, O., Han, T. S. & Son, M. Y. 2020. Intestinal Morphogenesis in Development,  
577 Regeneration, and Disease: The Potential Utility of Intestinal Organoids for Studying  
578 Compartmentalization of the Crypt-Villus Structure. *Front Cell Dev Biol*, 8, 593969.

579 Lee, C. F., Melkani, G. C. & Bernstein, S. I. 2014. The UNC-45 myosin chaperone: from worms  
580 to flies to vertebrates. *Int Rev Cell Mol Biol*, 313, 103-44.

581 Lehtimäki, J. I., Fenix, A. M., Kotila, T. M., Balistreri, G., Paavolainen, L., Varjosalo, M., Burnette,  
582 D. T. & Lappalainen, P. 2017. UNC-45a promotes myosin folding and stress fiber  
583 assembly. *J Cell Biol*, 216, 4053-4072.

584 Lubarsky, B. & Krasnow, M. A. 2003. Tube morphogenesis: making and shaping biological  
585 tubes. *Cell*, 112, 19-28.

586 Mammoto, T. & Ingber, D. E. 2010. Mechanical control of tissue and organ development.  
587 *Development*, 137, 1407-20.

588 Martin-Belmonte, F. & Mostov, K. 2008. Regulation of cell polarity during epithelial  
589 morphogenesis. *Curr Opin Cell Biol*, 20, 227-34.

590 Mazzolini, R., Dopeso, H., Mateo-Lozano, S., Chang, W., Rodrigues, P., Bazzocco, S., Alazzouzi,  
591 H., Landolfi, S., Hernandez-Losa, J., Andretta, E., et al. 2012. Brush border myosin Ia  
592 has tumor suppressor activity in the intestine. *Proc Natl Acad Sci U S A*, 109, 1530-5.

593 Müller, T., Hess, M. W., Schiefermeier, N., Pfaller, K., Ebner, H. L., Heinz-Erian, P., Ponstingl,  
594 H., Partsch, J., Rollinghoff, B., Köhler, H., et al. 2008. MYO5B mutations cause  
595 microvillus inclusion disease and disrupt epithelial cell polarity. *Nat Genet*, 40, 1163-  
596 5.

597 Ng, A. N., De Jong-Curtain, T. A., Mawdsley, D. J., White, S. J., Shin, J., Appel, B., Dong, P. D.,  
598 Stainier, D. Y. & Heath, J. K. 2005. Formation of the digestive system in zebrafish: III.  
599 Intestinal epithelium morphogenesis. *Dev Biol*, 286, 114-35.

600 Pernier, J., Kusters, R., Bousquet, H., Lagny, T., Morchain, A., Joanny, J. F., Bassereau, P. &  
601 Coudrier, E. 2019. Myosin 1b is an actin depolymerase. *Nat Commun*, 10, 5200.

602 Prospero, M. T., Lepine, P., Dingli, F., Paul-Gilloteaux, P., Martin, R., Loew, D., Knolker, H. J. &  
603 Coudrier, E. 2015. Myosin 1b functions as an effector of EphB signaling to control cell  
604 repulsion. *J Cell Biol*, 210, 347-61.

605 Revenu, C., Athman, R., Robine, S. & Louvard, D. 2004. The co-workers of actin filaments: from  
606 cell structures to signals. *Nat Rev Mol Cell Biol*, 5, 635-46.

607 Revenu, C., Ubelmann, F., Hurbain, I., El-Marjou, F., Dingli, F., Loew, D., Delacour, D., Gilet, J.,  
608 Brot-Laroche, E., Rivero, F., et al. 2012. A new role for the architecture of microvillar  
609 actin bundles in apical retention of membrane proteins. *Mol Biol Cell*, 23, 324-36.

610 Reymann, A. C., Boujemaa-Paterski, R., Martiel, J. L., Guerin, C., Cao, W., Chin, H. F., De La Cruz,  
611 E. M., They, M. & Blanchoin, L. 2012. Actin network architecture can determine  
612 myosin motor activity. *Science*, 336, 1310-4.

613 Robine, S., Huet, C., Moll, R., Sahuquillo-Merino, C., Coudrier, E., Zweibaum, A. & Louvard, D.  
614 1985. Can villin be used to identify malignant and undifferentiated normal digestive  
615 epithelial cells? *Proc Natl Acad Sci U S A*, 82, 8488-92.

616 Rossi, A., Kontarakis, Z., Gerri, C., Nolte, H., Holper, S., Krüger, M. & Stainier, D. Y. 2015. Genetic  
617 compensation induced by deleterious mutations but not gene knockdowns. *Nature*,  
618 524, 230-3.

619 Salas-Cortés, L., Ye, F., Tenza, D., Wilhelm, C., Theos, A., Louvard, D., Raposo, G. & Coudrier, E.  
620 2005. Myosin Ib modulates the morphology and the protein transport within multi-  
621 vesicular sorting endosomes. *J Cell Sci*, 118, 4823-32.

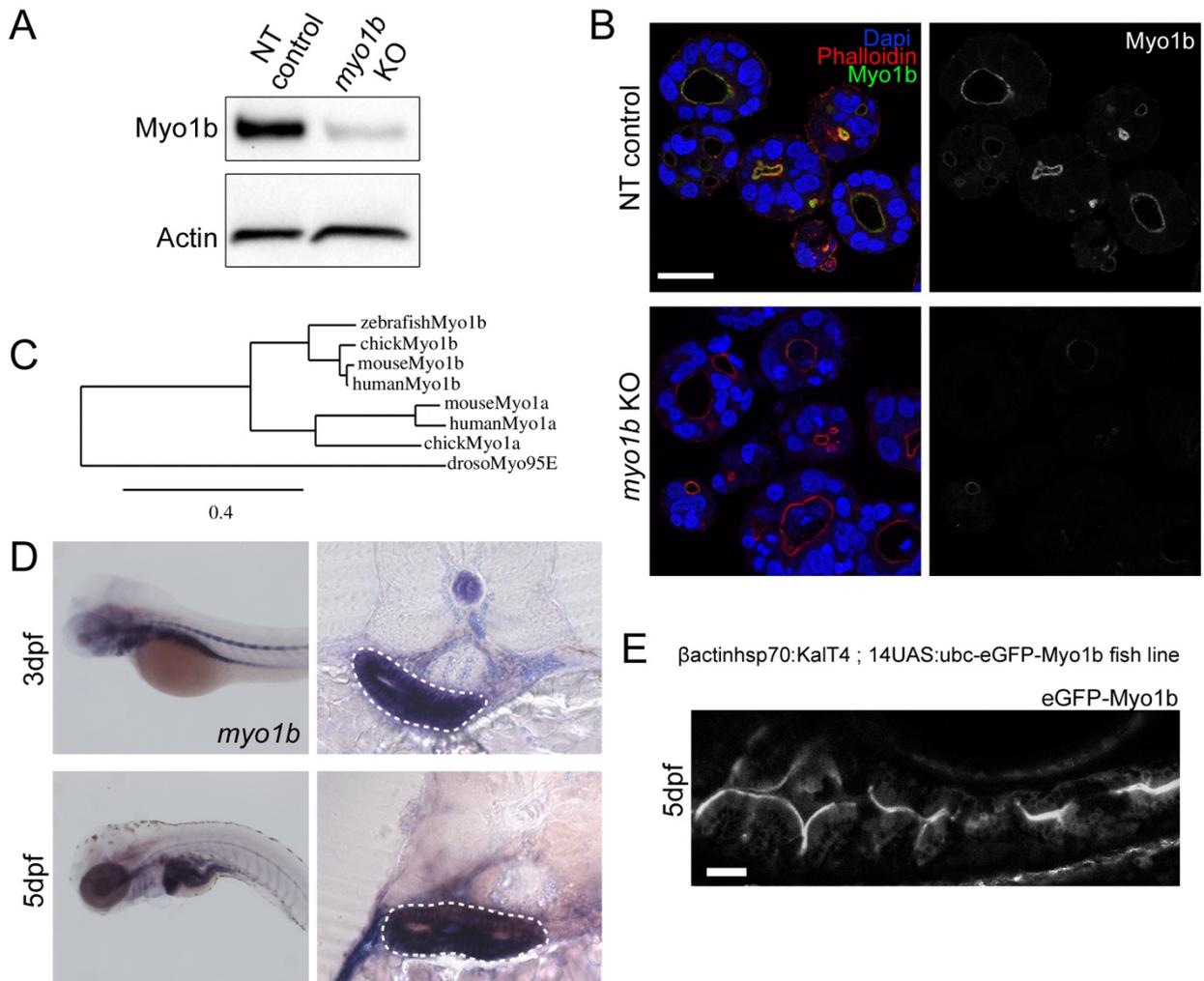
622 Sidhaye, J., Pinto, C. S., Dharap, S., Jacob, T., Bhargava, S. & Sonawane, M. 2016. The zebrafish  
623 goosepimples/myosin Vb mutant exhibits cellular attributes of human microvillus  
624 inclusion disease. *Mech Dev*, 142, 62-74.

625 Thisse, B. & Thisse, C. 2004. Fast Release Clones: A High Throughput Expression Analysis.  
626 *ZFIN Direct Data Submission* (<http://zfin.org/>). .

627 Tyska, M. J., Mackey, A. T., Huang, J. D., Copeland, N. G., Jenkins, N. A. & Mooseker, M. S. 2005.  
628 Myosin-1a is critical for normal brush border structure and composition. *Mol Biol Cell*,  
629 16, 2443-57.

630 Wallace, K. N., Akhter, S., Smith, E. M., Lorent, K. & Pack, M. 2005. Intestinal growth and  
631 differentiation in zebrafish. *Mech Dev*, 122, 157-73.  
632 Yamada, A., Mamane, A., Lee-Tin-Wah, J., Di Cicco, A., Prevost, C., Levy, D., Joanny, J. F.,  
633 Coudrier, E. & Bassereau, P. 2014. Catch-bond behaviour facilitates membrane  
634 tubulation by non-processive myosin 1b. *Nat Commun*, 5, 3624.  
635

636



Revenu et al., Figure 1

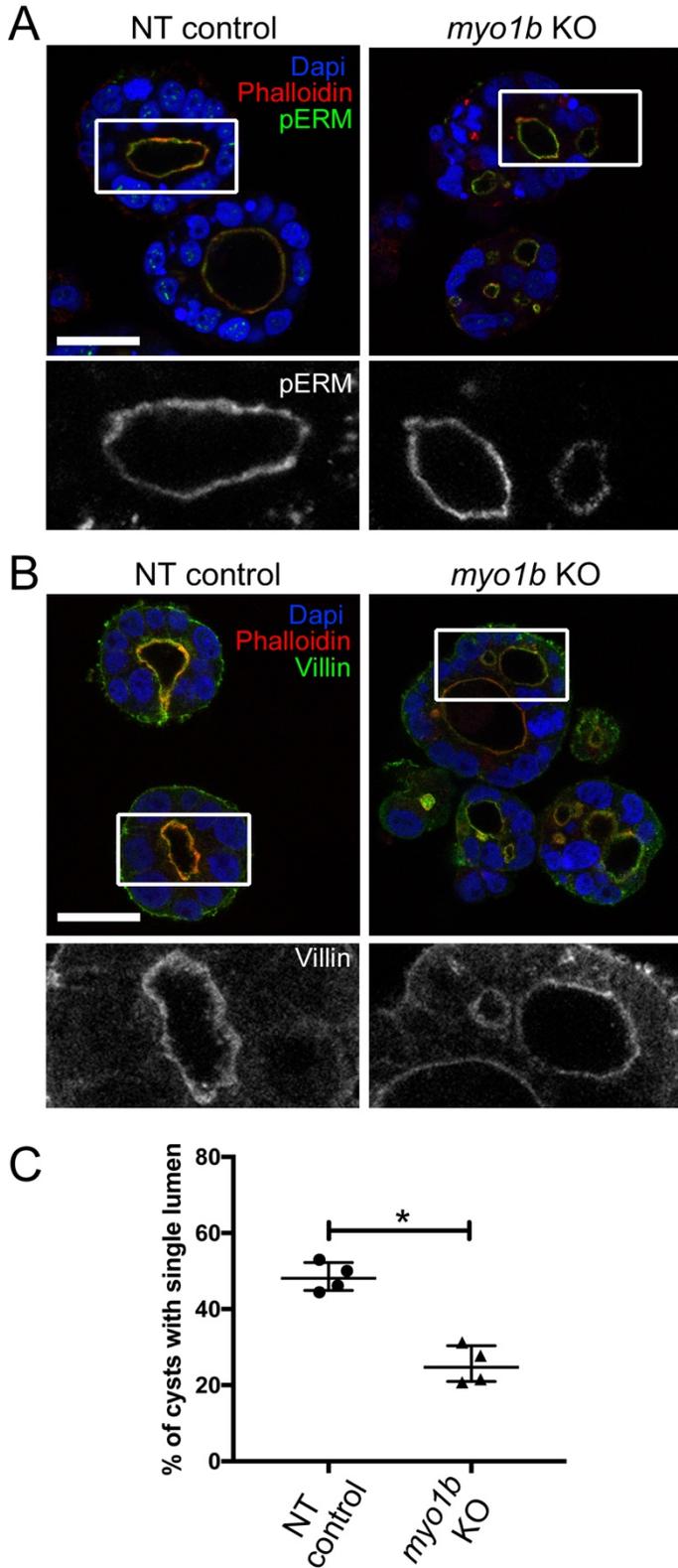
637

638 **Figure 1**

639 **Myo1b expression and apical localisation in gut epithelial cells.** **A-** Western blot analysis of  
 640 Myo1b expression in extracts from non-targeted (NT) control and *myo1b* targeted Caco-2 cells  
 641 (KO) using CRISPR/Cas9. **B-** Confocal sections of Caco-2 3D cultures stained for Myo1b, F-actin  
 642 (phalloidin) and nuclei (Dapi). **C-** Phylogenetic tree based on protein sequence of zebrafish, chick,  
 643 human and mouse Myo1b and Myo1a and drosophila Myosin95E. **D-** In situ hybridization for  
 644 *myo1b* transcripts on 3 and 5dpf zebrafish larvae whole mounts (left panel) and cross-sections at  
 645 the level of the intestinal bulb (right panels). On sections, the forming intestinal bulb is circled  
 646 with white dashed lines. **E-** Live, longitudinal (antero-posterior axis) confocal section of the  
 647 intestinal bulb of a 5dpf zebrafish larva expressing the transcription activator KalT4 driving the  
 648 expression of the eGFP-Myo1b transgene under the control of an upstream activating sequence  
 649 (UAS). The precise construction of the transgenes is annotated in the panel. Scale bars 30 $\mu$ m.

650

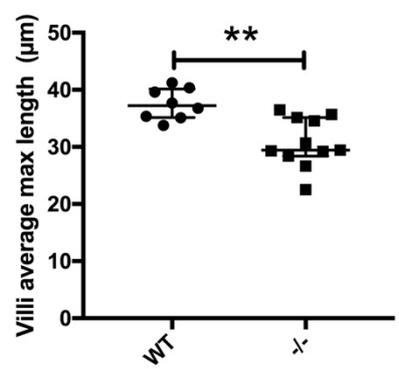
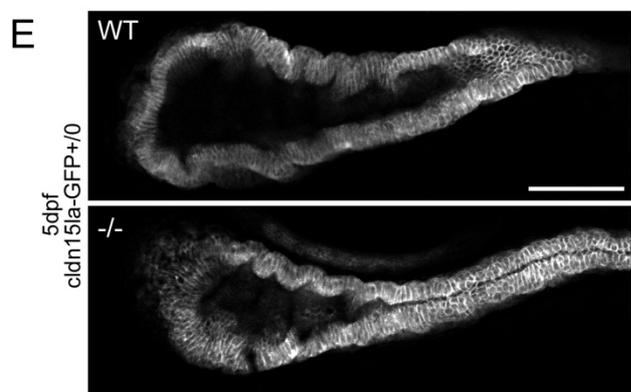
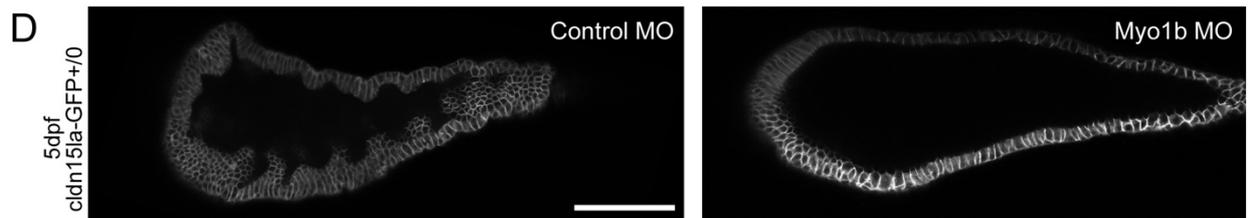
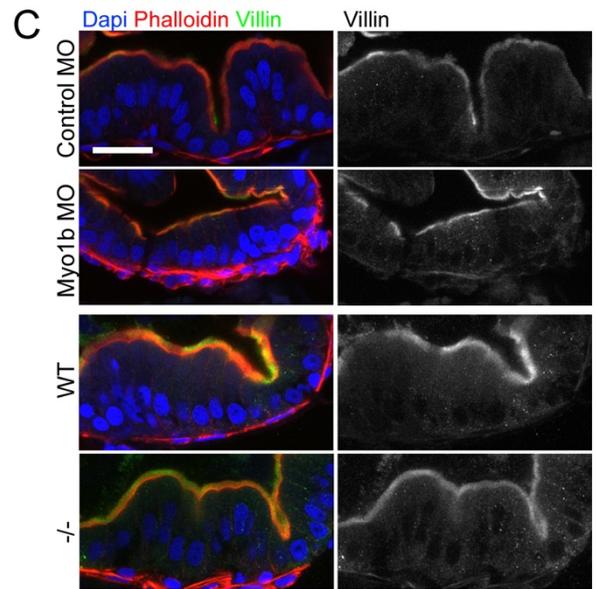
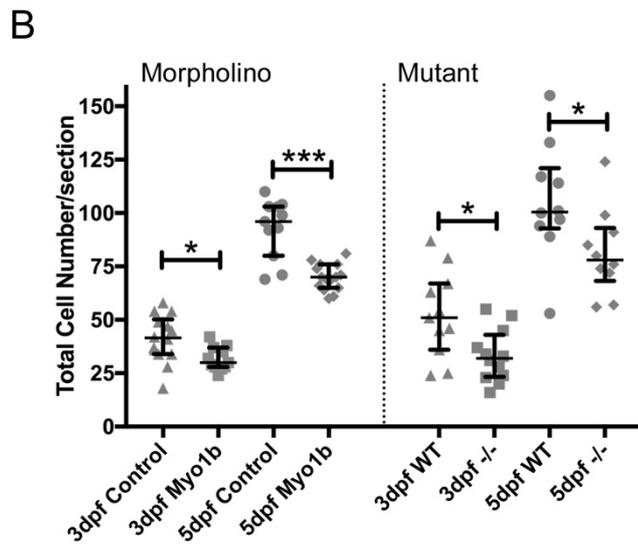
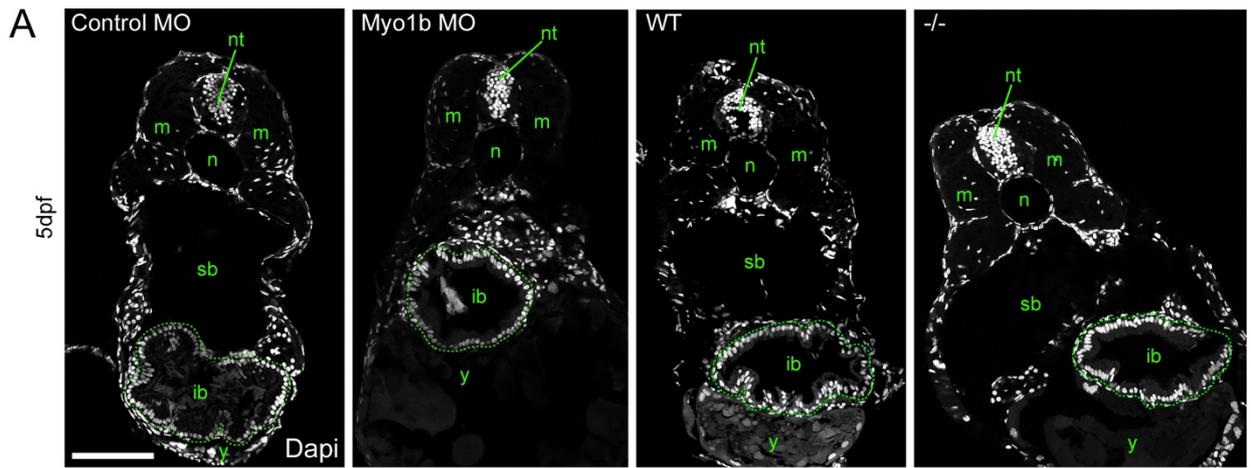
651  
652



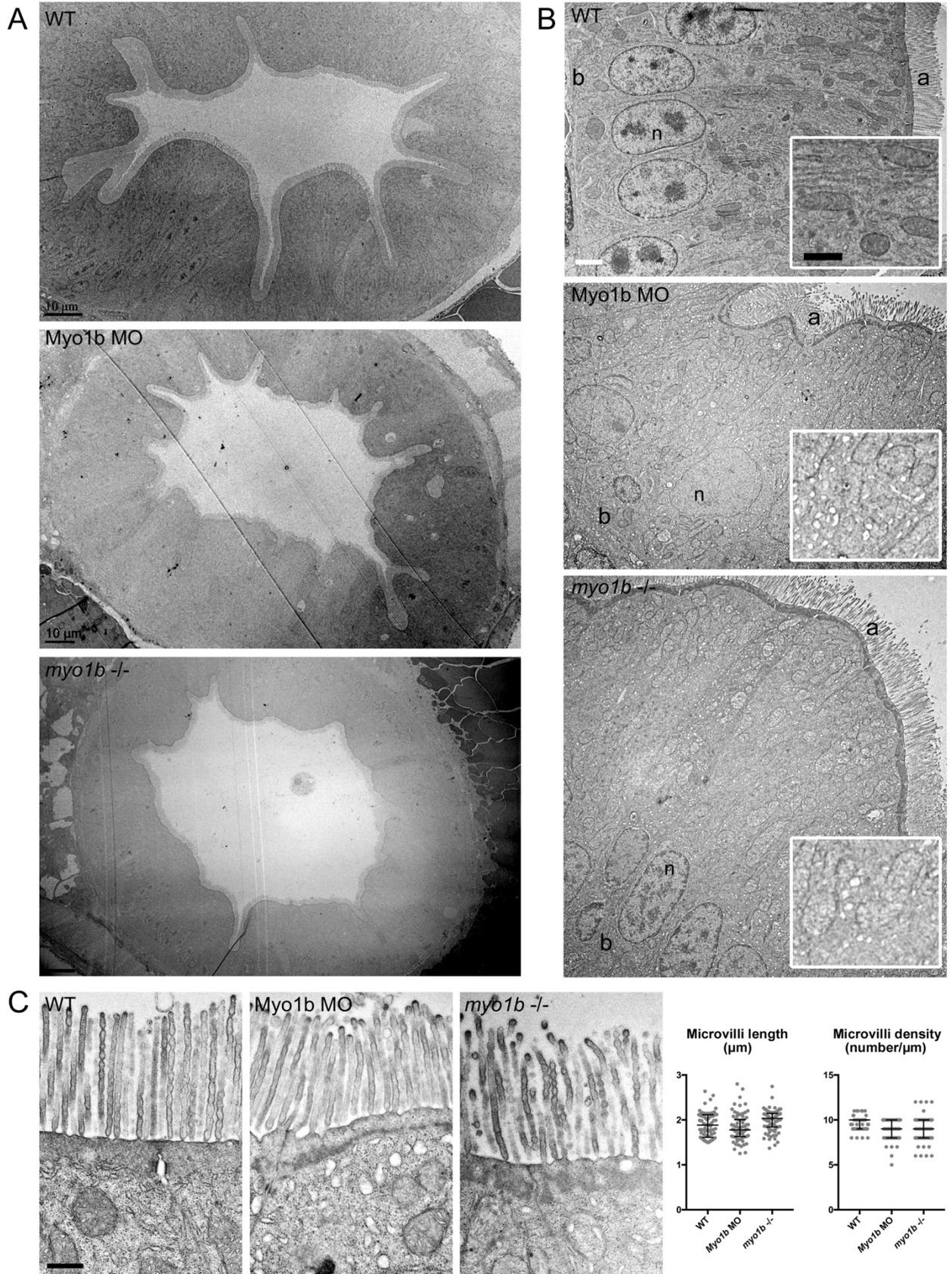
**Figure 2**

**Enterocyte 3D cyst organization is affected in the absence of Myo1b despite normal apico-basal polarization.** Confocal sections of NT control and *myo1b* KO Caco-2 3D cultures stained for the apical and microvilli markers phospho-Ezrin (pERM, **A**) and Villin (**B**). F-actin (phalloidin) and nuclei (Dapi) are stained, scale bars 30µm, boxed areas showed in insets are enlarged 2.5x. **C**- Quantification of the percentage of well-formed cysts with a single central lumen in NT control and *myo1b* KO Caco-2 3D cultures. Data represented are median and interquartile range, Wilcoxon test, \*p<0.05.

Revenu et al., Figure 2



675 **Figure 3**  
676 **Myo1b knock-down and knock-out impair intestinal bulb fold development.** A- Confocal  
677 single optical sections stained with nuclear labelling (Dapi) of 5dpf larvae injected with control  
678 and Myo1b Morpholinos (MO), and of 5dpf wild-type (WT) and *myo1b*<sup>-/-</sup> (-/-) larvae. ib intestinal  
679 bulb (circled with dashed lines), m muscles, n notochord, nt neural tube, sb swim bladder, y yolk.  
680 Scale bar=100µm. B- Quantifications from Dapi stained sections of the total number of cells per  
681 section at 3 and 5dpf in the four conditions. Data represented are median and interquartile range,  
682 Wilcoxon test, \*p<0.05, \*\*\*p<0.001. C- Confocal optical sections of the intestinal bulb of 5dpf  
683 larvae in the four conditions stained for the microvilli marker Villin, F-actin (phalloidin) and nuclei  
684 (Dapi) showing the preserved apico-basal polarity of enterocytes when Myo1b is affected. Scale  
685 bar 20µm. D- Single confocal planes of live 5dpf larvae expressing Cldn15la-GFP injected with  
686 control MO (left) and Myo1b MO (right). Note the flat epithelium in the Myo1b MO condition.  
687 Scale bar=100µm. E- Single confocal planes of live 5dpf WT and *myo1b*<sup>-/-</sup> larvae expressing  
688 Cldn15la-GFP and quantification of the average length of the 3 longest folds per intestinal bulb  
689 analysed. Scale bar=100µm. Data presented are median and interquartile range, Wilcoxon test,  
690 \*\*p<0.01.  
691



692 Revenu et al., Figure 4

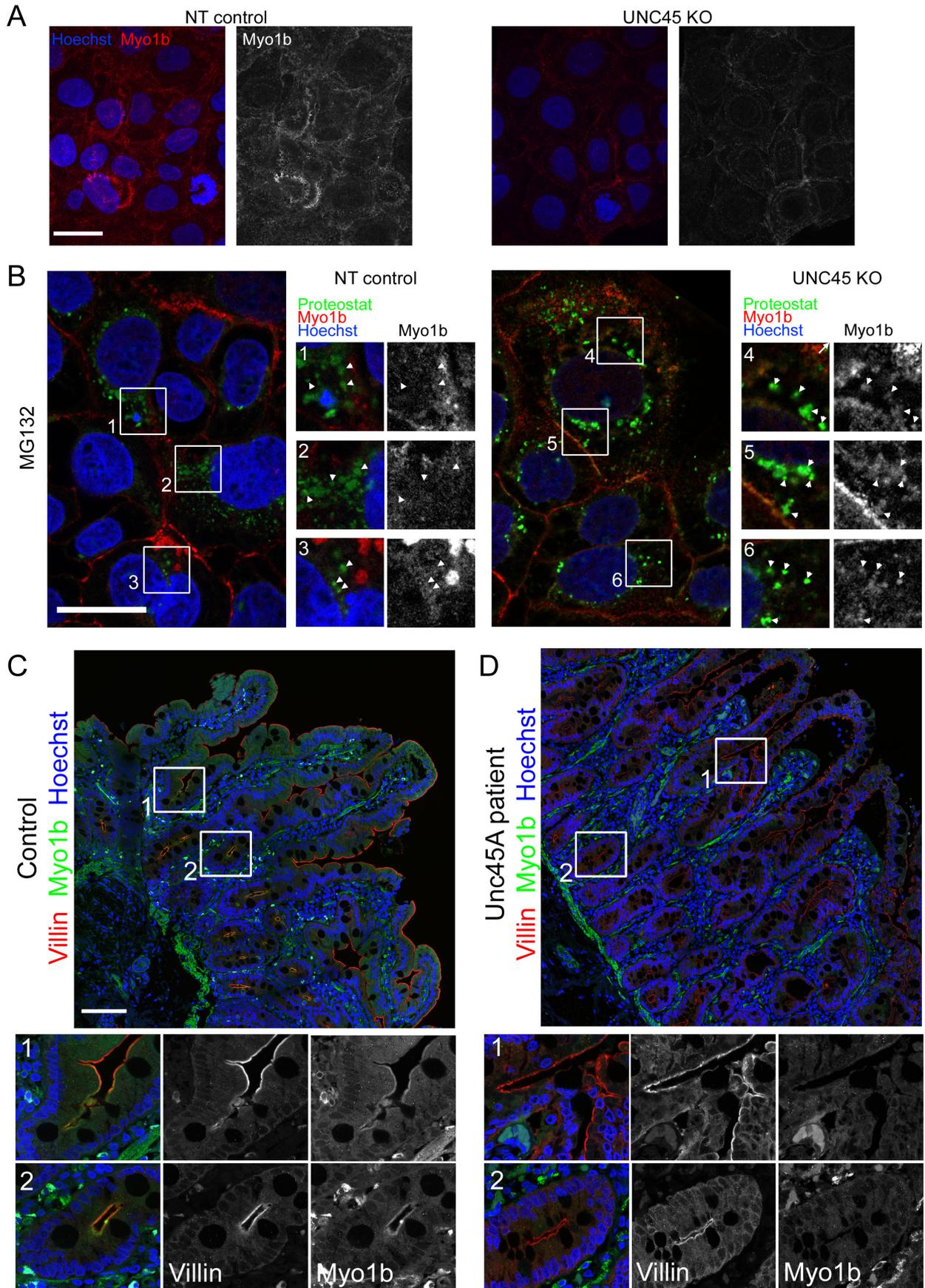
693

694

695 **Figure 4**

696 **Electron Microscopy confirms folding defects and shows affected trafficking.** Transmission  
697 electron micrographs of sections of intestinal bulbs from WT, Myo1b MO and *myo1b*<sup>-/-</sup> 5dpf  
698 larvae presenting a general view of the folds of the epithelium (**A**, scale bars 10µm) and of the  
699 apico-basally polarized enterocytes (**B**, scale bar 2µm ; b basal, a apical, n nuclei). Insets in **B**,  
700 show higher magnifications of the cytoplasm region to highlight the accumulation of vesicles in  
701 Myo1b MO and *myo1b*<sup>-/-</sup> samples, scale bar 1µm. **C**- Transmission electron micrographs of  
702 sections of intestinal bulbs from WT, Myo1b MO and *myo1b*<sup>-/-</sup> 5dpf larvae illustrating the  
703 organization of the brush border, and quantifications of the average length and density of the  
704 intestinal microvilli in the different conditions. Data presented are median and interquartile range.  
705 Scale bar 500 nm.

706



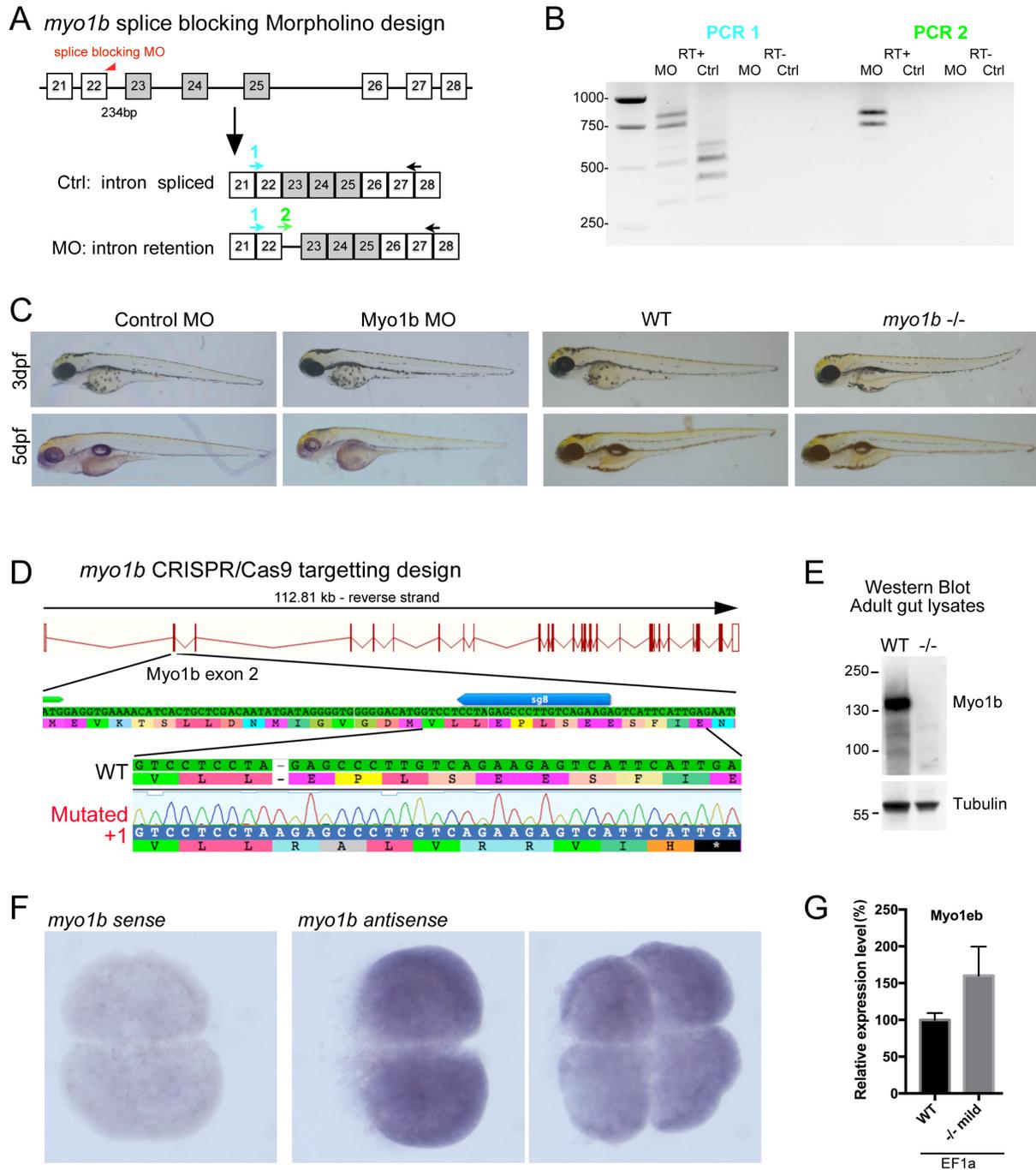
707

708

Revenu et al., Figure 5

709 **Figure 5**  
710 **Myo1b expression is destabilized in UNC45A depleted cells and in biopsies from UNC45A**  
711 **mutated patients. A-** Immunohistochemistry analyses of Myo1b in non-targeted (NT) control and  
712 UNC45A deficient (KO) Caco-2 cells show decreased Myo1b levels. Pictures are maximal  
713 projections of confocal stacks, Hoechst labels nuclei, scale bar 30µm. **B-** Confocal sections of NT  
714 control and UNC45A KO Caco-2 cells treated with the proteasome inhibitor MG132 and stained  
715 for Myo1b and the aggresome probe Proteostat. Hoechst labels nuclei, scale bar 30µm. Boxed  
716 areas showed in insets are enlarged 2x, arrowheads point at Proteostat-labelled protein aggregates  
717 and highlight colocalisation with Myo1b proteins in UNC45A KO cells. **C,D-** Confocal sections  
718 of a human biopsy from a healthy patient (control, C) and from a UNC45A LOF patient (D)  
719 immuno-labelled for the microvilli marker Villin and for Myo1b ; Hoechst labels nuclei, scale bar  
720 100µm. Boxed areas showed in insets are enlarged 3x, and highlight the apical localisation of  
721 Myo1b in control tissue (C) at the base of the villi (1) and in crypts (2), which is essentially lost in  
722 the UNC45A LOF tissue (D).

## Supplementary Figures and Table

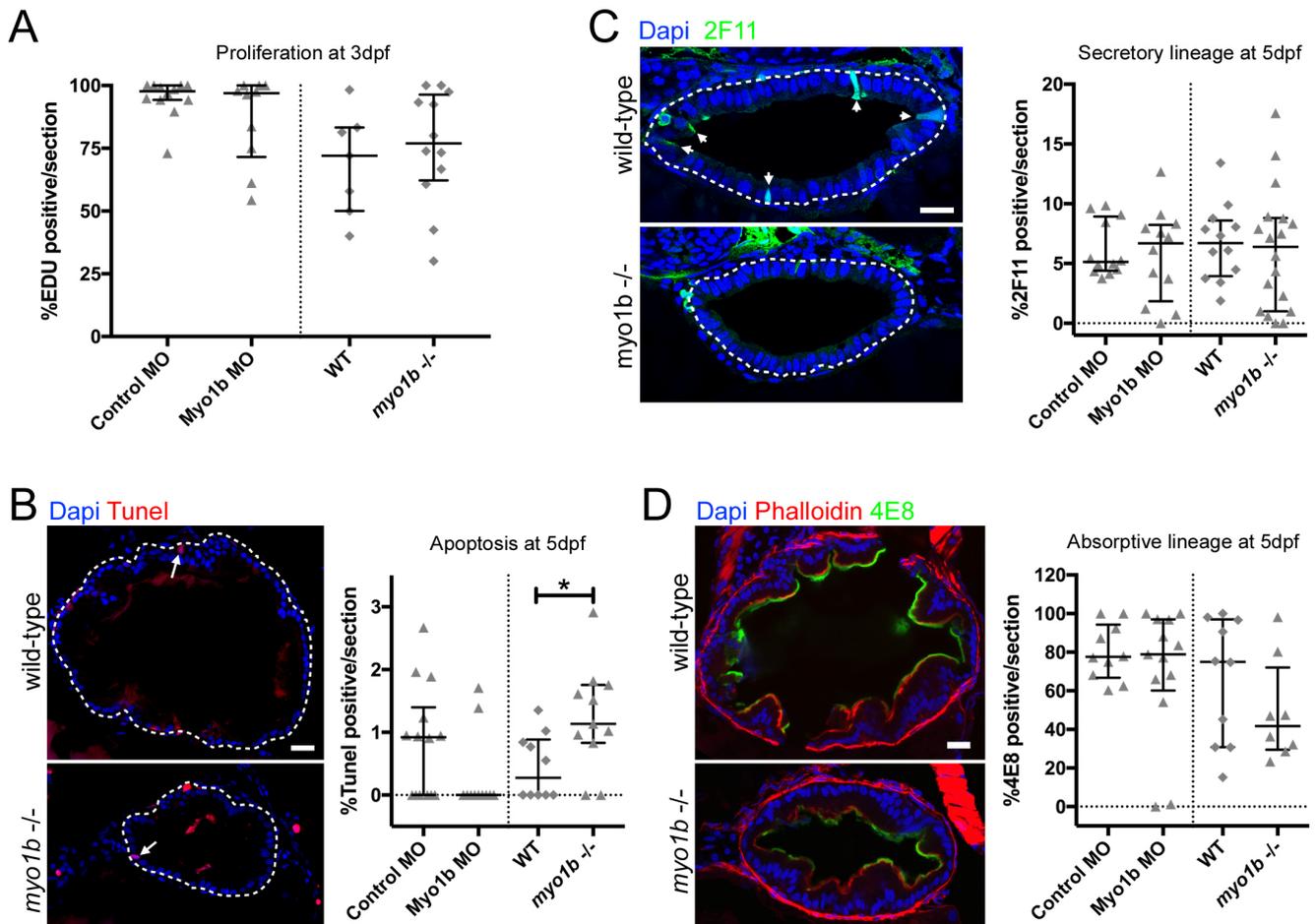


Revenu et al., Figure S1

### Figure S1

**Myo1b Morpholino and CRISPR mutant design and validation.** **A-** Schematics of the design and **B-** DNA gel of the RT-PCR performed to control Myo1b-MO knock-down efficiency. Higher bands amplified in PCR1 MO compared to control and bands amplified in PCR2 MO correspond to Myo1b cDNA retaining the intron targeted by Myo1b-sMO, as verified by sequencing. In B, the multiple bands amplified, both in control and MO conditions, correspond to expected splicing variants of exons 23, 24 and 25 (highlighted in grey)

as checked by sequencing. RT- is the control RT without superscript compared to RT+. **C-** Bright field pictures of 3 and 5dpf larvae presenting the phenotypes of control and Myo1b Morpholinos, WT and *myo1b*<sup>-/-</sup> larvae. **D-** Schematics of CRISPR/Cas9-mediated gene disruption at the *myo1b* genomic locus. The sgRNA (sgB, blue arrow) was targeting exon 2 downstream the start codon (ATG, green arrow). Compared to the WT sequence, the mutated allele displayed an insertion of 1bp generating a frame shift from amino-acid 21 and a premature STOP codon after 29 amino-acids. **E-** Western Blot with antibodies against Myo1b and Tubulin on lysates of dissected guts from WT and <sup>-/-</sup> adults. **F-** Negative control with a sense probe and in situ hybridisation with a *myo1b* anti-sense probe on wild-type embryos at 2 and 4 cell-stages showing maternal contribution for *myo1b* mRNA. **G-** RT-QPCR of *myo1eb* expression at 3dpf with EF1a used as reference gene, normalised on expression of the WT samples. Shown are mean and sem (WT=100.0±9.2, <sup>-/-</sup> =160.1±39.5 , n=6).



Revenu et al., Figure S2

### Figure S2

**Proliferation, apoptosis and differentiation are essentially unaffected in Myo1b MO and mutant conditions.** **A-** Quantifications from EDU and Dapi stained sections of the proportion of cells in S-phase at 3dpf do not reveal significant differences in the proliferative rate of Control vs Myo1b MO and of WT vs *myo1b*<sup>-/-</sup> samples. **B-** Confocal sections of intestinal bulbs stained for apoptosis (Tunel, red) at 5dpf in WT and *myo1b*<sup>-/-</sup> samples and quantifications of the proportion of apoptotic cells in the four conditions. **C,** **D-** Confocal sections of intestinal bulbs stained for differentiation markers in WT and *myo1b*<sup>-/-</sup> samples and quantifications of the proportion of differentiated cells in the four conditions. At 5dpf, neither differentiation of the secretory lineage (2F11, **C**) nor differentiation of the absorptive lineage (4E8, **D**) are significantly altered. For all quantifications, data represented are median and interquartile range, Wilcoxon test, \* $p < 0.05$ . For confocal images, nuclei are counterstained with Dapi (blue), bars=20 $\mu$ m.

Rank	Accession	Description	MW [kDa]	Coverage %	#Unique Peptides	#PSMs
1	P46735	Unconventional myosin-Ib OS=Mus musculus [MYO1B_MOUSE]	128,483	65,31	102	633
2	Q99KD5	Protein unc-45 homolog A OS=Mus musculus [UN45A_MOUSE]	103,3818	22,14	24	53

**Table S1**

**Mass spectrometry result of the GFP-Myo1b pull down assay.** UNC45A ranks second, directly after Myo1b. MW, molecular weight; #Unique Peptides, number of distinct peptide sequences identified; #PSMs (peptide spectrum matches) total number of identified peptide sequences for the protein.