



**HAL**  
open science

# Recherche des cibles de Rbf1 impliquées dans ses activités pro et anti-apoptotiques chez la drosophile

Marion Hoareau

► **To cite this version:**

Marion Hoareau. Recherche des cibles de Rbf1 impliquées dans ses activités pro et anti-apoptotiques chez la drosophile. Biologie moléculaire. Université Paris-Saclay, 2023. Français. NNT : 2023UP-ASL014 . tel-04224311

**HAL Id: tel-04224311**

**<https://theses.hal.science/tel-04224311>**

Submitted on 2 Oct 2023

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

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

# Recherche des cibles de Rbf1 impliquées dans ses activités pro et anti-apoptotiques chez la drosophile

*Search for Rbf1 targets involved in its pro- and anti-apoptotic activities in  
Drosophila*

## Thèse de doctorat de l'université Paris-Saclay

École doctorale n°577 : Structure et Dynamique des Systèmes Vivants (SDSV)  
Spécialité de doctorat : Sciences de la vie et de la Santé  
Graduate School : Life Sciences and Health (LSH)  
Référent : Université de Versailles-Saint-Quentin-en-Yvelines (UVSQ)

Thèse préparée dans l'unité de recherche **Laboratoire de Génétique et Biologie  
Cellulaire (LGBC)**, sous la direction d'**Isabelle GUENAL**, Professeure, et le co-  
encadrement de **Jessie COLIN**, Maîtresse de conférences

Thèse soutenue à Montigny-le-Bretonneux, le 30 mars 2023, par

**Marion HOAREAU**

## Composition du Jury

Membres du jury avec voix délibérative

<b>Sébastien BLOYER</b> Professeur, Université Paris-Saclay	Président
<b>Sophie LOUVET-VALLEE</b> Professeure, Sorbonne Université	Rapporteur & Examinatrice
<b>Véronique MONNIER</b> Maîtresse de conférences, Université Paris-Cité	Rapporteur & Examinatrice
<b>Elisabeth PETIT-TEIXEIRA</b> Professeure, Université Paris-Saclay	Examinatrice



**Titre :** Recherche des cibles de Rbf1 impliquées dans ses activités pro et anti-apoptotiques chez la drosophile

**Mots clés :** Rbf1 | Drosophile | Apoptose | Transcriptome | Disques imaginaux

**Résumé :** Le gène suppresseur de tumeurs *RB1* code la protéine pRb, dont l'activité est altérée dans la plupart des cancers humains. pRb est un co-facteur de transcription. Généralement décrit comme régulateur du cycle cellulaire, il est également impliqué dans d'autres processus comme l'apoptose. Cependant, ce rôle est moins bien caractérisé. Nous utilisons Rbf1, l'homologue de pRb chez la drosophile, pour étudier son rôle dans l'apoptose. Rbf1 induit l'apoptose quand il est surexprimé dans le disque imaginal d'aile larvaire mais protège de l'apoptose induite par le pro-apoptotique p53 dans les neurones. Son activité pro-apoptotique implique un contrôle transcriptionnel de régulateurs de l'apoptose dépendant du complexe dREAM qui comprend entre autres Rbf1 et le facteur de transcription dE2F2. L'activité anti-apoptotique de Rbf1 est bien moins caractérisée. Elle pourrait impliquer des régulateurs différents au sein du complexe dREAM et des gènes cibles différents. La compréhension des mécanismes de régulation médiés par ce complexe hautement conservé est possible notamment grâce aux études portant sur les organismes modèles, dont la drosophile.

Ce projet de thèse visait à identifier les cibles directes de Rbf1 dans un tissu dans lequel il est pro-apoptotique, et dans un tissu dans lequel il est anti-apoptotique. Afin de déterminer les régulations transcriptionnelles pouvant expliquer ce rôle opposé vis-à-vis de l'apoptose, nous avons utilisé des approches de ChIP-seq et de RNA-seq. J'ai mis au point les outils nécessaires pour ces approches. Notamment, un protocole d'isolement en masse des disques imaginaux d'aile, permettant de s'affranchir de la dissection et de la variabilité qu'elle induit, a été mis au point et publié. Bien que des difficultés techniques ne m'aient pas permis d'identifier de gène régulé différentiellement dans ces deux tissus, je me suis formée à l'analyse bio-informatique de données de séquençage et j'ai réalisé ces traitements et analyses.

**Title :** Search for Rbf1 targets involved in its pro- and anti-apoptotic activities in *Drosophila*

**Keywords :** Rbf1 | Drosophila | Apoptosis | Transcriptome | Imaginal discs

**Abstract :** The *RB1* tumor suppressor gene encodes the pRb protein, whose activity is altered in most human cancers. pRb is a transcriptional cofactor. Generally described as a cell cycle regulator, it is also involved in other vital processes, such as apoptosis. However, this role is less well characterized. We use Rbf1, the *Drosophila* homolog of pRb, to study its role in apoptosis. Rbf1 induces apoptosis when overexpressed in the larval wing imaginal disc but protects against apoptosis induced by the pro-apoptotic p53 in neurons. Its pro-apoptotic activity involves transcriptional control of dREAM complex-dependent apoptosis regulators, including Rbf1 and the transcription factor dE2F2. The anti-apoptotic activity of Rbf1 is much less characterized. It may involve different regulators within the dREAM complex and different target genes. Understanding the regulatory mechanisms mediated by this highly conserved complex is possible especially through studies in model organisms, including *Drosophila*.

This thesis project aimed to identify the direct targets of Rbf1 in a tissue in which it is pro-apoptotic and in a tissue in which it is anti-apoptotic. We used ChIP-seq and RNA-seq approaches to determine the transcriptional regulations that could explain this opposite role towards apoptosis. I developed the necessary tools to these techniques. In particular, a protocol for mass isolation of wing imaginal discs, allowing to avoid dissection and the variability it induces, has been developed and published. Although technical difficulties did not allow me to identify differentially regulated genes in these two tissues, I trained myself in bioinformatics analysis of sequencing data and performed these treatments and investigations.



# REMERCIEMENTS

Merci à **Isabelle**, tout d'abord pour ses cours à l'UVSQ qui m'ont fait me passionner pour la biologie cellulaire, puis pour m'avoir donné une chance lors d'un stage d'été, et enfin pour avoir encadré cette thèse dans les bons moments et les moins bons, mais toujours avec énergie et bonne humeur. Merci à **Jessie**, de m'avoir mis les pieds dans le plat de la biologie moléculaire et de m'avoir fait plus confiance sur ce projet que je n'aurais jamais confiance en moi-même. Merci pour votre écoute, vos réponses à mes nombreuses questions et surtout de m'avoir fait grandir aussi bien scientifiquement que personnellement.

Merci à tout le LGBC d'avoir été et de rester cette grande famille au-delà de simples relations de laboratoire. Merci à **Bernard** d'avoir répondu à toutes mes paniques des démarches de l'école doctorale. Merci à **Aurore** et à nos escapades légumières, merci à **Sylvina** la meilleure capitaine de curling, merci à **Nathalie** pour nos parties de badminton, aux **Sebs** et leurs jeux de mots, à **Maxence** et ses chandeliers, à la relève **Kamila**, à **Jean-Marc**, à **Arnaud**, à **Vincent**, à **José** et à tous ceux que j'ai hâte de connaître. Merci à **Delphine** et **Christine** de m'avoir accueillie de nombreux midis dans leurs bureaux, pour découvrir anecdotes et animaux du jour. Merci également à l'équipe EPIM pour leur accueil et leur confiance, je pars, mais pas très loin !

Si certaines n'ont pas été citées parmi les membres du LGBC, c'est parce qu'elles ont depuis bien longtemps outrepassé le rang de collègues, de « potes du boulot », pour être arrivées au rang d'amies. Je ne pourrais jamais assez les remercier, mais je vais essayer. Merci à **Juliette**, ma partenaire de crime, mon agent McCallaghan, ma Micheline, mon informatrice dinosaure (t'as vu Flo, j'ai mis le mot « dinosaure » !), celle qui me pousse des deux mains hors de ma zone de confort tout en me laissant en parfaite sécurité. Merci d'avoir été là pour les blind-tests, les soirées films et les podcasts, mais aussi mes moments de doutes, mes moments de colère et mes autres moments de doutes (il y en eu beaucoup). Merci à **Mélanie**, ma petite sœur de labo, d'avoir été la meilleure relève que le bureau 31 ait pu espérer. Merci pour les infos, les sorties et je ne pensais pas dire ça un jour, merci pour le thé ! Et surtout merci d'être venue me chercher dans ma coquille, je n'aurais voulu louper tous ces moments pour rien au monde. Merci à **Florine**, qui pourrait espérer rencontrer un jour une meilleure version de soi-même ? J'ai eu cette chance avec toi et je suis bien contente de l'avoir saisie ! Merci pour les cours de ski, de rire à mes jeux de mots nuls, de crier famine à midi pile et d'accompagner mes pintes, mais surtout, surtout, merci pour tout le reste.

Merci à tout le 3<sup>e</sup> étage passé et présent : **Hamadoun, Gina, Cassie, Marion, Elodie, Simon, Bilade** et tous les autres.

Merci à ceux qui ont accompagné mes cours mais aussi bien plus : la **Unicorn Team** et la **Crème de la Crème**. Merci à **Pierrick**, pour les soirées nanars, pour les soirées jeux vidéos, pour les soirées jeux de rôle et pour les soirées à refaire le monde et parler de la vie.

Un grand, énorme et évident merci à ma famille. Merci d'avoir toujours été là, n'importe où et n'importe quand. Je vous aime.

Merci à **Mélina**, d'avoir été là dans les bons moments comme dans les mauvais.

*Maintenant que tout ça est dit, il est temps de retourner aux bonnes vieilles habitudes : qu'est-ce qu'un rongeur dans l'espace ? (Un hamstéroïde)*

# TABLE DES MATIERES

<b>REMERCIEMENTS.....</b>	<b>2</b>
<b>LISTE DES ABREVIATIONS.....</b>	<b>7</b>
<b>INTRODUCTION .....</b>	<b>10</b>
<b>A- L' APOPTOSE.....</b>	<b>11</b>
I- Histoire de l'étude de l'apoptose .....	11
1. Découverte et morphologie .....	11
2. Découverte des mécanismes moléculaires .....	13
3. Le modèle drosophile dans l'étude de l'apoptose .....	15
II- Rôles et mécanismes de l'apoptose .....	17
1. L'apoptose : un mécanisme physiologique indispensable .....	17
2. Un mécanisme précis et régulé .....	18
a. Les acteurs de l'apoptose .....	18
Les caspases.....	19
Les protéines de la famille Bcl-2.....	20
Les inhibiteurs de caspases (IAPs) et leurs inhibiteurs .....	22
Les acteurs de l'apoptose chez la drosophile .....	23
b. L'apoptose : une cascade protéique bien huilée .....	25
<b>B- LES PROTEINES DE LA FAMILLE RB .....</b>	<b>29</b>
I- Du rétinoblastome à Rb .....	29
1. Du rétinoblastome.....	29
2. ...à la protéine Rb.....	30
II- pRb, un co-répresseur de transcription conservé .....	33
1. De nombreuses études pour caractériser les cibles de pRb et de ses homologues.....	33
2. Différents modes de régulation de la transcription .....	35
a. Par liaison à des protéines de régulation de la transcription .....	35
b. Par modification de la chromatine .....	37

Condensines et facteurs de remodelage de la chromatine.....	37
Modifications post-traductionnelles des histones.....	37
Un cas intermédiaire : le complexe NuRD.....	39
<b>III- Des rôles essentiels dans de nombreux processus .....</b>	<b>39</b>
1. pRb, régulateur majeur du cycle cellulaire .....	39
2. Pro- ou anti-apoptotique ? .....	41
a. <i>Les disques imaginaires d'aile et d'œil.....</i>	<i>41</i>
b. <i>Les RB, activateurs de l'apoptose mitochondriale dans les cellules en</i> <i>prolifération.....</i>	<i>42</i>
c. <i>Un rôle anti-apoptotique dans les cellules non prolifératives.....</i>	<i>46</i>
3. Et bien plus encore ! .....	48
a. <i>Développement et différenciation.....</i>	<i>48</i>
b. <i>Sénescence.....</i>	<i>49</i>
c. <i>Autophagie.....</i>	<i>49</i>
d. <i>Réparation de l'ADN.....</i>	<i>50</i>
<b>C- LE COMPLEXE DREAM.....</b>	<b>50</b>
<b>D- PROBLEMATIQUE DE LA THESE .....</b>	<b>77</b>
<b>MATERIELS ET METHODES .....</b>	<b>79</b>
<b>RESULTATS .....</b>	<b>86</b>
I- <i>Mise au point d'un protocole d'isolation en masse de disques</i> <i>imaginaires d'ailes.....</i>	<i>87</i>
II- <i>Outils de validation des cibles identifiées .....</i>	<i>113</i>
1. <i>Activité pro-apoptotique dans l'aile.....</i>	<i>113</i>
2. <i>Activité anti-apoptotique : de l'œil aux neurones .....</i>	<i>115</i>
a. <i>L'œil, un modèle difficilement utilisable.....</i>	<i>115</i>
b. <i>Autre souche, autre phénotype .....</i>	<i>117</i>
III- <i>Le ChIP-seq : une méthode d'identification des promoteurs liés par</i> <i>Rbf1</i>	<i>121</i>
1. <i>Mise au point de l'immunoprécipitation de chromatine .....</i>	<i>121</i>
a. <i>Anticorps : viser la protéine ou une étiquette ?.....</i>	<i>122</i>



b.	<i>Quelles billes d'immunoprécipitation ?</i> .....	122
c.	<i>Le début des ennuis : essais de ChIP contrôle VS Rbf1</i> .....	123
	Vérification de l'efficacité des billes d'immunoprécipitation.....	125
	Vérification de l'absence d'une étiquette HA dans la souche contrôle.....	126
	Un dernier espoir : une purification de l'ADN trop peu efficace.....	127
2.	Séquençage des fragments obtenus.....	130
3.	Traitement des données de séquençage.....	130
a.	<i>Contrôle qualité des données</i> .....	131
b.	<i>Alignement sur un génome de référence</i> .....	132
c.	<i>Filtrage des données</i> .....	132
d.	<i>Peak calling</i> .....	132
4.	Visualisation des données dans IGV et analyse ChIP-seq.....	133
	<b>IV- Le RNA-seq : une méthode d'identification des régulations</b>	
	<b>dépendantes de Rbf1</b> .....	134
1.	Tri magnétique ou tri par cytométrie en flux ?.....	134
2.	Traitement et analyse des données.....	139
a.	<i>Contrôle qualité</i> .....	139
b.	<i>Alignement sur le génome de référence</i> .....	142
	Création du génome dm6.....	142
	Alignement des reads sur le génome.....	142
	Création d'un index du génome.....	143
c.	<i>Conversion des reads alignés en fichiers de comptage des reads</i> ....	143
3.	Analyse différentielle des données de séquençage.....	143
a.	<i>Description des données brutes, variabilité</i> .....	143
b.	<i>Analyse différentielle</i> .....	145
	<b>DISCUSSION</b> .....	<b>150</b>
I-	Un protocole pour s'affranchir de la dissection.....	152
II-	ChIP-seq : des pistes d'amélioration.....	153
1.	Un problème récurrent de quantité d'ADN.....	153
2.	La ChIP pour étudier le complexe dREAM.....	156
III-	RNA-seq : de la purification d'ARN à la couverture des gènes.....	156
1.	Des difficultés d'obtention des ARN.....	156
2.	Des taux de duplication trop élevés.....	157

3. Analyse bio-informatique des données de RNA-seq.....	159
<b>IV- Perspectives.....</b>	<b>162</b>
1. Identifier des gènes cibles.....	162
2. Valider ces cibles.....	162
a. Vérifier si les gènes identifiés sont de réelles cibles de <i>Rbf1</i> .....	163
b. Evaluer l'implication des cibles dans le processus apoptotique .....	163
3. Etudier le rôle du complexe dREAM.....	166
<b>V- Conclusion.....</b>	<b>167</b>

## **ANNEXES ..... 169**

Annexe I : Article « Apoptosis Quantification in Tissue: Development of a Semi-Automatic Protocol and Assessment of Critical Steps of Image Processing », de Noiron <i>et al</i> , 2021 .....	170
---	-----

Annexe II : Script d'analyses de données RNA-seq, du contrôle qualité au comptage des reads pour l'échantillon CD8_1. ....	189
--	-----

Annexe III : Liste des gènes identifiés comme activés sans retrait des duplicats.....	191
---	-----

Annexe IV : Liste des gènes identifiés comme réprimés sans retrait des duplicats.....	200
---	-----

Annexe V : Liste des gènes identifiés comme activés avec retrait des duplicats.....	201
---	-----

Annexe VI : Liste des gènes identifiés comme réprimés avec retrait des duplicats.....	207
---	-----

## **BIBLIOGRAPHIE..... 208**

# LISTE DES ABBREVIATIONS

<b>A</b>	ANT	Adenine Nucleotide Translocator
	AO	Acridine Orange
	Apaf-1	Apoptotic peptidase activating factor 1
<b>B</b>	Bcl-2	B-cell leukemia (or lymphoma) gene number 2
	BH	Bcl-2 Homology
	BIR	Baculoviral IAP Repeat
<b>C</b>	CAD	Caspase-Activated DNase
	CARD	Caspase Activation and Recruitment Domain
	Ced	Cell death
	ChIP	Chromatin Immunoprecipitation
	ChIP-seq	Immunoprécipitation de chromatine suivie de séquençage
	CHR	Cell-cycle Homology Region
<b>D</b>	DED	Death Effector Domain
	DIAP	Death-associated Inhibitor of APoptosis
	DREF	DNA replication-related element binding factor
<b>E</b>	egl-1	Egg laying 1
<b>F</b>	FNL	Fraction Non Liée
<b>G</b>	GFP	Green Fluorescent Protein
	GMR	Glass Multiple Reporter
	GO	Gene Ontology
<b>H</b>	HAT	Histone AcétylTransférase
	HDAC	Histone DésACétylase
<b>I</b>	IAP	Inhibitor of Apoptosis Protein
	IGB	Integrated Genome Browser
	IGV	Integrative Genomics Viewer
	IP	Immunoprécipitation
<b>J</b>	JNK	Jun-Kinase

<b>L</b>	log2FC	Log <sub>2</sub> Fold Change
<b>M</b>	Mip	Myb-interacting protein
<b>N</b>	NGS	Next-Generation Sequencing
	NuRD	Nucleosomal and Remodeling Deacetylase
<b>P</b>	PLA	Proximity Ligation Assay
	PTPC	Permeability Transition Pore Complex
<b>R</b>	RAM	Rbf associated motif
	RB1	Retinoblastoma 1
	Rbf	Retinoblastoma family
	RHG	Reaper, Hid, Grim
	RING	Really Interesting New Gene
	RNA-seq	RNA-sequencing, séquençage des ARN
	ROS	Reactive Oxygen Species, Espèces Réactives de l'Oxygène
	RT-qPCR	Reverse Transcriptase quantitative Polymerase Chain Reaction
<b>S</b>	SAHF	Senescence-Associated Heterochromatin Foci
<b>T</b>	TSS	Transcription Start Site, site d'initiation de la transcription
	TUNEL	Terminal deoxynucleotidyl transferase dUTP Nick End Labeling
<b>U</b>	UAS	Upstream Activated Sequence
	UBC	UBiquitin-Conjugating
	UMI	Unique Molecular Identifier
<b>V</b>	VDAC	Voltage-Dependant Anion Channel
	VIH	Virus d'Immunodéficience Humaine
<b>Z</b>	ZNC	Zone of Non-proliferating Cells



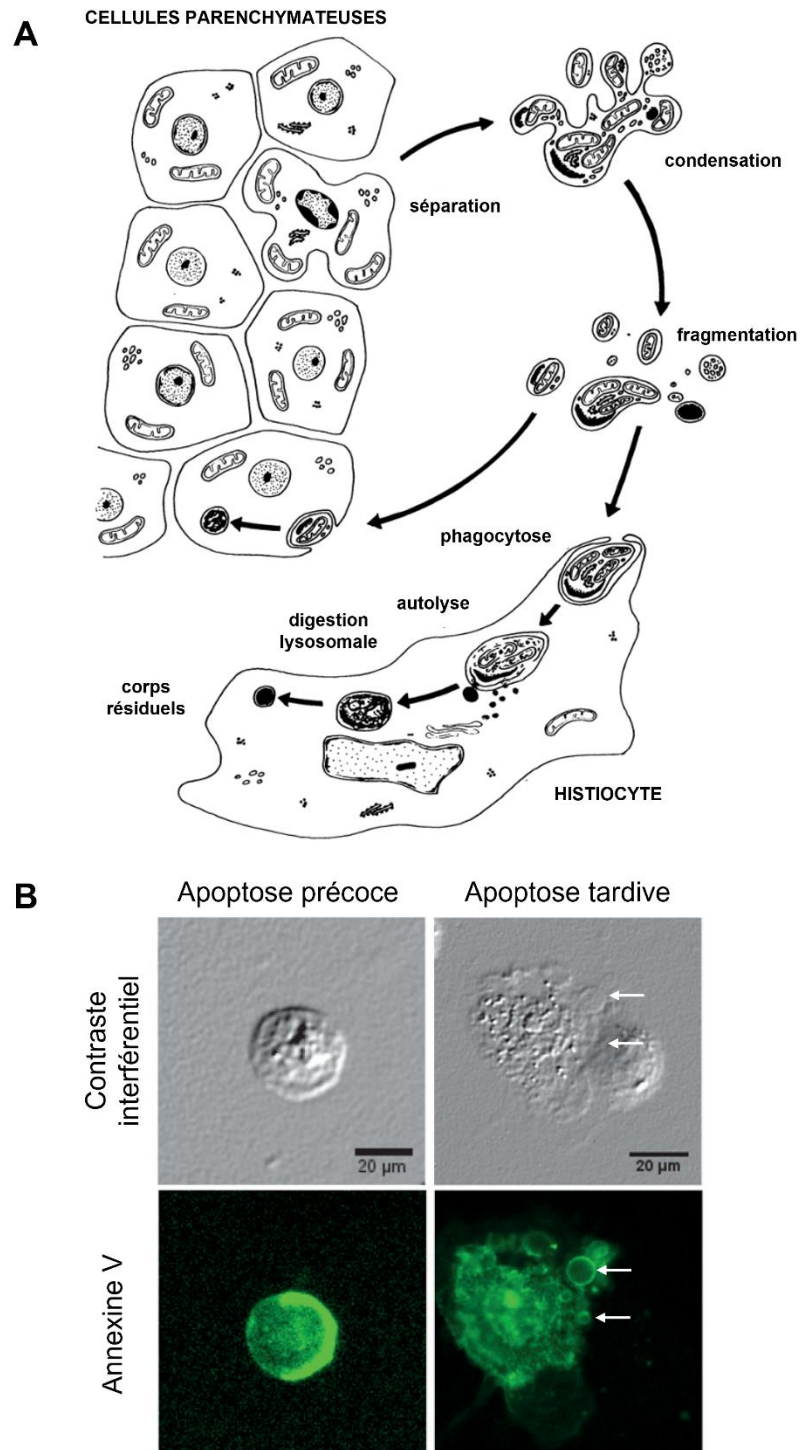
# **INTRODUCTION**

## A- L'APOPTOSE

### I- Histoire de l'étude de l'apoptose

#### 1. Découverte et morphologie

En 1972, John Kerr observe, dans des tissus de foie après interruption de l'apport sanguin d'hépatocytes, des cellules ayant une morphologie et un comportement différent des cellules en nécrose (Figure 1A). En effet, cette mort affecte des cellules isolées (et non pas des étendues cellulaires comme la nécrose). Histologiquement, il observe la formation de fragments cytoplasmiques contenant parfois des morceaux de noyaux à la chromatine fragmentée. La microscopie électronique permet de d'identifier quatre phases : la condensation du noyau et du cytoplasme, la séparation des fragments observés précédemment, puis leur phagocytose par les cellules voisines [1]. Ces observations avaient déjà été faites par des biologistes du développement [2]. En confrontant toutes leurs données, ils ont déterminé que cette mort, d'abord appelée « shrinkage necrosis » par John Kerr [3], est un mécanisme actif programmé ayant un rôle essentiel dans le développement. Ils utilisent alors le terme « apoptose », mot grec composé de « apo » (séparation) et « ptosis » (tombé) et qui désigne la perte des pétales chez les fleurs ou des feuilles chez les arbres [1]. Les facteurs d'initiation de l'apoptose sont peu connus en 1972, des horloges intrinsèques spécifiques du type cellulaire, un contrôle hormonal, des activateurs et des inhibiteurs [1] ou encore des stimuli internes (radicaux libres, infection virale) ou externes (radiations, hyperthermie) [4] sont évoqués.



**Figure 1 : Changements morphologiques caractéristiques de l'apoptose.** (A) Illustration des changements morphologiques de l'apoptose de John F. R. Kerr. *Traduit de Kerr et al., 1972.* (B) Le marquage Annexine V permet la détection de cellules en apoptose par liaison à la phosphatidyl sérine, qui est relocalisée à la membrane externe lors du processus apoptotique. Lors de l'apoptose précoce, les cellules PC-3 (lignée de cellules de cancer de la prostate) prennent une forme ronde condensée caractéristique. Plus tardivement, on peut observer les corps apoptotiques qui se détachent de la cellule (flèches blanches). *Adapté de Balvan et al., 2015.*



Si les travaux de John Kerr sont fondateurs dans l'étude moderne de l'apoptose [5], il ne décrit en réalité pas les premières observations de mort cellulaire programmée [6]. En effet, ce concept n'est pas récent puisqu'il était déjà évoqué en 1858 dans le Virchow's Cellular Pathology sous le terme de nécrobiose, définie comme une « mort lente » ou la « mort des tissus au sein d'un organisme vivant ». L'identification en 1877 de la « coagulation necrosis », une forme de nécrose des cellules de l'épithélium trachéo-bronchial retrouvée dans la diphtérie a ouvert la voie à de nombreuses recherches sur la mort cellulaire à partir des années 1880. De nombreux termes émergent alors : pycnose, caryolyse, margination de la chromatine, ... En 1884, l'apparition des colorants cellulaires a permis à Walther Flemming d'introduire le terme « mort cellulaire spontanée » : en étudiant des follicules ovariens chez les mammifères, il s'aperçoit que de nombreux noyaux sont en fragmentation au niveau de l'épithélium. Ces cellules présentent les caractéristiques classiques de l'apoptose connue aujourd'hui, auxquelles il donne le nom de « chromatolyse ». L'observation de ces phénomènes devient très populaire et de nombreux articles témoignant de cette morphologie émergent dans les années qui suivent [6].

En 1914, la quantité de données générées permet à Ludwig Gräper, un scientifique allemand, de publier un article dans lequel il émet l'idée que ce phénomène de chromatolyse sert à contrebalancer la mitose. Pour le prouver, il travaille sur la vésicule vitelline dont la taille diminue au cours du développement. Il pose alors les hypothèses suivantes : cette diminution est due soit à une diminution de la taille des cellules, soit à une diminution de leur nombre par élimination de certaines cellules. C'est cette deuxième hypothèse qu'il valide en observant de la mort par chromatolyse. Malheureusement, cette étude passe inaperçue, probablement en raison du contexte historique, et le terme « chromatolyse » est repris par les neuropathologistes pour désigner des cassures dans les granules de Nissl. Cependant, il reste populaire parmi les embryologistes qui observent des phénomènes correspondant parfaitement aux changements morphologiques typiques de l'apoptose [7].

A la fin des années 1950, la découverte des lysosomes entraîne l'apparition du concept de « suicide cellulaire » qui propose que certaines cellules soient tuées « de l'intérieur », par opposition à la nécrose induite par un traumatisme extérieur [6]. Les recherches à ce sujet permettent dans les années 1960 d'identifier un nouveau mécanisme de suicide cellulaire avec libération de radicaux libres. [8]. C'est ici que nous revenons à John Kerr et à l'apparition du terme « apoptose ».

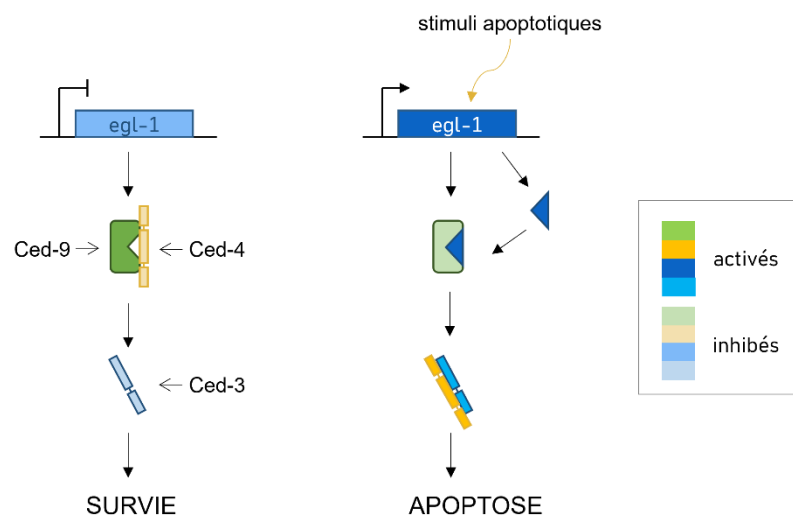
## 2. Découverte des mécanismes moléculaires



La découverte des mécanismes moléculaires est amenée par les chercheurs travaillant sur le nématode *Caenorhabditis elegans*. Lors du développement de l'hermaphrodite, 1090 cellules somatiques sont générées. Parmi elles, 131 sont destinées à mourir [9,10]. Cette mort est dite programmée car elle a lieu toujours dans les mêmes

cellules, avec la même séquence de changements morphologiques correspondants à ceux observés par John Kerr (condensation de la cellule et de son noyau, phagocytose par les cellules voisines ; Figure 1) [1,10].

L'étude de mutations entraînant une modification du nombre de cellules apoptotiques observées au cours du développement permet d'identifier la famille des gènes *ced* (cell death) : *ced-3*, *ced-4* et *ced-9* [11]. Grâce à des mutants et des tests d'épistasie, leur rôle pro- ou anti- apoptotique, ainsi que leur position dans la cascade génique sont déterminés [12]. A l'aide d'expérience de biochimie complémentaires, la cascade suivante a été établie : en absence d'un signal de mort, la protéine Ced-4, pro-apoptotique, est séquestrée par la protéine Ced-9, membre anti-apoptotique de la famille Bcl-2. Lorsqu'un signal de mort survient, le gène *egl-1* (*e*gg *l*aying) est exprimé et la protéine Egl-1 qui en résulte, également membre de la famille Bcl-2, inhibe Ced-9. Ainsi, la libération de Ced-4 lui permet d'interagir avec Ced-3 et de l'activer. Ced-3 apparaît comme l'effecteur final de l'apoptose chez *C. elegans* [11] (Figure 2).



**Figure 2 : Cascade protéique menant à l'apoptose chez le nématode *C. elegans*.** Le programme d'apoptose repose sur les protéines de la famille Bcl-2 Egl-1 (pro-apoptotique) et Ced-9 (anti-apoptotique), l'activateur de caspase Ced-4 et la caspase Ced-3). Sans stimuli apoptotique, l'anti-apoptotique Ced-9 séquestre l'activateur de caspase Ced-4, qui ne peut alors pas activer la caspase Ced-3. A la réception d'un stimulus, le gène *egl-1* est activé et le pro-apoptotique Egl-1, en inactivant Ced-9, va permettre la libération de Ced-4 qui va activer Ced-3 et déclencher l'apoptose.

En 1994, Hengartner et Horvitz rapprochent *ced-9* du proto-oncogène mammalien *bcl-2*, car leurs fonctions semblent similaires. En effet, la comparaison de leurs séquences et de leurs structures, ainsi que le sauvetage d'un phénotype *ced-9* mutant par la surexpression de *bcl-2* humain permet de conclure que *ced-9* et *bcl-2* sont des homologues fonctionnels [13,14]. Par la suite, d'autres homologues sont découverts parmi les acteurs de l'apoptose, ce qui permet de mettre en évidence des familles de protéines conservées ainsi que leur rôle dans les différentes voies d'apoptose. On peut citer

en exemple notoire la famille des caspases, effecteurs de l'apoptose dont le représentant chez le nématode *C. elegans* n'est autre que Ced-3 [15].

Cependant, si de nombreux progrès ont été faits dans la description puis la compréhension des mécanismes gouvernant la mort cellulaire programmée depuis les observations de Virchow en 1858, l'étude de l'apoptose est loin d'être terminée. En effet, si la cascade apoptotique semble relativement simple chez *C. elegans*, le mécanisme d'apoptose est loin de l'être. On distingue deux voies : intrinsèque, dite également « mitochondriale », et extrinsèque, dite « des récepteurs de mort ». L'existence de ces deux voies, leurs interactions, ainsi que l'existence de multiples homologues chez les mammifères et leurs redondances, fait de l'apoptose un mécanisme présentant encore bien des zones d'ombre. Pour faciliter son étude, d'autres organismes modèles ont émergé, comme la drosophile *Drosophila melanogaster*. Ce modèle permet d'appréhender des systèmes protéiques plus complexes que ceux de *C. elegans*, permettant une approche plus complète des éventuelles interactions protéiques et des mécanismes de la cascade, tout en s'affranchissant de la redondance complexe qu'on peut trouver chez les mammifères.

### 3. Le modèle drosophile dans l'étude de l'apoptose



*Drosophila melanogaster*, aussi appelée « mouche du vinaigre », est un diptère bien connu dans les foyers pour coloniser les fruits bien mûrs. La drosophile est également un organisme modèle utilisé depuis plus de cent ans. En effet, son cycle de vie court (10 jours, Figure 3), la capacité des femelles à pondre plusieurs centaines d'œufs et l'absence de recombinaison génétique chez les mâles en font un modèle d'étude idéal pour la génétique. De plus, l'élevage de ces organismes est facile et relativement peu coûteux [16].

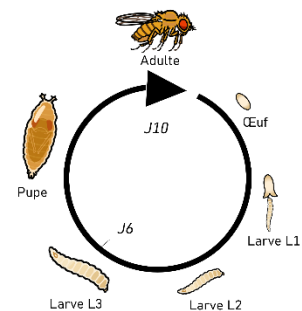
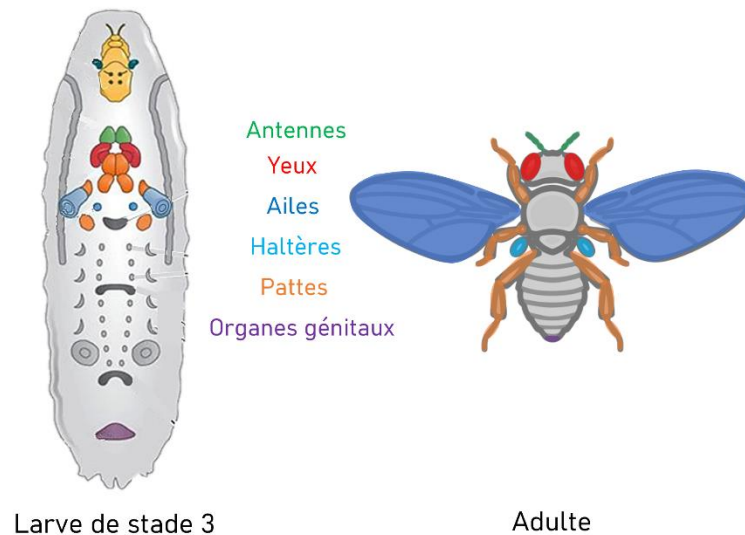


Figure 3 : Cycle de développement de la drosophile à 25°C.

C'est en 1910 que Thomas Hunt Morgan, un généticien américain, obtient le premier mutant *white* (aux yeux blancs) dont la mutation est transmise à la descendance. Grâce à celui-ci, il confirme les lois de l'hérédité de Mendel [17]. Depuis, la drosophile est devenue un modèle d'étude *in vivo* phare, très utilisé pour les études génétiques mais également dans de nombreuses autres disciplines de la biologie : développement, biologie cellulaire et moléculaire, ...

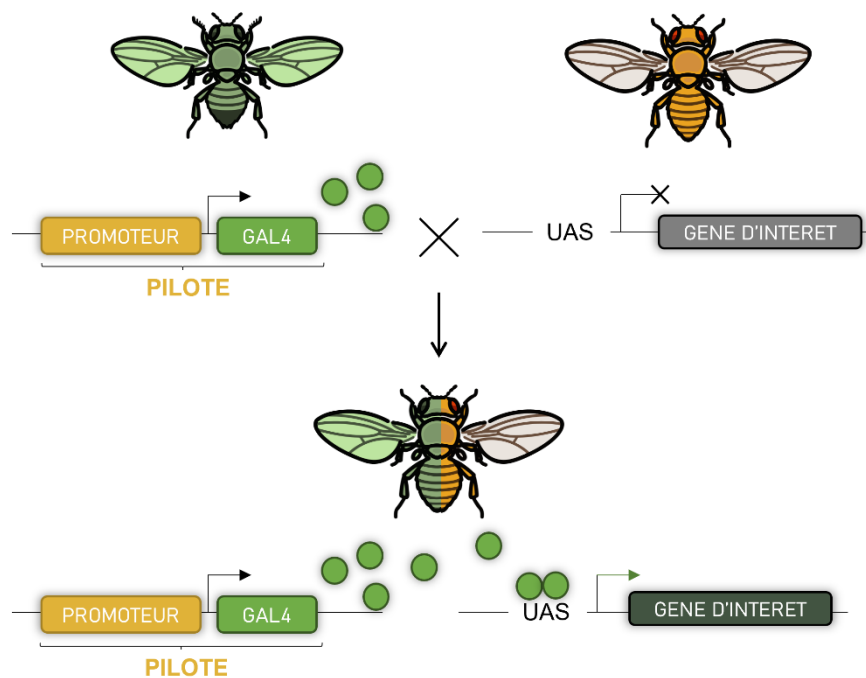
Le séquençage complet du génome de la drosophile a été réalisé en 2000, celui-ci contient 180 millions de bases, dont environ un tiers d'hétérochromatine. Les 120 millions de bases d'euchromatine restantes sont réparties majoritairement sur deux grands autosomes (chromosomes 2 et 3) et le chromosome X, le quatrième chromosome ne contenant qu'environ 1 million de bases d'euchromatine. On compte environ 13 000 gènes codants pour des protéines [18]. L'étude de ce

génomique a permis de mettre en évidence un nouvel avantage du modèle drosophile : sa grande conservation avec le modèle mammifère, qui s'accompagne d'une moindre redondance, rendant plus aisée la compréhension des mécanismes biologiques. En effet, une étude menée en 2001, comparant 929 gènes impliqués dans des maladies humaines au génome de la drosophile, révèle que 714 d'entre eux, soit 77 %, sont des homologues de 548 gènes de drosophile [19].



**Figure 4 : Correspondance entre disques imaginaux de la larve L3 et organes externes chez l'adulte.** Les organes correspondants aux disques imaginaux sont visualisés par la même couleur. Adapté de <http://flirtyfwies.blogspot.com/2007/10/life-cycle-of-drosophila.html>.

Un grand nombre de maladies étant associé à une dérégulation du processus apoptotique, *Drosophila melanogaster* devient, à la suite de *C. elegans*, un modèle de choix pour l'étude de l'apoptose. En effet, de l'apoptose développementale est observée chez la drosophile lors des stades d'embryogenèse et de métamorphose. Notamment, les « disques imaginaux » sont particulièrement sujets à cette apoptose. Ce sont des structures larvaires internes précurseurs des organes externes chez l'adulte, comme les ailes, les pattes ou encore les yeux (Figure 4). Le système UAS-GAL4, un outil génétique très puissant, permet l'expression contrôlée dans le temps et l'espace d'un gène d'intérêt. Ainsi, il est possible de surexprimer un gène dans un tissu particulier, ou à un stade de développement particulier de l'individu [20] (Figure 5). Cela se révèle indispensable dans l'étude de l'apoptose car une induction ubiquitaire de mort cellulaire entraînerait la mort de l'individu entier. Grâce à ce système et à tous les outils classiques de biologie (RNAi, protéines étiquetées, ...), il est possible d'induire et d'étudier la mort cellulaire dans des organes non essentiels à la vie de l'individu, comme les ailes, et ainsi d'observer des phénotypes aux différents stades de la vie.



**Figure 5 : Système UAS-GAL4.** A 25°C, une lignée portant le gène GAL4 sous contrôle d'un pilote permettant une expression précise dans l'espace et le temps est croisée avec une lignée portant un gène d'intérêt sous contrôle d'une séquence UAS (Upstream Activated Sequence). Chez la descendance, le produit GAL4 est exprimé dans le domaine et la fenêtre d'expression du pilote. Dans les tissus ou durant les stades concernés, le GAL4 se fixe sur les séquences UAS, activant la transcription du gène d'intérêt.

## II- Rôles et mécanismes de l'apoptose

### 1. L'apoptose : un mécanisme physiologique indispensable

L'apoptose est un mécanisme indispensable à la formation et au maintien des individus, au même titre que la prolifération ou la différenciation. En effet, elle a un rôle complémentaire mais opposé à la mitose, permettant ainsi le maintien de l'homéostasie des organismes [4,21,22].

Durant le développement, l'apoptose est indispensable à la genèse de nombreux organes et systèmes, dans quasiment tous les organismes. On peut par exemple citer l'élimination des cellules interdigitales pour la formation des doigts chez les mammifères [22] (Figure 6) ou le développement du système nerveux durant lequel les cellules n'établissant pas de synapses fonctionnelles meurent [23].

Le rôle essentiel de l'apoptose ne s'arrête pas à la naissance de l'individu. En effet, à l'âge adulte, elle est également nécessaire à de nombreux processus comme le renouvellement cellulaire pour éliminer les cellules vieillissantes ou endommagées, la dégénérescence des follicules ovariens lors du cycle menstruel [24] ou encore le développement et le fonctionnement du système immunitaire [25].

On estime que chaque jour, 10 milliards de cellules sont créées pour compenser la perte cellulaire par apoptose [26].

Toutefois, le mantra de Peter Parker, « un grand pouvoir implique de grandes responsabilités », s'applique à cette situation. En effet, compte tenu de son rôle essentiel dans de nombreux processus vitaux, les dérégulations de l'apoptose ont bien souvent pour conséquences de nombreuses pathologies. On peut distinguer deux groupes de pathologies liées aux dérégulations de l'apoptose : celles impliquant une augmentation de l'apoptose, et celles impliquant une diminution de l'apoptose.

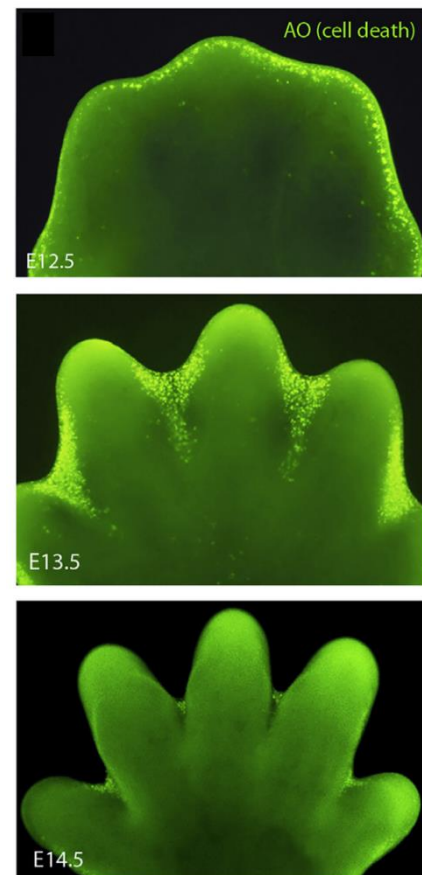
Lorsque la balance apoptose/prolifération penche du côté de l'apoptose, il y a un excès d'apoptose, et donc une perte de cellules trop importante. Cela peut être lié à une activation indésirable d'un pro-apoptotique ou à l'inactivation d'un anti-apoptotique. Dans cette catégorie on peut citer par exemple les maladies neuro-dégénératives comme les maladies de Parkinson ou d'Alzheimer [27] ou l'infection au virus d'immunodéficience humaine (VIH) [28].

A l'inverse, lorsque la balance penche du côté prolifération, il y a un défaut d'apoptose et donc des cellules surnuméraires. Cela peut être lié à l'activation d'un anti-apoptotique ou l'inactivation d'un pro-apoptotique. Le cancer est la pathologie la plus connue de ce groupe : lors de la survenue d'un cancer, les cellules cancéreuses ne meurent pas et prolifèrent de manière incontrôlée, créant ainsi des tumeurs [29]. Les maladies auto-immunes surviennent également sur ce modèle, lorsque des cellules du système immunitaires ciblant les marqueurs du soi ne sont pas éliminées [30].

## 2. Un mécanisme précis et régulé

### a. Les acteurs de l'apoptose

Depuis l'identification du mécanisme à quatre acteurs de l'apoptose chez *C. elegans*, des milliers d'études sur l'apoptose ont été réalisées, et environ 20 000 autres sont publiées chaque année. Ainsi, un grand nombre de mécanismes, régulations et protéines ont pu être identifiés, chez de nombreux



**Figure 6 : Implication de l'apoptose dans le développement des doigts chez la souris.** Un marquage acridine orange (AO) a été réalisé sur des pattes d'embryon de souris à différents stades de développement entre 12,5 et 14,5 jours. Le marquage apoptotique se trouve particulièrement au niveau des cellules interdigitales à 13,5 jour : celles-ci sont éliminées pour donner la forme des doigts. Adapté de Wood and Martin, 2017.

organismes modèles. On retrouve également des mécanismes pouvant être apparentés à de l'apoptose dans la mort cellulaire programmée chez les plantes. Si l'existence de l'apoptose reste un point de discordance chez les biologistes végétaux, on trouve des protéines caspases-like et des protéines ayant une structure similaire aux Bcl-2 ; ainsi que les étapes de l'apoptose animale comme la condensation de la chromatine ou la formation de corps apoptotiques [31].

En réalité, si les quatre acteurs du modèle nématode sont une base solide des processus apoptotiques des mammifères et de la drosophile, ils ne constituent qu'un niveau de complexité très faible. En effet, bien plus d'acteurs existent chez ces organismes. Pour une meilleure efficacité et une réponse plus rapide aux stimuli apoptotiques, certains de ces acteurs sont déjà présents dans la cellule, comme ced-4 et ced-3, et n'attendent que leur activation pour enclencher la mort de la cellule. C'est le cas par exemple des caspases et des protéines pro-apoptotiques de la famille Bcl-2, qui sont maintenues inactives par les inhibiteurs de caspases et les membres anti-apoptotiques de la famille Bcl-2, respectivement.

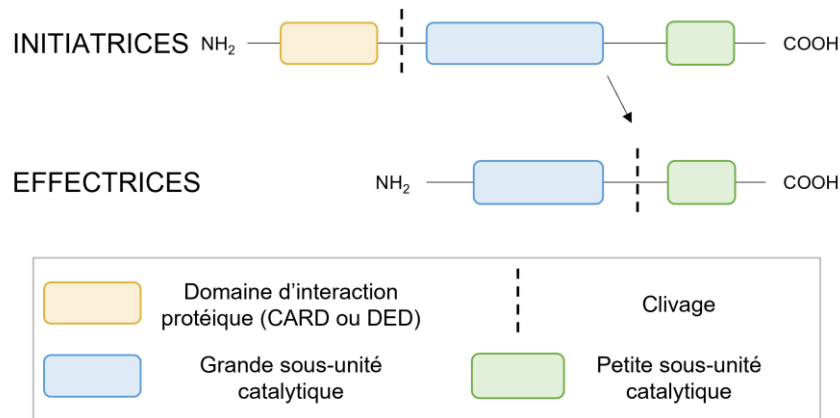
### Les caspases

Dans les années 80, Ced-3 est identifiée comme étant l'effecteur final de l'apoptose [32]. Au fil des années, d'autres protéines de même type, des protéases dépendantes des cystéines capables de cliver après un résidu aspartate et d'induire l'apoptose, ont été découvertes dans d'autres organismes. Dans un souci de clarté, la communauté scientifique s'est entendue en 1996 pour désigner ces protéines sous le terme « caspase », pour cystéine protéase clivant après un acide aspartique [33]. Lors du processus apoptotique, deux types de caspases sont nécessaires : les effectrices et les initiatrices.



Les caspases initiatrices activent les caspases effectrices par clivage. Chez les mammifères, les caspases initiatrices sont au nombre de trois : les caspases 8, 9 et 10. La 8 et la 10 interviennent lors de la voie extrinsèque de l'apoptose (ou voie des récepteurs de mort). La caspase 9 est la caspase initiatrice de la voie intrinsèque ou mitochondriale [34].

Les caspases initiatrices sont produites sous forme de pro-caspases activables par clivage. Grâce à un domaine d'interaction CARD (Caspase Activation and Recruitment Domain) ou DED (Death Effector Domain), elles s'assemblent au sein d'un complexe qui permet leur auto-activation. Dans le cas de la voie mitochondriale de l'apoptose, ce complexe est appelé « apoptosome ». Il contient sept protéines Apaf-1 associées au cytochrome c. L'apoptosome permet une auto-activation de la caspase 9 [35,36]. Une fois activées, les caspases initiatrices se fixent aux caspases effectrices et les clivent, ce qui, à leur tour, permet leur activation [37] (Figure 7).



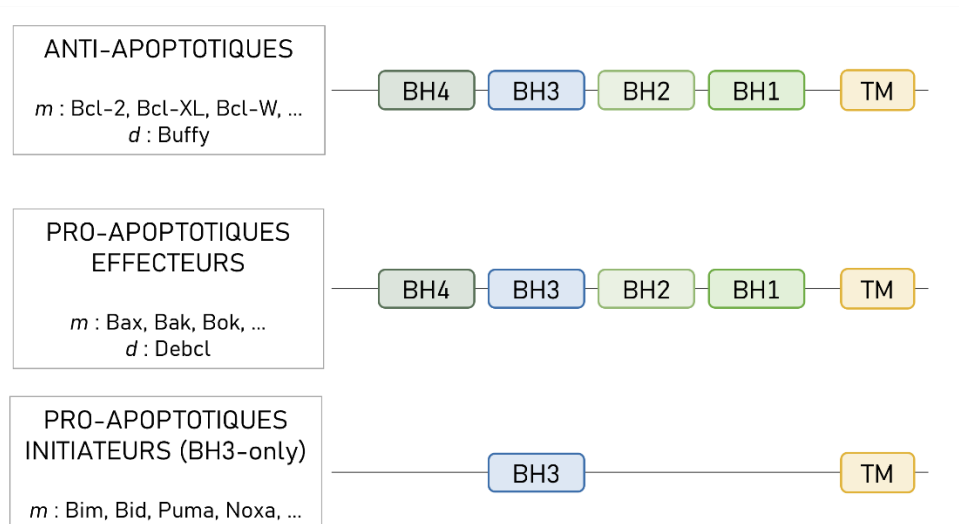
**Figure 7 : Domaines fonctionnels des caspases.** Pour être activées, les procaspases initiatrices s'oligomérisent via leurs domaines d'interaction protéique et sont clivées. Elles peuvent alors aller activer les caspases effectrices par clivage entre la grande et la petite sous-unité catalytique.

Les caspases effectrices, après activation, clivent de nombreux composants cellulaires, responsables des changements morphologiques caractéristiques de l'apoptose (Figure 1B) [38]. Parmi leurs cibles, on peut citer les lamines nucléaires, dont le clivage est impliqué dans la condensation de la chromatine et du noyau, des protéines du cytosquelette (actine, gelsoline, ...) ou encore de ICAD, l'inhibiteur de la DNase CAD (Caspase-Activated DNase), ce qui induit la fragmentation de l'ADN [37]. Chez *C.elegans*, Ced-3 joue à la fois le rôle de caspase initiatrice et effectrice : elle s'auto-active par clivage. Cette activation est contrôlée par Ced-4, l'activateur de caspases [32]. Chez les mammifères, on trouve trois caspases effectrices : les caspases 3, 6 et 7 [34]. La 3 et la 7 restent inactives jusqu'à ce que les caspases initiatrices activées coupent le linker entre les petites et les grandes sous-unités catalytiques [37]. La caspase 6, quant à elle, semble plutôt activée soit par clivage par la caspase 3 [39], soit par auto-clivage [40].

### Les protéines de la famille Bcl-2

Les protéines de la famille Bcl-2 ont un rôle de régulation de l'apoptose. Le clonage d'une région chromosomique dont la translocation est associée à l'apparition de lymphomes folliculaires a permis la découverte de l'oncoprotéine Bcl-2 (B-cell leukemia -or lymphoma- gene number 2). Celle-ci a ensuite donné son nom à la famille de protéines [41]. Celle-ci compte seize membres chez les mammifères et est caractérisée par la présence d'un ou plusieurs domaines appelés domaines BH (Bcl-2 Homology). Selon le nombre de domaines BH et le rôle de la protéine vis-à-vis de l'apoptose, on distingue trois sous-familles dans la famille Bcl-2 (Figure 8).





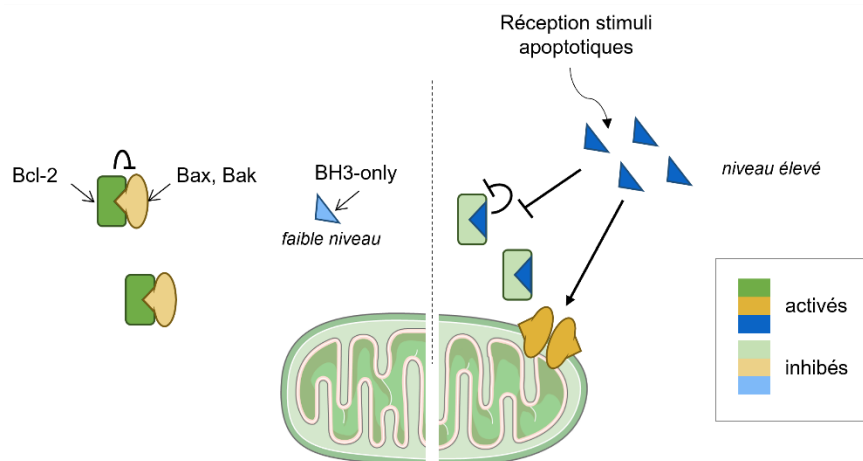
**Figure 8 : Domaines des protéines de la famille Bcl-2.** Les protéines de la famille Bcl-2 sont caractérisées par la présence de 1 à 4 domaines BH et d'un domaine transmembranaire (TM). Elles sont divisées en trois sous-familles : les anti-apoptotiques, les pro-apoptotiques multi-BH, et les pro-apoptotiques « BH3-only ». Ces régulateurs sont absents de la cascade chez le nématode *C. elegans*, mais présents chez la drosophile (d) et chez les mammifères (m). Adapté de Delbridge et al., 2016.

*Les anti-apoptotiques.* Les anti-apoptotiques de la famille Bcl-2 contiennent de multiples domaines BH. On y trouve notamment Bcl-2, dont le fonctionnement ressemble à celui de Ced-9 chez le nématode [13]. Ces protéines se lient aux protéines pro-apoptotiques de la même famille et les séquestrent pour empêcher leur activité [42], notamment de relargage du cytochrome c [43] (Figure 9).

*Les pro-apoptotiques effecteurs.* Certains pro-apoptotiques de la famille Bcl-2 contiennent également de multiples domaines BH. Leurs représentants les plus connus sont Bax et Bak, qui n'ont pas d'homologues chez *C. elegans*. Lorsqu'ils sont libérés des Bcl-2 anti-apoptotiques, les pro-apoptotiques multi-BH subissent un changement de conformation et sont relocalisés à la mitochondrie [44]. Nous avons choisi de représenter quatre domaines BH pour les pro-apoptotiques effecteurs sur la Figure 8, mais la présence d'un domaine BH4 chez les membres de cette sous-famille ne fait pas consensus dans la communauté scientifique. L'interaction des pro-apoptotiques multi-BH avec les lipides de la membrane mitochondriale, et notamment la cardiolipine, semble indispensable ; mais il n'est pas clair si leur activation est induite par cette interaction ou si elle est antérieure. Malgré la capacité des protéines de la famille Bcl-2 à former des pores sur des membranes synthétiques, le rôle des protéines Bax et Bak à la mitochondrie n'est pas élucidé [45]. La formation de pores dans la membrane mitochondriale externe, permettant le relargage de facteurs apoptogéniques comme le cytochrome c, reste l'hypothèse la plus répandue [43] (Figure 9). p53, principalement décrit pour son rôle de suppresseurs de tumeurs, possède un rôle pro-apoptotique : il est capable de stimuler l'apoptose par différents mécanismes impliquant notamment les protéines de la famille Bcl-2. En effet, p53 active la transcription des gènes pro-apoptotiques de la famille Bcl-2 et réprime la transcription de gènes anti-apoptotiques comme *bcl-2*. Il pourrait également jouer un rôle dans la libération des Bcl-2

pro-apoptotiques, ainsi que dans la perméabilisation de la membrane mitochondriale en activant Bax et Bak [46,47]. Cette activation est considérée comme le point de non-retour du processus apoptotique [41].

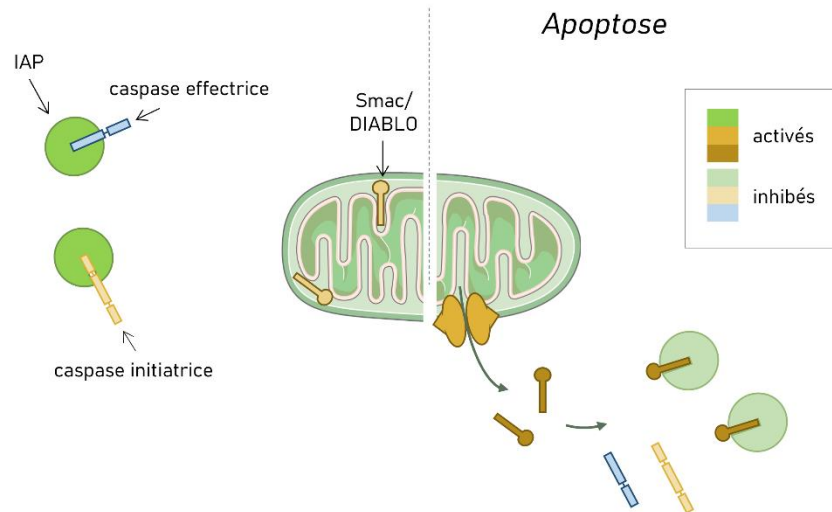
*Les BH3-only.* Les BH3-only, qui ne possèdent qu'un seul domaine BH, sont également des pro-apoptotiques. Chez les mammifères, les plus connus sont Bim et Puma. Elles jouent un rôle notamment lors de l'initiation de l'apoptose. En effet, lors de la réception de stimuli apoptotiques, les BH3-only sont stimulées ou parfois présentes en plus grande quantité [48] (Figure 9). Cette stimulation peut être transcriptionnelles ou post-transcriptionnelles, on sait par exemple que Puma et Noxa sont des cibles transcriptionnelles de p53 [49,50]. Les BH3-only déclenchent l'apoptose par deux mécanismes. D'une part, elles ont la capacité de se fixer sur les protéines Bcl-2 anti-apoptotiques, libérant, par effet de titration, les pro-apoptotiques Bax et Bak et leur permettant de se relocaliser à la mitochondrie. D'autre part, elles peuvent également activer directement les pro-apoptotiques Bax et Bak [51] (Figure 9).



**Figure 9 : Modèle d'activation et d'interactions des protéines Bcl-2.** En conditions physiologiques (gauche), les protéines pro-apoptotiques (Bax, Bak) sont maintenues inactives par interaction avec les protéines anti-apoptotiques (Bcl-2). Les BH3-only sont à un faible niveau. Lors de la réception de signaux apoptotiques, les BH3-only sont stimulées et leur quantité ou leur activité. En se fixant sur les Bcl-2, elles libèrent les pro-apoptotiques qui peuvent alors s'oligomériser pour former des pores à la membrane mitochondriale.

### Les inhibiteurs de caspases (IAPs) et leurs inhibiteurs

Les IAPs (Inhibitor of Apoptosis Protein) sont également des régulateurs de l'apoptose. C'est une famille de protéines anti-apoptotiques caractérisée par la présence d'un domaine BIR (baculoviral IAP repeat), nécessaire et suffisant pour leur activité anti-apoptotique [52]. Neuf IAPs ont pu être identifiées chez les humains. Parmi elles, certaines possèdent également des domaines CARD, RING (Really Interesting New Gene) et/ou UBC (UBiquitin-Conjugating) leur permettant de se lier aux caspases (Figure 10), voire d'exercer une activité ubiquitine-ligase pour permettre leur dégradation [53]. Les IAPs agissent en aval de Bax et Bak, les Bcl-2 pro-apoptotiques multi-BH [54].



**Figure 10 : Mode d'action des IAPs et inhibiteurs des IAPs.** En conditions physiologiques (gauche), les caspases initiatrices et effectrices sont maintenues inactives par interaction avec les IAPs. Les inhibiteurs des IAPs (Smac/DIABLO et Omi/HtrA2) sont séquestrés dans l'espace intermembranaire de la mitochondrie. Lorsque le processus apoptotique est engagé, les inhibiteurs des IAPs sont libérés de la mitochondrie via les pores impliquant Bax et Bak. Ils peuvent alors interagir avec les IAPs et libérer les caspases. Celles-ci seront ensuite prises en charge par l'apoptosome et/ou activées par clivage.

Hors conditions apoptotiques, les inhibiteurs des IAPs sont localisés dans l'espace intermembranaire mitochondrial. Lors de la réception de stimuli apoptotiques, ces inhibiteurs sont relargués dans le cytosol, permettant de faire basculer l'équilibre vers la réalisation de l'apoptose. Chez les mammifères, deux inhibiteurs des IAPs ont été identifiés : Smac/Diablo [55,56] et Omi/HtrA2 [57]. Lors de la perméabilisation de la membrane par Bax et Bak, ces protéines, normalement localisées dans l'espace intermembranaire, sont relâchées dans le cytosol en même temps que le cytochrome c [56,58]. Elles vont alors se lier aux IAPs pour libérer les caspases (Figure 10) et déclencher le processus d'apoptose.

### Les acteurs de l'apoptose chez la drosophile



Chez *Drosophila melanogaster*, la recherche de mutants présentant une diminution d'apoptose développementale a permis en 1994 de lier les gènes contenus dans la délétion H99 à une activité de régulation de l'apoptose. En effet, dans les embryons de drosophile contenant cette délétion, White *et al.* ont pu observer que des cellules censées être éliminées au cours de l'embryogenèse restaient bien vivantes. Par des expériences de complémentation fonctionnelle, ils ont déterminé que cette délétion recouvrait trois gènes : *reaper* (*rpr*), *hid* et *grim*, groupées sous le nom de « RHG », codant pour des protéines pro-apoptiques [59].


Ces recherches de mutants affectant l'apoptose ont également permis de mettre en évidence des IAPs chez la drosophile, nommées DIAP1 et DIAP2 (Death-associated inhibitor of apoptosis 1/2). Ces IAPs agissent en aval des RHG [60]. Plus tard, deux autres IAPs seront identifiées : dBruce et Deterin,




mais celles-ci sont moins bien connues. Ces quatre protéines possèdent les domaines BIR caractéristiques des IAPs, DIAP1 et 2 possèdent également des domaines RING E3 ubiquitine ligase.

A cette époque, les mécanismes d'apoptose chez *C. elegans* étaient déjà bien connus. Les caspases existant chez le nématode et les mammifères, il était très probable qu'il en existe également chez la drosophile. De plus, l'apoptose induite par Rpr, Hid et Grim est sensible à l'inhibiteur de caspases p35 [61]. Cependant, l'étude de leur séquence ne permet pas de mettre en évidence des domaines caractéristiques des caspases ou même des activateurs de caspases. C'est à la suite du séquençage du génome de la drosophile en 2000 [18] que sept caspases sont identifiées : trois initiatrices (Dredd, Dronc et Strica) et quatre effectrices (drICE, Dcp-1, Decay et Damm) [62,63]. En voulant étudier des mécanismes d'activation de caspases homologues à ceux des mammifères et de *C. elegans*, Kanuka *et al.* ont observé que la surexpression de *ced-4* chez la drosophile induisait de l'apoptose. Cette observation a mené à la découverte de Dapaf-1 (aussi appelé Dark), l'homologue de la protéine Apaf-1 (Apoptotic peptidase activating factor 1) [64]. Parmi les caspases initiatrices, Dronc est la seule à avoir un domaine CARD, qui permet son interaction avec Dark. Dredd, grâce à son domaine DED, est plus impliquée dans la voie d'apoptose extrinsèque. Quant à Strica, elle reste mystérieuse car très peu étudiée, elle ne possède pas les domaines classiques des caspases initiatrices. Les caspases effectrices les plus étudiées sont Drice et Dcp-1 [65].

Comme chez les mammifères, les caspases sont inhibées par les IAPs. En particulier, DIAP1 interagit avec Dronc grâce à ses domaines BIR, la poly-ubiquitine et permet ainsi sa dégradation. Cependant, ce n'est pas la seule partenaire de DIAP1. En effet, Hid peut également se fixer sur les domaines BIR de DIAP1 et ainsi, par compétition, empêcher la fixation de la caspase et permettre son activité de clivage [66]. Les protéines RHG sont donc des inhibiteurs de IAPs.

Pour ce qui est des protéines Bcl-2, il n'en existe que deux chez *Drosophila melanogaster* : Debcl, considéré comme pro-apoptotique [67] (aussi connu sous le nom de Dbok, Drob-1 ou Dborg-1), et Buffy, en général anti-apoptotique. Comme Bcl-2 avec Bax et Bak, Buffy interagit directement avec Debcl pour le séquestrer et ainsi inhiber son activité pro-apoptotique [68]. Cependant, leurs rôles pro- et anti-apoptotiques ne sont pas aussi arrêtés que chez les mammifères. En effet, leur mode d'action semble dépendre du tissu dans lequel ils sont exprimés : dans les neurones, *debcl* protège de la dégradation des neurones induite par les protéines à polyglutamines, alors que *buffy* favorise cette neurodégénérescence [62,69].



				
Initiation	<b>BH3-only</b>	egl-1	-	Bim, Puma, Noxa, ...
	<b>Bcl-2 anti-apoptotiques</b>	ced-9	Buffy	Bcl-2, Bcl-XL, ...
	<b>Bcl-2 pro-apoptotiques</b>	-	Debcl	Bax, Bak, Bok, Bid, ...
	<b>Apoptosome</b>	-	Dark	Apaf1
Exécution	<b>Caspases initiatrices</b>	ced-4	Dredd, Dronc, Strica	Caspases 2, 8, 9, 10
	<b>Caspases effectrices</b>	ced-3	drICE, Dcp-1, Decay, Damm	Caspases 3, 6, 7
Régulation	<b>IAPs</b>	-	DIAP1, DIAP2	cIAP1, cIAP2, NIAP, ...
	<b>Inhibiteurs des IAPs</b>	-	Rpr, Hid, Grim	Smac/DIABLO, Omi/HtrA2

**Figure 11 : Multiplication des acteurs selon l'organisme.** Le modèle drosophile offre une complexité intermédiaire entre le nématode et les mammifères.

Ainsi, l'étude de l'apoptose révèle l'intérêt de la drosophile comme organisme modèle. En effet, elle offre un niveau de complexité intermédiaire entre les mécanismes mis en évidence chez le nématode *C. elegans* et les mammifères. Les mécanismes complexes, comme les régulateurs de la famille Bcl-2 ou la régulation liée aux IAPs, sont conservés, tout en limitant la redondance fonctionnelle (Figure 11).

b. L'apoptose : une cascade protéique bien huilée



En conditions physiologiques, tous les acteurs nécessaires au déroulement de l'apoptose sont présents, prêts à entrer en scène (Figure 12). Grâce aux protéines régulatrices, les protéines actrices du déroulement de l'apoptose sont maintenues inactives par les protéines anti-apoptotiques : Bax et Bak sont séquestrées par les protéines Bcl-2 [42] et les caspases par les IAPs [52], elles-mêmes protégées des inhibiteurs des IAPs piégés dans la mitochondrie [55,57].

Les stimuli apoptotiques peuvent être variés : stress cytotoxique, dommages à l'ADN trop importants, ... Lors de la réception de ces signaux, les protéines BH3-only sont activées [48] et déclenchent la libération et l'activation des pro-apoptotiques Bax et Bak, soit par interaction avec leurs inhibiteurs Bcl-2, soit par activation directe [51]. Bax et Bak sont alors relocalisés à la membrane mitochondriale externe, où ils forment des oligomères permettant la perméabilisation de la membrane

mitochondriale [70]. Des facteurs apoptogéniques sont alors relargués dans le cytosol, notamment le cytochrome c et les inhibiteurs des IAPs, Smac/Diablo et Omi/HtrA2 [4,58].

Les inhibiteurs des IAPs se fixent sur leurs cibles et libèrent ainsi les pro-caspases séquestrées [57]. Parmi elles, la pro-caspase 9, initiatrice. Via leurs domaines CARD, plusieurs pro-caspases 9 sont recrutées à l'apoptosome, un complexe multi-protéique formé de 7 à 8 protéines Apaf-1, de dATP et du cytochrome c libéré de la mitochondrie. Ce complexe permet un rapprochement des caspases entre elles, ainsi que leur activation par clivage [36].

Les caspases 9 activées sont libérées de l'apoptosome et se fixent aux procaspases effectrices pour les cliver, et ainsi les activer. Celles-ci peuvent alors aller exercer leur activité protéase sur de nombreux substrats, entraînant ainsi les changements morphologiques caractéristiques de l'apoptose (condensation de la chromatine et du noyau, fragmentation de l'ADN, formation des corps apoptotiques) et la mort de la cellule [37].



Chez la drosophile, le processus est très similaire [62] (Figure 13). Cependant, un point majeur diffère : la perméabilisation de la membrane mitochondriale externe. En effet, lors de l'initiation de l'apoptose, si les protéines Debcl sont effectivement présentes à la mitochondrie [71], il n'y a toutefois pas de libération du cytochrome c. Cette localisation mitochondriale pourrait permettre le recrutement de Drp1, l'acteur principal de la fission mitochondriale, et ainsi promouvoir la fission. Cependant, le rôle de Debcl dans l'apoptose n'est pas clair et celui-ci semble intervenir dans différentes voies d'apoptose. En effet, l'action d'un inhibiteur de caspases peut ou non inhiber la mort induite par *debcl* ; et l'activité pro-apoptotique de Debcl ne semble pas dépendante des inhibiteurs de IAPs (RHG) [72]. Ces derniers, quant à eux, sont régulés transcriptionnellement ou post-traductionnellement [73]. Leur action sur les IAPs DIAP1/2 permet la libération des caspases initiatrices Dronc, qui vont s'insérer dans un apoptosome constitué de 8 protéines Dark (Apaf-1) [64]. L'activation des caspases initiatrices au sein de l'apoptosome ne nécessite pas le cytochrome c [62,63]. Les caspases Dronc activées vont ensuite elle-mêmes activer les caspases effectrices drICE, également libérées de DIAP1/2 [74,75]. Les caspases effectrices clivent alors les composants cellulaires.

Bien qu'il n'y ait pas de libération du cytochrome c, la mitochondrie reste comme chez les mammifères le point central de la voie d'apoptose intrinsèque chez la drosophile [62]. En effet, certains acteurs, comme les protéines de la famille Bcl-2 Debcl et Buffy, y sont présents lors du déclenchement de la réponse apoptotiques. D'autres acteurs y convergent, comme par exemple les protéines RHG pour séquestrer les DIAPs et libérer les caspases Dronc et drICE. Cette localisation est indispensable à l'activité pro-apoptotique de ces inhibiteurs des IAPs puisqu'en l'absence des

domaines ou séquences d'adressage nécessaires à leur adressage à la mitochondrie, il n'y a pas de déclenchement de l'apoptose [76]. De façon commune avec l'apoptose des mammifères, la fission mitochondriale semble jouer un rôle important dans ce processus [77]. Elle est permise par la localisation ou relocalisation des protéines de l'apoptose à la mitochondrie : en effet, Debcl permet le recrutement de Drp1, l'activateur principal de la fusion ; et Rpr, qui fait partie des RHG, séquestre et inhibe la protéine inductrice de fusion Marf [78]. Cependant, le rôle exact de la fission mitochondriale est encore méconnu.

Ces cascades protéiques permettent donc une réponse rapide et efficace à des stimuli apoptotiques grâce à un système de régulation des acteurs de l'apoptose basé sur un équilibre entre des activateurs et des répresseurs post-traductionnels. Cependant, la régulation de l'apoptose peut également être réalisée directement au niveau de la transcription de ces acteurs. Parmi ces régulateurs transcriptionnels, on peut notamment trouver les protéines de la famille RB, dans un rôle bien moins connu que celui de la régulation du cycle cellulaire.

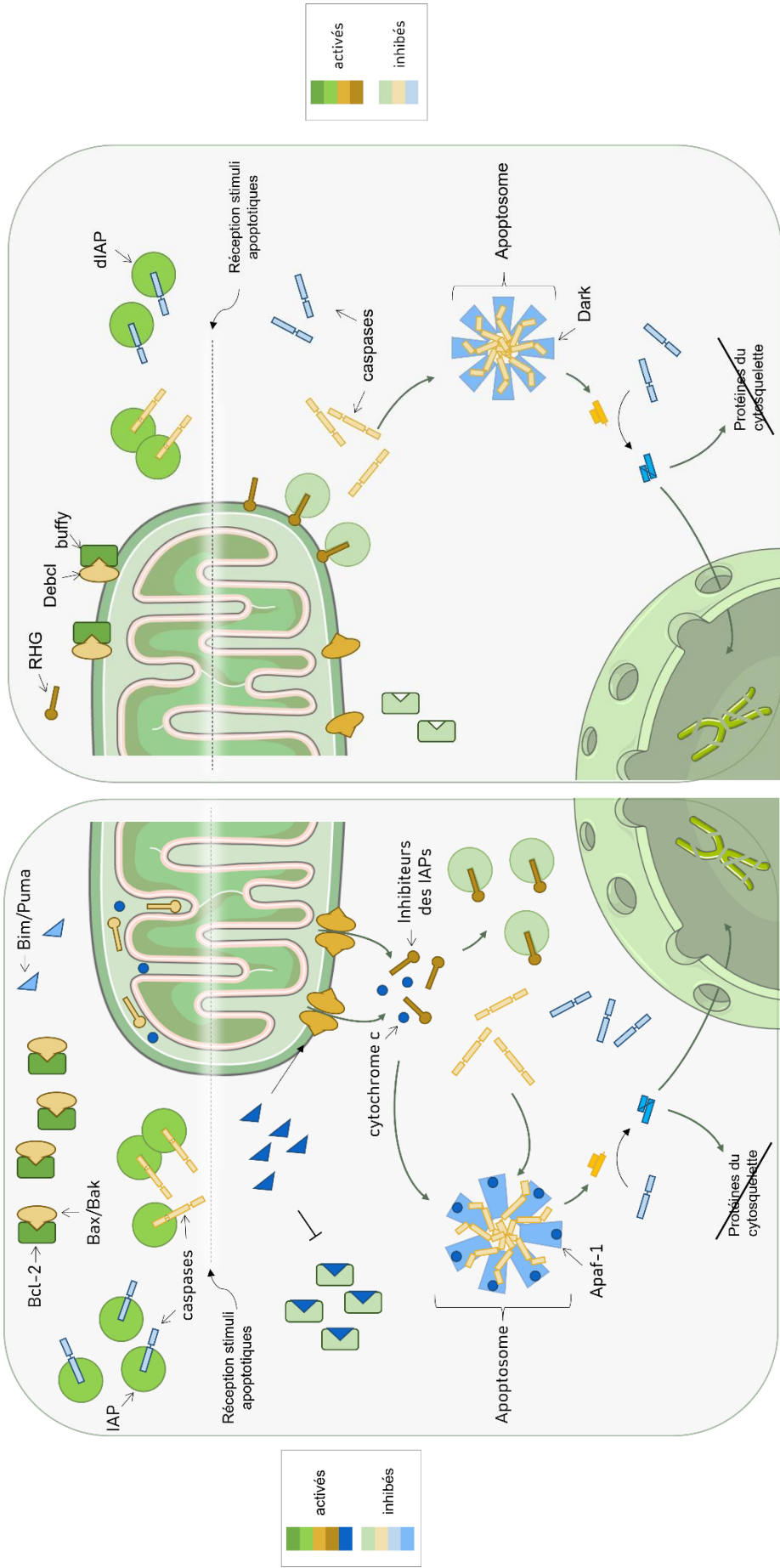


Figure 13 : Le processus apoptotique chez la drosophile

Figure 12 : Le processus apoptotique chez les mammifères



## B- LES PROTEINES DE LA FAMILLE RB

### I- Du rétinoblastome à Rb

#### 1. Du rétinoblastome...

Le rétinoblastome est une tumeur pédiatrique oculaire qui prend naissance à partir des rétinoblastes, les cellules immatures de la rétine [79]. Il a été décrit pour la première fois en 1809 [80]. Aujourd'hui, il touche un enfant sur 15 à 20 000 naissances, ce qui représente 9000 nouveaux cas par an dans le monde [81] et 50 en France [82]. Cette maladie a une importance majeure en pédiatrie, puisqu'elle représente 4 % de tous les cancers pédiatriques. C'est la tumeur intraoculaire la plus fréquente chez l'enfant [79,83]. On en distingue deux formes : le rétinoblastome unilatéral, lorsqu'un seul œil est touché, et le bilatéral, lorsque les deux yeux sont touchés.

L'âge moyen de diagnostic est de 18 mois pour le bilatéral, un peu plus tard (24 mois) pour l'unilatéral. Le signe clinique principal est la leucocorie, c'est-à-dire un reflet blanc observé à travers la pupille, sur une photo sans flash par exemple. 95 % des enfants sont diagnostiqués avant l'âge de 5 ans [83].

Les traitements reposent sur de la chimiothérapie, de la radiothérapie et parfois une énucléation selon le stade et l'avancement de la tumeur. Si le rétinoblastome est traité à temps, les chances de survie sont très bonnes (95 à 98 %), ainsi que le sauvetage de l'œil (90 %) et de la vision (80%) [83]. En revanche, un rétinoblastome non traité entraîne une perte progressive de la vue et un risque de colonisation de l'orbite, du nerf optique et du cerveau qui engage le pronostic vital [81,83]. Il peut également y avoir des métastases dans le reste du corps [84].

Aucune prédisposition de genre ou d'ethnie n'a pu être mise en évidence. En revanche, l'apparition plus précoce des rétinoblastomes bilatéraux par rapport aux unilatéraux a conduit Alfred G. Knudson Jr. sur la piste génétique avec sa « 2-hit hypothesis » (hypothèse en deux coups). En étudiant les dossiers de 48 cas survenus entre 1944 et 1969, il observe que 25 à 30 % de ces cas sont bilatéraux et tous héréditaires et dans les 70 à 75 % restants (unilatéraux), 15 à 20 % sont héréditaires. Par une étude statistique, il propose que deux événements mutationnels soient nécessaires pour l'initiation tumorale. Chez les patients ayant déjà une mutation héréditaire, un seul événement supplémentaire serait nécessaire à l'apparition d'un rétinoblastome, ce qui explique la survenue plus précoce de ces cancers héréditaires [85].

L'étude de plusieurs patients présentant un « syndrome du chromosome 13q » a permis de lier la survenue d'un rétinoblastome avec une délétion sur le bras long du chromosome 13. Cette délétion, en plus d'une forte prédiction au rétinoblastome, est associée à des retards mentaux et à des anomalies

congénitales comme par exemple des microcéphalies ou des microphthalmies [86,87]. Ces données suggéraient que le gène responsable du rétinoblastome se situe dans cette région génomique. C'est en 1987 que ce gène, appelé *RB* a été identifié sur le chromosome 13, puis cloné [88].

*RB* est le premier gène suppresseur de tumeurs à avoir été identifié [89]. Par la suite, son étude a permis de déterminer que l'inactivation de la voie RB est un prérequis au développement de nombreux autres cancers comme certains cancers pulmonaires, l'ostéosarcome [90] ou les cancers génito-urinaires. En effet, on considère que plus de 90% des carcinomes pulmonaires à petites cellules [91] et 30 à 60 % des cancers de la vessie ou de la prostate sont liés à une dérégulation de la voie RB [92,93].

## 2. ...à la protéine Rb

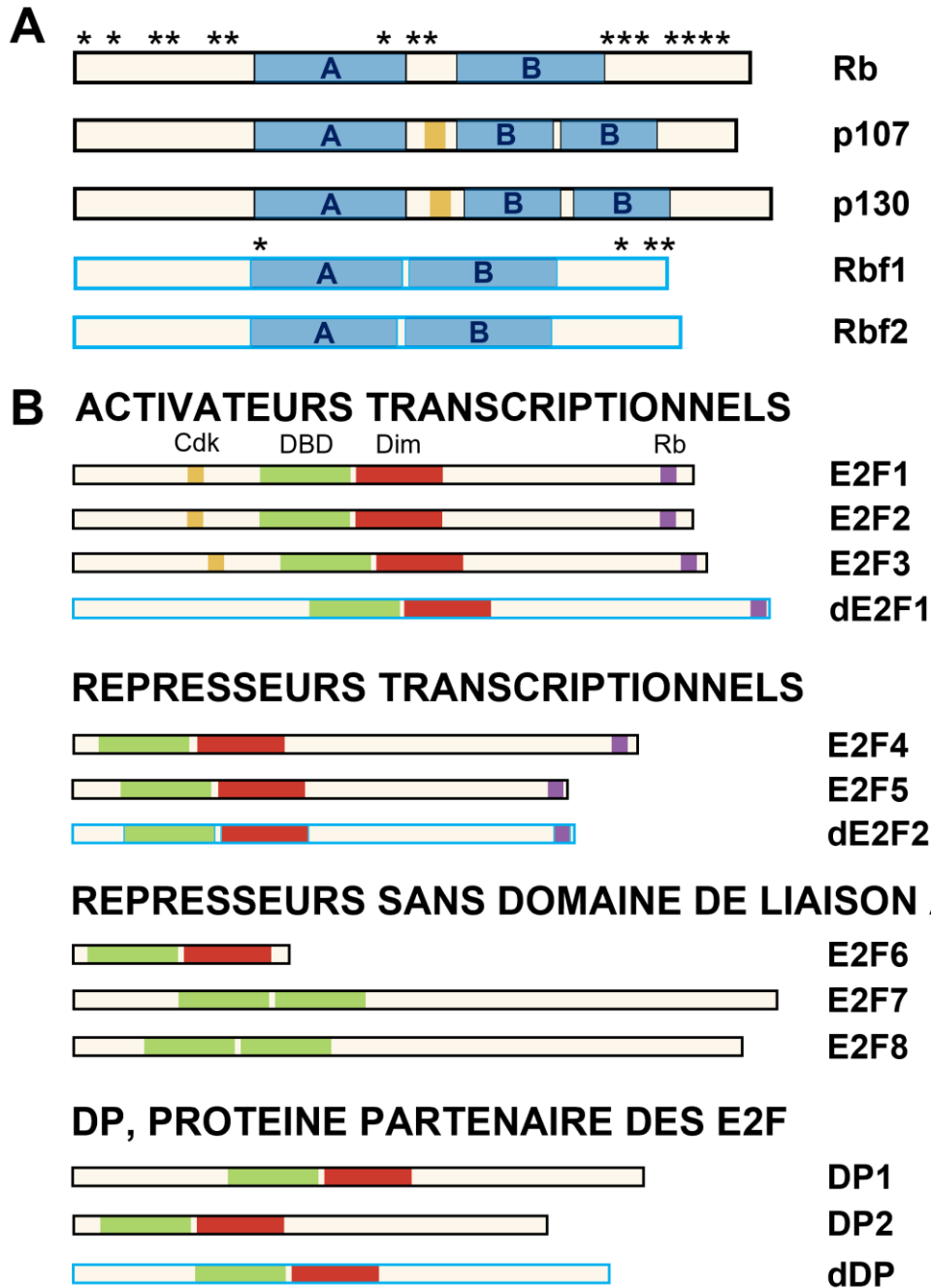
### Nomenclature

*RB1* : gène du rétinoblastome, mammifère  
*pRb* : protéine du rétinoblastome, mammifère  
*rbf1* : gène du rétinoblastome, drosophile  
*Rbf1* : protéine du rétinoblastome, drosophile

Le gène *RB1* code pour la protéine du même nom. La protéine Rb (ou pRb) fait partie, avec p107 et p130, de la famille des protéines à poche. Cependant, les cancers impliquant des mutations de *p107* ou *p130* sont bien plus rares

que celles impliquant *RB1* [91]. Ceci pourrait être dû à une homologie de séquence différente entre ces trois facteurs : si *p107* et *p130* partagent plus de 50% d'identité, *RB1* ne partage que 30 à 35% d'identité avec ces gènes [94]. Ce sont des co-facteurs de transcription, dont la séquence se trouvant dans le domaine poche permet des interactions protéines-protéines. Un domaine spacer sépare la poche en deux sous-domaines, les poches A et B (Figure 14A) [95].

En plus de ces domaines particuliers, les protéines de la famille RB contiennent de nombreux sites de phosphorylation (Figure 14A), utiles notamment pour la régulation de leur rôle dans le cycle cellulaire [96,97]: les formes hyperphosphorylées ont une capacité d'interaction diminuée avec leurs protéines cibles. De plus, des données récentes suggèrent l'existence d'un « code de mono-phosphorylation » : les sites de phosphorylation de pRb contrôleraient son association avec ses partenaires, et ainsi la régulation des gènes associés [98].



**Figure 14 : Famille des protéines RB et E2F/DP chez les mammifères et *Drosophila melanogaster*.** (A) Famille des protéines à poche. Chez les mammifères, elle est constituée de Rb, p107 et p130. Chez la drosophile, de RBF1 et RBF2. A et B représentent les domaines « poche » conservés nécessaires aux interactions protéine-protéine. Les \* correspondent aux sites de phosphorylation. En jaune : domaine de liaison aux cyclines. Contour bleu : homologues chez la drosophile. (B) Famille des protéines E2F. Les mammifères possèdent huit E2F (E2F1-8) et deux DP (DP1 et DP2). Chez la drosophile (contour bleu), on trouve deux E2F (dE2F1 et dE2F2) et une DP (dDP). Elles possèdent toutes un domaine de liaison à l'ADN (DBD, vert) et un domaine de liaison aux DP (Dim, rouge) sauf E2F7 et E2F8. Les dE2F et E2F1-4 ont un domaine de liaison à Rb (Rb, violet). En lien avec leur rôle dans le cycle cellulaire, les E2F activatrices de la transcription ont un domaine de liaison aux cyclines/cdk (Cdk, jaune). *Adapté de Du and Pogoriler, 2006*

Chez la souris adulte, pRb est exprimée de façon ubiquitaire. En revanche, lors de l'embryogenèse, les niveaux de transcrite de *RBI* sont régulés temporellement et spatialement : il est exprimé à hauts niveaux dans le système nerveux, les cellules sanguines et le cristallin [90]. pRb joue un rôle indispensable lors du développement car des souris mutées pour *RBI* meurent entre 13 et 15 jours de gestation en présentant de nombreux défauts de différenciation dans les muscles et les globules rouges, ainsi que des phases S ectopiques [99]. Quant à elle, l'expression de *p107* est limitée aux tissus comprenant des cellules en prolifération, tandis que celle de *p130* est faible et diffuse [100]. Les mutations de *p107* ou *p130* n'entraînent pas de défauts de développement majeurs et conduisent à la naissance d'individus viables [101].

En tant que co-facteur de transcription, l'activité de pRb passe par l'interaction avec de nombreuses protéines : plus de cent ont été rapportées comme interagissant avec pRb, la plupart *via* son domaine poche [102]. Les partenaires les plus étudiées sont les protéines de la famille E2F.

Les protéines de la famille E2F sont des facteurs de transcription qui agissent en hétérodimères contenant une sous-unité E2F et une sous-unité DP. Toutes possèdent un domaine de liaison à l'ADN et un domaine de dimérisation Dim (à l'exception des E2F 7 et 8). Chez les mammifères, il existe 2 DP (DP1 et DP2) et 8 E2F (Figure 14B). Ces dernières sont réparties en trois sous-groupes [95]:

- Activateurs : E2F 1 à 3. Ils possèdent une séquence de liaison aux cdk et une séquence de localisation nucléaire à l'extrémité N-terminale.
- Répresseurs : E2F 4 et 5. Ils possèdent une séquence d'export nucléaire et sont transloqués dans le noyau lorsqu'ils sont en complexe avec Rb.
- Répresseurs sans domaines de liaison à Rb : E2F 6 à 8. Leur action est indépendante de la famille Rb.

Les activateurs se lient préférentiellement à pRb, tandis que les répresseurs à p107 (E2F4) et p130 [103].

Les protéines de la famille Rb, ainsi que leurs partenaires et leurs interactions, sont très conservés. Des comparaisons de séquences génomiques ont révélé des homologues de *RB* dans tous les groupes majeurs d'eucaryotes, laissant penser que ce gène existait déjà chez le dernier ancêtre commun des eucaryotes il y a 1,6 à 2 milliards d'années [104]. Ainsi, on trouve des homologues de ces protéines chez la drosophile, le nématode [105], ou encore les plantes comme celles du genre *Arabidopsis* [106].

Chez *Drosophila melanogaster*, ces homologues sont présents avec une moindre redondance. Il existe 2 E2F (dE2F1 activateur, dE2F2 répresseur [107]), 1 DP (dDP [108]) et 2 protéines de la famille RB (Rbf1 et Rbf2 [97]) (Figure 14). Si évolutivement, ces deux dernières sont plus proches de p107 et p130, le mode d'action de Rbf1 semble de rapprocher de celui de pRb [109]. En effet, comme chez les mammifères, Rbf1 se lie aux deux dE2Fs alors que Rbf2 ne se lie qu'au répresseur dE2F2 [107,109].

De plus, la perte de fonction de *rbf1* est létale comme celle de *RB* chez la souris. Enfin, Rbf1 est présente chez l'adulte, alors que les niveaux de Rbf2 sont réduits chez la femelle adulte et indétectables chez le mâle. Dans l'embryon de drosophile, l'expression de *rbf1* est assez constante, avec un pic de transcrits entre 0 et 4h, puis des transcrits plus concentrés dans les intestins, l'épiderme et le système nerveux central durant le développement. Chez la larve, l'expression de *rbf1* est uniforme dans les disques imaginaux de pattes, d'ailes et d'haltères. Dans le disque imaginal d'œil-antenne, elle est retrouvée principalement dans les régions flanquant le sillon morphogénétique [100].

## II- pRb, un co-répresseur de transcription conservé

### 1. De nombreuses études pour caractériser les cibles de pRb et de ses homologues

pRb est connu pour être principalement un répresseur transcriptionnel. La nature de co-facteur de transcription de pRb et de ses homologues soulève immédiatement la question de ses cibles transcriptionnelles et des facteurs qui lui sont associés. Ainsi, de nombreuses études génomiques se sont intéressées aux profils des gènes régulés par pRb et ses homologues (Figure 15).

En 2010, Agustin Chicas *et al.*, dans le but de déterminer le rôle de pRb dans la sénescence, réalisent des expériences d'immunoprécipitation de chromatine suivie de séquençage (ChIP-seq) dans des fibroblastes humains en croissance, quiescents ou sénescents. Ils mettent ainsi en évidence respectivement 4525, 4214 et 8336 sites de fixation de pRb, dont 1281 redondants dans chaque type cellulaire [110]. Plus récemment, d'autres expériences de ChIP-seq ont réalisées avec des formes sauvage, mono-phosphorylées et non phosphorylables de pRb, cette fois-ci dans des cellules d'épithélium pigmentaire humain, 9069 à 12660 sites de fixations ont ainsi pu identifiés [111].

Deux études utilisant des techniques similaires ont été réalisées chez *Drosophila melanogaster*, dans des larves de 3<sup>e</sup> stade (L3) [112] ou des embryons [113], permettant de déterminer les régions génomiques associées à Rbf1 et/ou dE2F1 et dE2F2. Elles ont mis en évidence plus de mille sites de liaison de Rbf1, associés à presque 2000 gènes. Les 1890 gènes de drosophile ainsi identifiés chez l'embryon correspondent à 2310 orthologues humains. En les comparant avec les données obtenues en fibroblastes humains [110], 27 à 46 % (570 à 974) d'entre eux ont été identifiés comme étant également des cibles de pRb [113].

Des expériences de transcriptomique ont également été réalisées chez le nématode *C. elegans* en embryon et en larves de premier et quatrième stades (L1 et L4) en contexte mutant pour l'unique représentant des protéines à poches, lin-35. 1949 gènes ont une expression altérée à au moins un des stades. De manière cohérente avec le rôle connu de répresseur transcriptionnel, 64 à 81 % de ces gènes sont activés en contexte lin-35 mutant. Environ 30 % des gènes régulés par les protéines à poche chez l'humain et ayant des orthologues chez *C. elegans* sont régulés par lin-35 [105].

La plupart des sites identifiés se trouvent à proximité du site d'initiation de la transcription (TSS), révélant une fixation à proximité des promoteurs (Figure 15). La comparaison des sites de fixation de Rbf1 en larve L3 avec une liste de gènes dérégulés en cellules S2 traités avec des ARN interférents ciblant Rbf1 [114] a permis de confirmer que la majorité des répressions transcriptionnelles médiées par Rbf1 sont directes. En effet, 84 % des gènes dont l'expression est augmentée en contexte d'inhibition de Rbf1 apparaissent dans ses sites de fixation, ces gènes sont donc directement réprimés par Rbf1 en contexte sauvage [112]. L'étude plus récente de Sanidas *et al.* en cellules humaines s'éloigne de ce modèle en ne montrant que les promoteurs ne représentent que 37 % des sites de fixation de pRb sauvage, le reste étant des régions inter- et intragéniques, et notamment des enhanceurs [111].

Publication	Organisme et type cellulaire	Technique	Nombre de gènes cibles	Position par rapport au TSS	Motifs associés
Chicas <i>et al.</i> , 2010	Fibroblastes humains en croissance, quiescents ou sénescents	ChIP-seq	4525, 4214 ou 8336 (dont 1281 redondants)	80 % dans les 5 kb avant le TSS	E2F
Korenjak <i>et al.</i> , 2012	Drosophile, larves L3	ChIP-on-chip	2130, dont 84 % avec une expression augmentée en mutant Rbf1 ou Rbf2 (comparaison avec gènes dérégulés en cellules Kc)	82 % entre 1 kb en amont et 200 bp en aval du TSS	E2F, DRE, CHR
Acharya <i>et al.</i> , 2012	Drosophile, embryons	ChIP-seq	1890	95 % dans les 2 kb autour du TSS	E2F, DREF, FOXJ2, RAM
Sanidas <i>et al.</i> , 2022	Cellules d'épithélium pigmentaire rétinien humain	ChIP-seq	12660 pour pRb sauvage	37,6 % entre 1 kb en amont et 100 pb en aval du TSS pour pRb sauvage, le reste dans des régions non promotrices	E2F, SP1, ETS, c-Jun, CTCF

**Figure 15 : Tableau récapitulatif de quelques études de génomique visant à identifier les cibles de pRb et de ses homologues.** Pour rappel, le génome humain contient environ 21 000 gènes et le génome de drosophile environ 13 000.

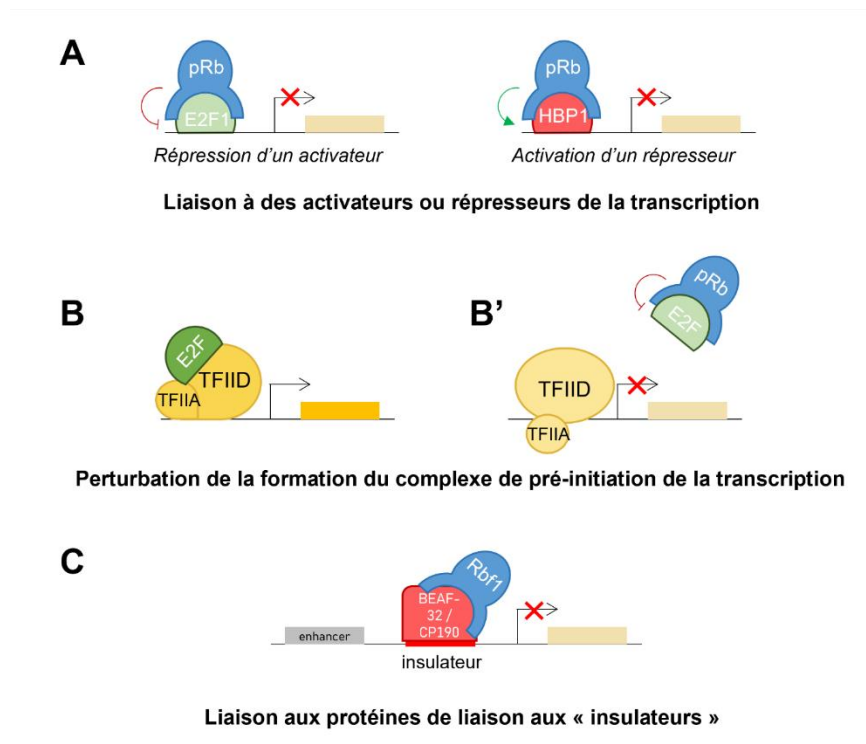
La famille des protéines à poche ne possédant pas de domaine de liaison à l'ADN, la fixation aux promoteurs des gènes régulés se fait par l'intermédiaire de protéines pouvant se lier à l'ADN, comme des facteurs de transcription. La recherche de motifs connus dans les promoteurs des gènes régulés révèle une forte représentation des séquences cibles des facteurs E2Fs, partenaires principaux des RBs, dans tous les organismes [105,110,112,113] (Figure 15). Ce ne sont cependant pas les seuls, puisque des motifs similaires à ceux liés par DREF (DNA Replication-related Element-binding Factor)

[112,113] et par le facteur de transcription FOXJ2 de la famille Forkhead [113] sont mis en évidence, ainsi qu'un motif CHR (Cell-cycle gene Homology Region) lié par des orthologues de Mip120 [115], un partenaire des RBs au sein du complexe de régulation transcriptionnelle DREAM/dREAM. Un nouveau motif a également été identifié et nommé RAM pour « motif associé à Rbf » (Rbf Associated Motif) [113].

Les analyses par Gene Ontology (GO) ont permis de confirmer le rôle de pRb et de ses homologues dans la régulation du cycle cellulaire et notamment la réplication [105,110,113]. Cependant, d'autres processus moins connus sont mis en évidence par ces analyses : réparation de l'ADN, modification de la chromatine ou encore développement (neurogenèse, ovogenèse, polarité cellulaire, spermatogénèse) [105,113,116,117]. La comparaison des cibles identifiées en fibroblastes humains [110] avec celles identifiées en embryons de drosophile a permis de mettre en évidence une forte conservation, avec notamment plus de 200 voies de signalisation conservées et indispensables telles que les voies Hippo, JNK, Wnt ou Notch [113].

## 2. Différents modes de régulation de la transcription

### a. Par liaison à des protéines de régulation de la transcription



**Figure 16 : Régulation de la transcription par liaison des protéines RB à des protéines de régulation de la transcription.** (A) pRb peut se lier à des activateurs (E2F1) ou des répresseurs (HBP1) de la transcription pour les réprimer ou les activer, respectivement. (B) Lors de l'activation des gènes, les facteurs de transcription de la famille TFII sont recrutés aux promoteurs pour permettre l'initiation de la transcription. Il peut recruter des facteurs de transcription comme les E2Fs. (B') En présence de pRb, celle-ci bloque la formation du complexe en séquestrant les E2Fs, ou en entraînant l'assemblage de protéines dans une conformation inactive. *Adapté de Ross et al., 1999* (C) Chez la drosophile, la présence de protéines de liaison aux « insulateurs » est significativement corrélée à des promoteurs liés par Rbf1 et/ou Rbf2.

pRb et ses homologues ne possédant pas de domaine de liaison à l'ADN, elles doivent interagir avec d'autres protéines pour réguler l'expression des gènes. Dans ce but, pRb est connue pour ses interactions avec de nombreux facteurs de transcription *via* des domaines d'interaction présents sur ces facteurs (Figure 14 et ref [118]). Comme évoqué dans le paragraphe précédent, des études génomiques ont mis en évidence des motifs redondants dans les séquences associées aux pRb. On les retrouve associées à des facteurs de transcription comme la famille bHLH (MyoD, NeuroD1), GATA1 [113] ou encore les C/EBPs [119], mais le plus connu et le plus représenté est la famille des E2Fs, et en particulier les E2Fs activateurs (E2F1 à 3 chez les mammifères, dE2F1 chez la drosophile).

pRb étant majoritairement un répresseur transcriptionnel, la régulation de l'expression des gènes cibles des E2Fs se fait par répression de l'action de ce dernier. Pour cela, deux stratégies ont été mises en évidence.

Premièrement, comme bien connu dans la régulation de la transition G1/S du cycle cellulaire (voir partie B-III-1.), la forme hypophosphorylée de pRb peut se lier au domaine activateur des E2Fs activateurs et ainsi bloquer leur activité [103,120]. Cette forme active peut également se lier à d'autres protéines, comme le répresseur transcriptionnel HBP1 (Figure 16A). On sait par exemple que cette protéine nécessite pRb ou p130 pour réprimer le promoteur de *N-MYC*, un gène nécessaire au cycle cellulaire [121]. Il est d'ailleurs intéressant de noter que ce promoteur est également activé par les E2Fs [94]. La progression du cycle cellulaire, et donc l'état de phosphorylation de pRb semble également jouer sur ses interactions et sa localisation sur le génome. En effet, des expériences de ChIP-seq avec des formes sauvages, mono-phosphorylées et non phosphorylables de pRb ont montré une relocalisation de pRb, présent au niveau des promoteurs et colocalisant avec E2F1 en G1 et au niveau des enhancers, colocalisant avec c-Jun dans les cellules cyclantes. On trouve également une superposition des sites de liaison pRb avec CTCF, un régulateur négatif de la transcription non spécifique des promoteurs ou des enhancers. Cette superposition semble spécifique de gènes ayant une fonction neuronale [111].

Deuxièmement, pRb réprime la transcription par toutes les ARN polymérases. En effet, l'interaction de pRb avec des facteurs nécessaires au recrutement de ces protéines inhibe la transcription [122]. Par exemple, le recrutement de pRb aux promoteurs bloque l'assemblage du complexe de pré-initiation et inhibe les facteurs de transcription adjacents. Ce complexe contient l'ARN polymérase II et des facteurs de transcription de la famille TFII et est nécessaire à la transcription. Il recrute ensuite différents autres composants, comme des facteurs de transcription tels que les E2Fs et leurs partenaires DPs (Figure 16B). La présence de pRb bloque la formation de ce complexe, et donc la transcription du gène cible. Cette inhibition pourrait se faire par masquage du site d'interaction entre E2F4 et les TFII, ou en entraînant un assemblage de protéines dans une



conformation inactive [123] (Figure 16B'). pRb est également impliquée dans la répression des gènes dépendants des ARN polymérase I et III [122].

Chez la drosophile, on trouve un troisième type de protéine partenaire des RBs : les protéines de liaison aux « insulateurs ». Les « insulateurs » sont des éléments d'ADN possédant une activité de blocage des enhanceurs, ils peuvent perturber la communication entre un enhanceur et un promoteur lorsqu'ils sont placés entre ces deux éléments. La présence de motifs BEAF-32 et CP190, les protéines de liaison aux « insulateurs » de drosophile, est significativement corrélée à des promoteurs liés par Rbf1 et/ou Rbf2 [124] (Figure 16C). Cette association se fait par l'intermédiaire du complexe de régulation transcriptionnelle dREAM (voir partie C).

### b. Par modification de la chromatine

La configuration de la chromatine est déterminante pour la régulation de l'expression des gènes. L'hétérochromatine, état condensé de la chromatine, correspond à des régions de faible transcription. Le niveau de condensation dépend à la fois de l'activité de facteurs de remodelage de la chromatine et de modifications post-traductionnelles appliquées sur les histones. pRb et ses homologues agissent sur ces deux événements pour réguler la transcription des gènes.

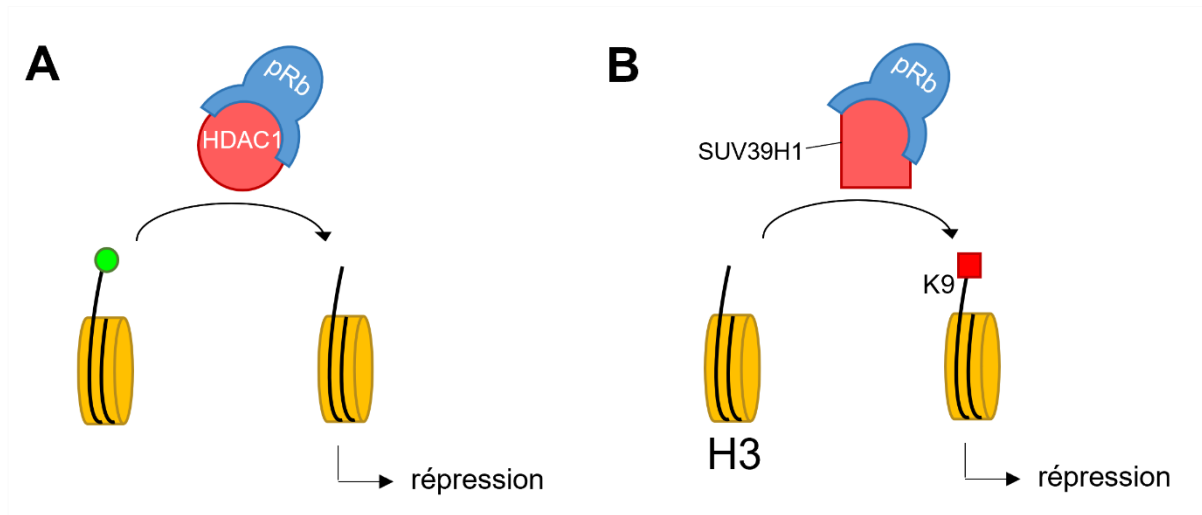
#### Condensines et facteurs de remodelage de la chromatine

Les complexes Condensine, constitués des condensines I et II, sont nécessaires à la condensation et la stabilité de la chromatine dans les phases précoces de la mitose. pRb et ses homologues interagissent spécifiquement avec la condensine II pour favoriser la condensation [125–127].

Les facteurs de remodelage de la chromatine sont des protéines ATP-dépendantes capables de modifier l'accessibilité de l'ADN nucléosomal en modulant les liens histones-ADN [128]. Les SWI/SNF, une des familles de protéines de remodelage les plus étudiées, sont des complexes de remodelage des nucléosomes chez la levure, qui ont pour homologues chez les mammifères les protéines BRG1 et hBRM. Ces deux protéines sont capables d'interagir physiquement avec les protéines de la famille RB, et BRG1 est nécessaire à l'arrêt du cycle induit par pRb dans certaines lignées cellulaires [129].

#### Modifications post-traductionnelles des histones

Les histones sont des protéines formant des octamères autour desquels est enroulé l'ADN. Certains résidus d'histones sont les cibles d'enzymes qui y appliquent des modifications post-traductionnelles qui forment un « code histone ». Ce code, composé d'acétylations, de méthylations, de phosphorylations et d'ubiquitinations, est lu par diverses protéines et interprété comme différents états chromatiniens. L'acétylation et la méthylation sont les deux modifications les plus étudiées [130].



**Figure 17 : Modifications post-traductionnelles des histones médiées par pRb.** (A) pRb interagit avec HDAC1, qui va retirer les marques acétyl activatrices sur les histones et ainsi entraîner la répression transcriptionnelle. (B) pRb interagit avec la méthyltransférase SUV39H1 qui appose une marque répressive me-K9-H3. Ces deux mécanismes sont également retrouvés chez la drosophile.

Les histones hyperacétylés sont en général associés à des gènes activement transcrits, les hypoacétylés à des gènes réprimés. Les protéines responsables de ces marques sont les histones acétyltransférases (HAT), qui apposent des acétyl, et les histones désacétylases (HDAC), qui les retirent.

pRb interagit avec HDAC1 pour diriger l'activité désacétylase spécifiquement sur les promoteurs E2Fs (Figure 17A) et ainsi retirer les marques activatrices dans certaines conditions [131], comme la sénescence induite par Ras ou un arrêt du cycle induit par p16<sup>INK4a</sup> [103]. Cette interaction est également retrouvée chez la drosophile, où RBF interagit avec l'histone désacétylase dRPD3 [132].

La méthylation joue un rôle différent, puisque c'est la position de l'ajout du méthyl par des méthyltransférases qui va déterminer son effet sur l'état chromatinien. Ainsi, certaines marques sont associées à de l'hétérochromatine, comme la méthylation sur la lysine 9 de l'histone H3 (me-K9-H3) ; alors que la méthylation de l'histone H3 cette fois-ci sur la lysine 4 (me-K4-H3) et non sur la 9 est associée à des gènes activement transcrits.

pRb interagit avec l'histone méthyltransférase SUV39H1 qui appose spécifiquement des marques répressives me-K9-H3 [133] (Figure 17B). Celles-ci permettent le recrutement de HP1, une protéine identifiée comme étant un composant de l'hétérochromatine chez la drosophile [134]. Ce mécanisme a également été identifié chez le nématode *C. elegans*, chez lequel l'homologue de HP1, *hpl-2*, interagit génétiquement avec le gène *lin-35* qui code pour l'homologue de pRb [135].

Des complexes E2F1-pRb ont également été associés au "silencing" d'éléments transposables *via* le recrutement de l'histone méthyltransférase EZH2 et de son homologue chez la drosophile E(Z), une sous-unité du complexe de répression transcriptionnelle Polycomb (PRC2 chez la drosophile). Elle permet l'apposition de marques répressives H3-K27 (di ou tri-méthylation) [132,136].

Plusieurs de ces protéines peuvent également former des complexes pour apposer différentes marques répressives : pRb et E2F1 sont retrouvées avec la méthyltransférase DNMT1 et l'histone désacétylase HDAC1 dans la répression de promoteurs cibles des E2Fs [137]. De façon indépendante des E2Fs, les membres de la famille RB sont également des régulateurs de la structure chromatinienne par régulation de la méthylation H4K20 [138].

### Un cas intermédiaire : le complexe NuRD

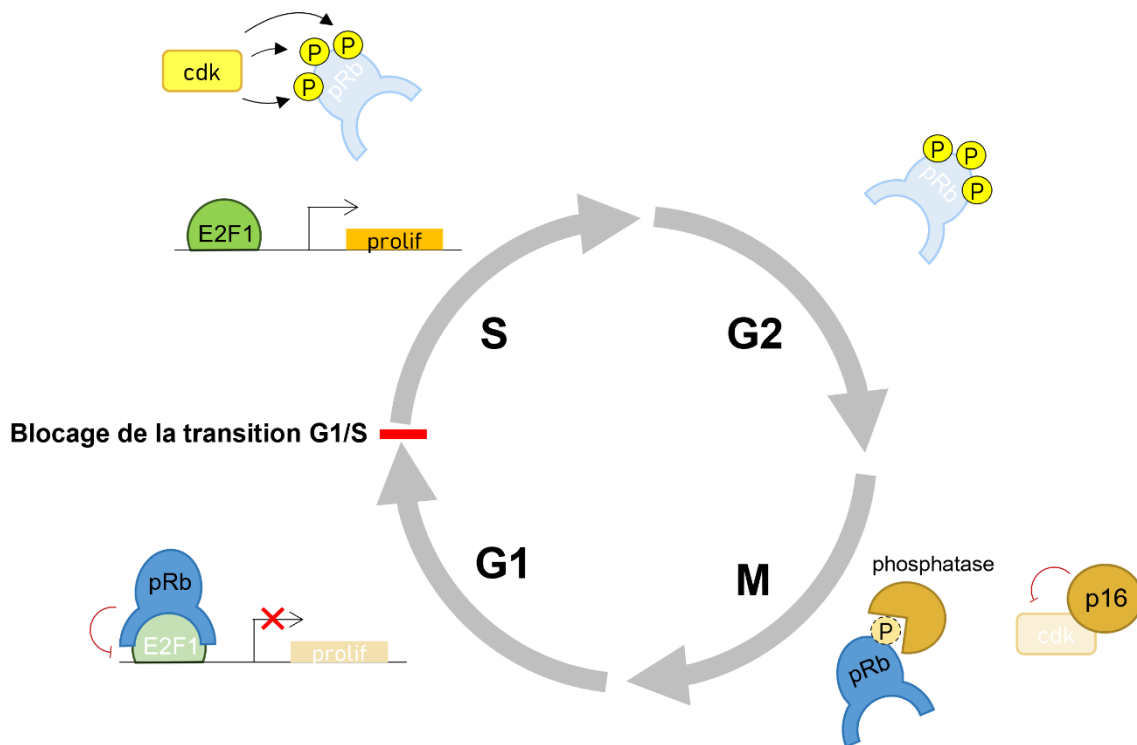
Le complexe NuRD (Nucleosomal and Remodeling Deacetylase) contient à la fois des facteurs de remodelage ATP-dépendants et des activités d'histone désacétylase grâce aux HDAC1 et 2. Chez la souris, ce complexe permet le "silencing" de l'élément transposable LINE-1. pRb permet le recrutement des composants du complexe NuRD par l'intermédiaire d'une phosphorylation sur la sérine 811 [98], et une perte de fonction de *RB* est associée à la délocalisation du noyau vers le cytosol de Mi2- $\beta$ , une des sous-unités de ce complexe [139].

## III- Des rôles essentiels dans de nombreux processus

### 1. pRb, régulateur majeur du cycle cellulaire

Le rôle le plus connu et le plus étudié des protéines de la famille RB est son rôle dans le cycle cellulaire. En effet, en inhibant les protéines E2Fs, pRb est un gardien de la transition G1/S. En effet, les E2Fs régulent un grand nombre de gènes codant pour des protéines nécessaires à l'entrée en phase S, comme des protéines impliquées dans la réplication.

La régulation de pRb et de son interaction avec les E2Fs se fait par l'intermédiaire de phosphorylations. Lorsqu'une cellule passe en phase S, pRb est inactivée par le biais de phosphorylations par les complexes cycline E/cdk2 [140]. Elle ne peut alors plus interagir avec E2F, qui est libéré et active ses promoteurs cibles [141]. Lorsqu'une cellule entre en mitose, une phosphatase clive les groupements phosphate additionnels et les CDKN, comme p16, limitent une nouvelle phosphorylation de pRb, uniquement mono-phosphorylée par les complexes cycline D/cdk4/6 [140,141]. Ainsi hypophosphorylée, elle interagit avec E2F pour réprimer son activité et donc ses gènes cibles : la progression du cycle de G1 vers S est alors inhibée [141,142] (Figure 18).



**Figure 18 : Régulation de pRb au cours du cycle cellulaire.** Durant la phase S, l'hyperphosphorylation de pRb par les cdk la rend inactive et l'empêche de réprimer E2F1, qui peut alors activer les gènes de prolifération. Lors de la mitose, une phosphatase clive les phosphates présents sur pRb, et une CDKN comme p16 empêche les cdk de phosphoryler à nouveau pRb. Cette forme hypophosphorylée active réprime E2F1 et les gènes de la prolifération, bloquant ainsi le passage en phase S.

La composition du complexe famille RB/E2F change selon la phase du cycle [143] :

- Dans les cellules quiescentes (phase G0), p130 est en complexe avec les E2Fs inhibiteurs (E2F4 et 5).
- Dans les cellules cyclantes, on trouve p107 en complexe avec E2F4 en G0/G1, et pRb avec E2F1-3 durant la phase S.

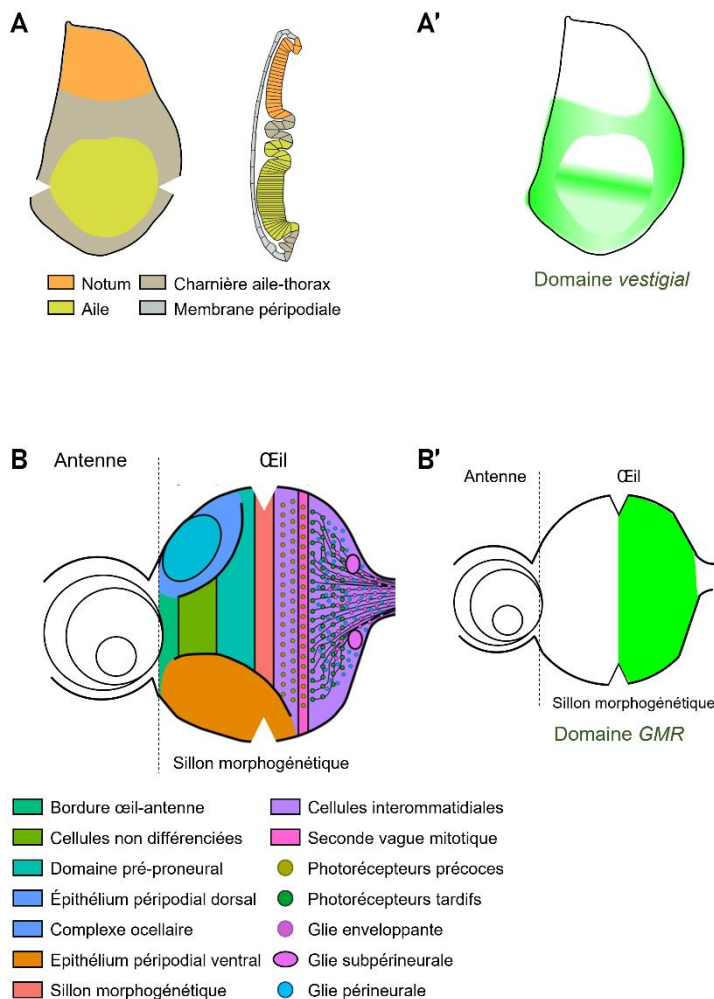
Si pRb a un rôle transcriptionnel primordial pour le passage de la transition G1/S, ce n'est pas son seul rôle dans le cycle cellulaire. En effet, elle régule également les origines de réplication lors de la phase S. Chez la drosophile, lors de la réception de signaux développementaux dans les cellules de l'ovaire, des clusters de gènes sont amplifiés en initiant à répétition la réplication de l'ADN à des sites spécifiques. La ploïdie dans ces cellules peut alors atteindre 16n. Lorsque les Rbf sont mutées, on observe des défauts de réplication, avec des cellules contenant 32n ADN, soit au moins une phase S supplémentaire [144]. Cette implication des Rbf dans le contrôle de l'endoréplication est confirmée par l'apparition de défauts d'endoréplication dans le développement des soies chez des mutants dE2F1 ne pouvant plus interagir avec les Rbf [145]. Grâce à des expériences de co-immunoprécipitation, ce rôle des Rbf a pu être élucidé : en formant des complexes avec dE2F1 et DmORC, un complexe d'initiation de la réplication, il permet une régulation négative de l'amplification des gènes. Ce

complexe a également été retrouvé dans l'embryon, suggérant un rôle général de répression des origines de réplication pour le complexe Rbf-dE2F1 [144]. Le rôle et la présence de ces complexes n'a pas été investiguée chez les mammifères, mais l'existence d'homologues pour les complexes dE2F1-Rbf et pour le complexe DmORC (ORC chez les mammifères) suggère que des mécanismes similaires pourraient exister.

## 2. Pro- ou anti-apoptotique ?

Le rôle des protéines de la famille Rb est bien moins étudié dans l'apoptose que dans le cycle cellulaire. Cependant, pRb et ses homologues peuvent promouvoir ou supprimer l'apoptose à travers des mécanismes à la fois directs et indirects [146,147]. Ces différents rôles semblent être dépendants du contexte cellulaire. Le laboratoire d'accueil s'intéresse particulièrement aux activités de Rbf1, l'homologue de pRb chez la drosophile, dans l'apoptose des cellules des disques imaginaux d'aile et d'œil de la drosophile.

### a. Les disques imaginaux d'aile et d'œil



**Figure 19 : Organisation des disques imaginaux d'aile et d'œil-antenne.** (A) Schéma du disque imaginal d'aile (gauche) et d'une section antéro-postérieure (droite). La zone de cellule non prolifératives (ZNC) se situe le long de l'axe antéro-postérieur matérialisé entre les têtes de flèches blanches. Adapté de Tripathi et Irvine, 2022. (B) Schéma du disque imaginal d'œil-antenne. Le sillon morphogénétique se situe le long de l'axe dorso-ventral matérialisé entre les têtes de flèches blanches. Adapté de Ariss et al., 2018. Les deux structures sont orientées selon un axe vertical dorso-ventral (haut/bas) et horizontal antéro-postérieur (droite/gauche). Les domaines d'expression des pilotes *vestigial* et *GMR* sont représentés respectivement dans les figures (A') et (B').

Le disque imaginal d'aile est le précurseur larvaire de l'aile, mais également de la charnière aile/thorax et d'une partie du thorax (notamment le notum). Il est constitué d'un épithélium monocouche recouvert d'une membrane péripodiale, structure qui ne participe pas directement à l'aile adulte mais favorise le repliement de l'épithélium lors de la formation de l'aile (Figure 19A). Le disque d'aile est caractérisé par une importante prolifération cellulaire. En effet, durant les stades larvaires, sa taille augmente de plus de mille fois, grâce à une augmentation du nombre de cellules [148]. Cette caractéristique nous permet d'étudier un tissu en prolifération. Cependant, parmi ces nombreuses cellules cyclantes, on peut trouver une zone particulière de non-prolifération appelée ZNC (Zone of Non-proliferating Cells), qui se situe à la frontière dorso-ventrale [149]. Le développement de l'aile adulte à partir du disque imaginal d'aile début lors des premiers stades du développement de la puppe [148].

Le disque imaginal d'œil est associé au disque imaginal d'antenne. Durant les premiers stades larvaires, tout comme le disque imaginal d'aile, le disque d'œil est composé d'un épithélium monocouche recouvert d'une membrane péripodiale [150]. A partir du troisième stade larvaire, on observe la formation d'un sillon à la marge postérieure du disque d'aile : le sillon morphogénétique. Ce sillon se déplace ensuite vers l'extrémité antérieure du disque et entraîne dans son sillage une vague de différenciation. Ainsi, les cellules antérieures à ce sillon ne sont pas différenciées et prolifèrent de façon asynchrone, alors que certaines cellules situées en postérieur initient une différenciation en photorécepteurs. Les cellules non différenciées en photorécepteurs subissent alors une seconde vague de différenciation en cellules interommatidiales (Figure 19B) [151,152]. Cette particularité fait du disque imaginal d'œil un modèle d'étude de la coordination de la prolifération cellulaire et de la différenciation.

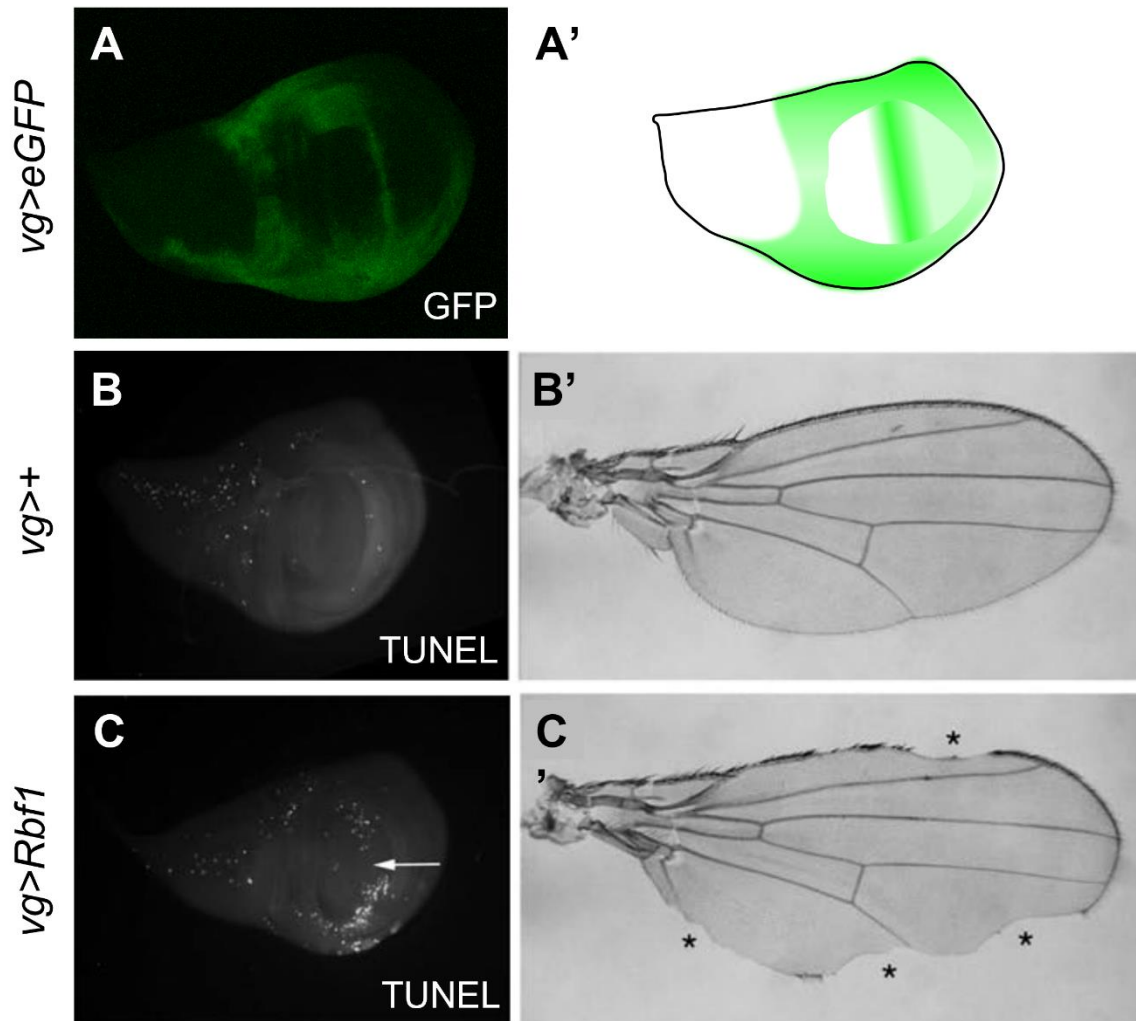
#### b. Les RB, activateurs de l'apoptose mitochondriale dans les cellules en prolifération

Les premiers mutants  $RB^{-/-}$  présentant une augmentation des niveaux d'apoptose dans certains organes chez l'embryon de souris, les chercheurs ont d'abord conclu à une activité anti-apoptotique de pRb [99]. Cette observation semblait pourtant contraire à son rôle connu de suppresseur de tumeurs, et des hypothèses se sont tournées vers une dérégulation du cycle cellulaire dans les neurones [153]. Dix ans plus tard, des études ont pu montrer que cet apparent effet anti-apoptotique chez les mutants  $RB$  est en réalité un effet non autonome-cellulaire qui provient des dérégulations du cycle cellulaire des cellules environnantes. En effet, la surprolifération du tissu placentaire entraîne des hypoxies dans les tissus embryonnaires [154]. Cependant, d'autres observations ont pu déterminer un réel rôle anti-apoptotique des protéines de la famille RB dans certains tissus [155,156]. Ce rôle anti-apoptotique sera discuté dans la section suivante.

Depuis, en accord avec son activité de suppresseur de tumeurs, de nombreuses études ont démontré un rôle pro-apoptotique de pRb. Notamment, un mutant de pRb résistant à la phosphorylation et donc constitutivement actif induit de la mort cellulaire dans des cellules en culture [157].

Si on retrouve ici un mode de régulation déjà connu de pRb (la phosphorylation) [157], son mode d'action n'est pour autant pas toujours conforme à son rôle classique de répresseur transcriptionnel. En effet, dans les cellules en prolifération, on trouve des complexes pRb-E2F1 associés à l'HAT pCAF en réponse à des dommages à l'ADN. Ces complexes sont liés à des promoteurs de gènes pro-apoptotiques comme *caspase 7* qui code pour la protéine du même nom. On retrouve également des complexes pRb-HDAC1 réprimant des gènes du cycle cellulaire [147].

Le laboratoire d'accueil travaille depuis de nombreuses années sur l'apoptose induite par Rbf1, l'homologue de pRb chez la drosophile. Ainsi, il a pu montrer que la surexpression de Rbf1 dans le disque imaginal d'aile par l'intermédiaire du pilote *vestigial* induit de l'apoptose dans ce tissu, visible par un marquage TUNEL (Terminal deoxynucleotidyl transferase dUTP Nick End Labeling). Chez l'adulte, cette apoptose durant le développement se traduit par l'apparition d'encoches à la marge de l'aile, dont le nombre et l'étendue sont corrélés à la quantité d'apoptose induite [158] (Figure 20). Cette mort est dépendante des caspases, de dE2F2 et de dDP, mais indépendante de l'interaction avec dE2F1 [159].



**Figure 20 : Rbf1 est pro-apoptotique dans le disque imaginal d'aile.** (A) Le domaine vestigial (vg) est visible grâce à l'expression GFPGFP et s'étend sur la poche de l'aile, il peut être schématisé comme en (A'). Disque imaginal d'aile (B) et aile adulte (B') contrôles. Lorsqu'on surexprime Rbf1 dans le domaine vestigial, on observe un marquage TUNEL (C) qui témoigne de cellules en apoptose et qui se traduit dans l'aile adulte par l'apparition d'encoches à la marge de l'aile (C'). La flèche blanche indique une zone centrale de la poche où les cellules ne sont pas marquées. *Adapté de Milet et al., 2010.*

La drosophile a permis d'élucider une partie de la cascade menant à l'apoptose induite par Rbf1. En effet, l'activité de répresseur transcriptionnel de pRb et de ses homologues a conduit notre équipe à l'hypothèse suivante : les complexes Rbf1/dE2F2 peuvent réprimer des gènes anti-apoptotiques. Ils ont pu alors montrer qu'une surexpression de Rbf1 dans le disque imaginal d'aile entraîne une diminution des niveaux d'ARNm de deux facteurs anti-apoptotiques : *buffy*, membre anti-apoptotique de la famille Bcl-2 ; et l'inhibiteur de caspases *dIAP1*. Des expériences utilisant un gène rapporteur ont permis de démontrer que, contrairement à *buffy* dont la répression transcriptionnelle par Rbf1 est directe, celle de *diap1* est indirecte. En effet, l'ARNm de *dIAP1* est dégradé par la protéine issue du gène *how*, dont la transcription est directement activée par Rbf1 [159].



Comme chez les mammifères, la mitochondrie est centrale dans l'apoptose chez la drosophile. L'implication de protéines telles que Buffy, de la famille Bcl-2, et Diap1, suggérait que l'apoptose induite par Rbf1 pouvait être mitochondriale (Figure 21).

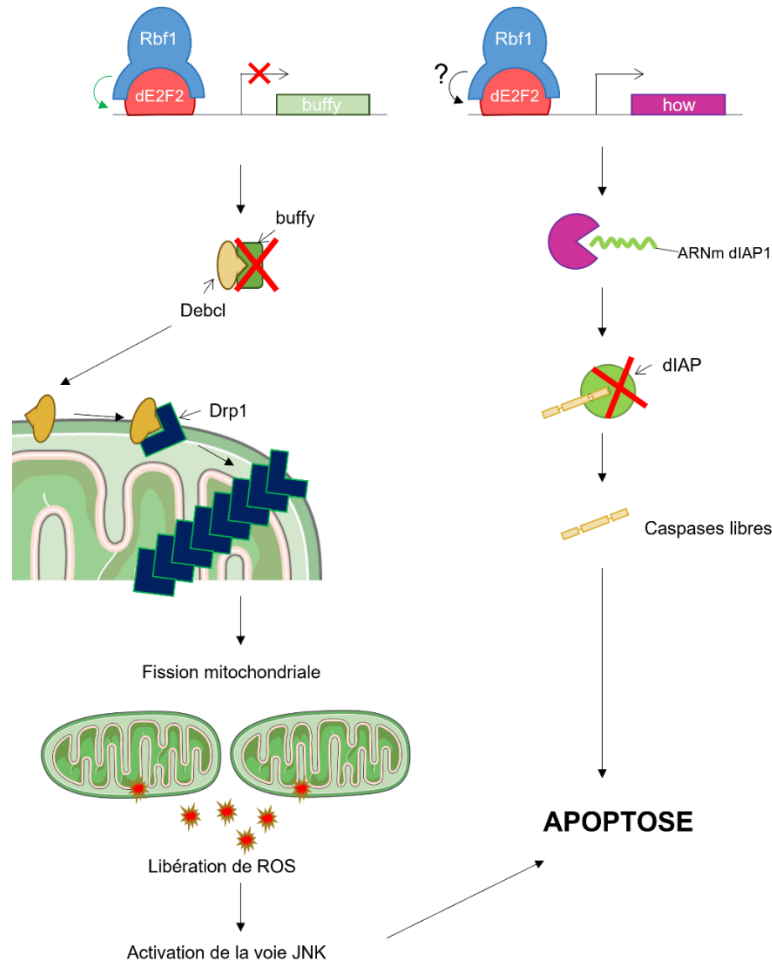


Figure 21 : Apoptose induite par Rbf1 chez *Drosophila melanogaster*

Afin de poursuivre d'en poursuivre la caractérisation, le laboratoire s'est ensuite intéressé à Debcl, l'autre membre de la famille Bcl-2 chez la drosophile, pour sa part pro-apoptotique. Comme vu dans le chapitre "les régulateurs de l'apoptose, Debcl, aussi connue sous le nom de Drob-1, dBorg-1 ou encore BOK, est nécessaire à l'induction de la mort cellulaire dans certaines conditions, notamment l'apoptose induite par des dommages à l'ADN [160] ; et son expression ectopique induit de l'apoptose [67]. Buffy et Debcl s'inhibe mutuellement par interaction physique [68,160]. Dans la cascade d'apoptose induite par Rbf1, des expériences d'interactions génétiques ont permis de placer Debcl en aval de Buffy. Ainsi, l'inactivation transcriptionnelle de Buffy par Rbf1 permet l'activation de Debcl [161] (Figure 21). Cette activation entraîne la production d'espèces réactives de l'oxygène (ROS : Reactive Oxygen Species), responsables de l'activation de la voie Jun-Kinase (JNK) [161–163].

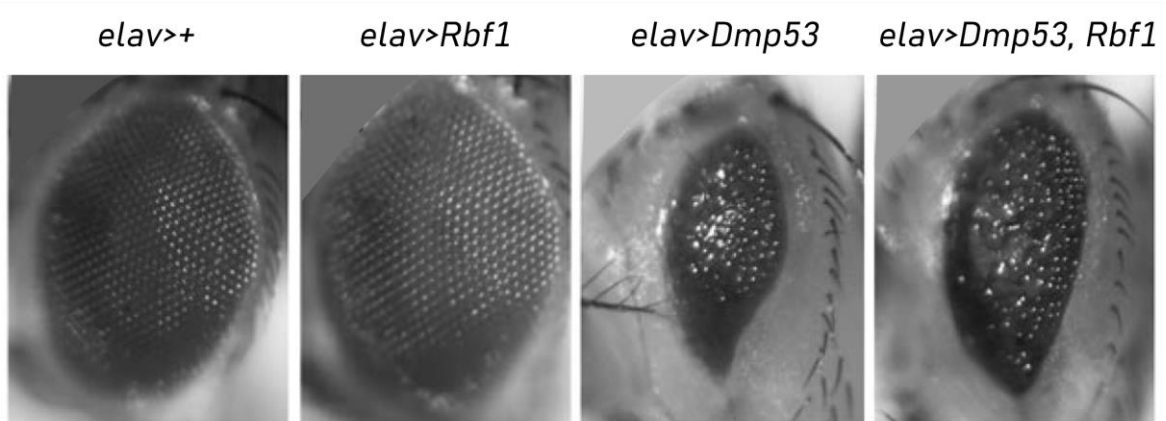
La dynamique mitochondriale est également un acteur majeur de l'apoptose chez la drosophile. En effet, lors de l'induction de l'apoptose, on observe de la fission mitochondriale, médiée entre autres par la protéine Drp1, requise pour la production de ROS (Figure 21). Drp1 est requis pour l'apoptose induite par Debcl, et des expériences de co-immunoprécipitation et de PLA (Proximity Ligation Assay) ont démontré une interaction physique entre Drp1 et Debcl [161] (Figure 21).

Chez les mammifères, une condition nécessaire au bon déroulement de la fission mitochondriale est la relocalisation de BAX (Bcl-2 pro-apoptotique) à la mitochondrie, qui colocalise alors avec Drp1 et la mitofusine 2 [164,165]. S'il n'y a pas de preuve d'une interaction physique entre BAX et Drp1, on sait cependant que BAX stimule la sumoylation de Drp1 et ainsi son association aux membranes mitochondriales [161]. pRb pourrait indirectement participer à la fission mitochondriale par une fonction non nucléaire. En effet, il a été montré que pRb interagissait directement avec BAX à la mitochondrie et permettait la perméabilisation de la membrane mitochondriale externe, et ainsi le relargage du cytochrome c. Cette fonction est indépendante des activités transcriptionnelles de pRb [166]. Cette localisation non nucléaire avait déjà été observée chez les mammifères, mais n'a pas été montrée avec l'homologue de drosophile de pRb [167].

pRb et ses homologues ont donc une fonction pro-apoptotique. L'apoptose induite par pRb est une apoptose mitochondriale, qui passe par une activation directe (mammifère) ou indirecte (drosophile) des protéines pro-apoptotiques de la famille Bcl-2. Cette activation permet la libération de ROS et l'activation de la voie JNK, ainsi que la fission mitochondriale. Ces étapes mènent à la mort de la cellule. Cependant, l'activité pro-apoptotique de pRb n'est observée que dans des tissus en prolifération. En effet, une surexpression de Rbf1 dans les cellules de la ZNC du disque imaginal d'aile, qui contient des cellules qui ne prolifèrent pas, n'induit pas d'apoptose [158].

### c. Un rôle anti-apoptotique dans les cellules non prolifératives

Comme évoqué dans la section précédente, malgré un rôle principalement anti-apoptotique, les protéines de la famille Rb peuvent également protéger de l'apoptose dans certains tissus.

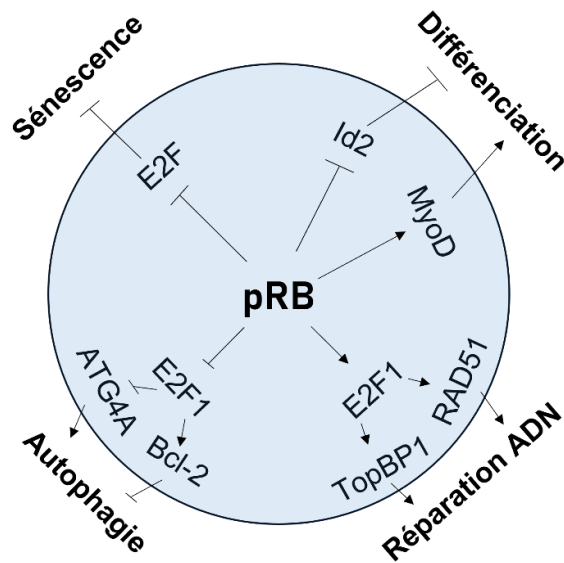


**Figure 22 : Rbf1 est anti-apoptotique dans les neurones.** Rbf1 et/ou Dmp53 sont surexprimés grâce au pilote pan-neuronal *elav*. Adapté de Milet et al, 2010.

Chez la drosophile, la surexpression de Dmp53, l'homologue du pro-apoptotique p53, par l'intermédiaire du pilote pan-neuronal *elav* induit de l'apoptose visible par marquage TUNEL et se traduisant par une réduction de l'œil adulte. Lors de la surexpression de Rbf1 *via* ce même pilote, on n'observe ni marquage TUNEL ni réduction de l'œil. Rbf1 n'a donc pas de rôle pro-apoptotique dans les neurones, et en particulier dans l'œil de drosophile. Plus étonnant encore, la co-surexpression de Dmp53 et de Rbf1 entraîne un sauvetage partiel du phénotype d'œil associé à Dmp53, avec un œil moins réduit. Rbf1 est donc anti-apoptotique dans ce tissu [158] (Figure 22).

L'activité anti-apoptotique de pRb et de ses homologues est, comme son activité pro-apoptotique, peu décrite par rapport aux activités des protéines de la famille RB dans le cycle cellulaire. Les observations disponibles semblent cependant indiquer que pRb serait anti-apoptotique dans les cellules non prolifératives, comme les neurones. En effet, ce rôle n'est pas spécifique du disque imaginal d'œil puisque la surexpression de Rbf1 grâce au pilote *eyeless*, qui permet une expression antérieure à l'apparition du sillon morphogénétique, induit une apoptose des cellules du disque imaginal d'œil [158].

### 3. Et bien plus encore !



**Figure 23 : Implication de pRb dans des processus cellulaires autres que le cycle cellulaire et l'apoptose.** pRb peut inhiber ou activer des facteurs de transcription, dont les E2Fs pour promouvoir la différenciation, la sénescence, l'autophagie et la réparation de l'ADN. *Inspiré de Khidr et al., 2006.*

#### a. Développement et différenciation

Une des étapes clés du développement est la différenciation terminale, durant laquelle les cellules précurseurs s'engagent dans un programme de différenciation. Elles expriment tout d'abord des marqueurs de différenciation précoces. Dans certains tissus, on observe également une expansion clonale, et certaines cellules peuvent se retirer du cycle cellulaire de manière irréversible. Ces cellules deviennent alors plus résistantes à l'apoptose et expriment des gènes tardifs spécifique du tissu. Enfin, des changements morphologiques s'opèrent jusqu'à acquisition du phénotype totalement différencié [94].

Durant ces processus, pRb et ses homologues sont indispensables à la mise en place de nombreuses structures, car les souris déficientes pour pRb meurent de 13 à 15 jours de gestation avec des défauts dans les développements érythroïdes, neuronaux et du cristallin [99]. Dans ces souris, l'initiation de la différenciation est faite car on trouve une expression des gènes de différenciation précoce, mais les cellules n'atteignent pas le stade complètement différencié [94]. Lors du développement de la rétine, une perte de pRb entraîne des divisions ectopiques [168].

Les protéines RB jouent un rôle de coordination des cellules direct ou indirect dans le développement. Ces actions se font principalement par interaction avec des protéines annexes impliquées dans la différenciation. Par exemple, lors de la neurogenèse, pRb séquestre Id2, un inhibiteur de la différenciation, et permet ainsi le maintien du timing normal de différenciation [169]. Dans les neurones, une mutation de pRb provoque des problèmes de développement des cellules de la rétine [170]. Chez la drosophile, une double mutation de Rbf et de la voie Hippo entraîne une

dédifférenciation dans l'œil [171]. Dans le muscle squelettique, pRb augmente aussi l'activité de MyoD, un facteur important pour la différenciation musculaire [172] (Figure 23).

### b. Sénescence

La sénescence est un arrêt irréversible de la prolifération lorsque la cellule sort définitivement du cycle cellulaire. Une cellule sénescence présente une morphologie particulière (cellule large et plate), une résistance accrue à l'apoptose et une expression de gènes altérée avec l'activation d'inhibiteurs du cycle cellulaire (CDKNs p21 et p16). Dans les cellules en culture, cet état est déclenché après un trop grand nombre de divisions ou par des stimuli comme des dommages à l'ADN ou un stress oxydant. *In vivo*, ce mécanisme est considéré comme étant une réponse antiproliférative au stress et un mécanisme suppresseur de tumeurs [110]. En effet, la sénescence est établie et maintenue par les voies p53 et p16-pRb, qui sont des voies de suppresseurs de tumeurs [173].

La voie p16-pRb est en général induite par des stimuli entraînant des dommages à l'ADN. p16 est une CDKN et peut donc maintenir pRb dans sa forme active hypophosphorylée en inactivant les CDKs. pRb peut alors inhiber E2F [142] (Figure 23). Des expériences de transcriptomique comparant des cellules en croissance, quiescentes ou sénescences en présence d'ARN interférents visant pRb ont montré que celle-ci régule une sous-population spécifique de gènes cibles de E2F pendant la sénescence, notamment des acteurs clés de la machinerie de réplication de l'ADN comme des éléments du complexe de pré-réplication [110].

La voie p16-pRb génère également des foyers d'hétérochromatine associés à la sénescence (SAHF, Senescence-Associated Heterochromatin Foci) qui permettent la répression de gènes de prolifération. L'apparition de ces SAHF coïncide avec une augmentation des marques répressives me-K9-H3 et le recrutement de pRb et HP1 $\gamma$ , marqueur de l'hétérochromatine, aux promoteurs cibles des E2Fs [174].

### c. Autophagie

L'autophagie est un processus cellulaire physiologique de recyclage des composants intracellulaires. Elle peut également être déclenchée en cas de stress cellulaire, et elle est considérée comme une voie « suppresseur de tumeurs » car certaines protéines suppresseur de tumeurs induisent l'autophagie, comme PTEN et p53. A l'inverse, certains oncogènes comme Bcl-2 et TOR inhibent l'autophagie [175]. Chez les souris, l'inactivation de régulateurs clés de l'autophagie entraîne la formation de tumeurs spontanées [176].

La réintroduction de pRb dans des cellules déficientes pour pRb entraîne une diminution de Beclin, un activateur de l'autophagie et une diminution de marqueurs de l'autophagie comme la formation d'autophagosomes. Ces mécanismes passent par la stimulation de gènes de l'autophagie

(ATG4A, ATG7) et la diminution de régulateurs négatifs (Bcl-2). Ces régulations se font grâce à l'inhibition de E2F1 par liaison de pRb (Figure 23), activée par la CDKN p16 [177].

A l'inverse, E2F1 antagonise l'autophagie médiée par pRb et déclenche l'apoptose. L'équilibre de la voie E2F/RB est essentiel au maintien de l'homéostasie tissulaire : un excès de pRb entraîne de l'autophagie et un excès de E2F1 conduit à l'apoptose [177].

#### d. Réparation de l'ADN

Les cassures double-brin de l'ADN peuvent être générées par divers stress cellulaires. Elles sont sources d'instabilité génétique. Deux mécanismes peuvent réparer ces cassures : la recombinaison homologue et la ligature d'extrémités non homologues. Ces deux mécanismes font intervenir des complexes protéiques aux extrémités lésées afin de réparer la cassure [178].

Chez les mammifères et les plantes, on trouve des complexes E2F1/pRb qui colocalisent avec les marqueurs de la réparation des cassures d'ADN double-brin, comme  $\gamma$ H2AX et BRCA1 [179]. La localisation de E2F1 dans ces foyers nécessite sa phosphorylation, qui va permettre la fixation de TopBP1 [180], une protéine impliquée dans les réactions des topoisomérases dans les cassures double brin (Figure 23). Cette interaction E2F1/TopBP1 est stabilisée par pRb [181].

De plus, des expériences de déplétion de pRb ou E2F1 ont montré qu'ils sont nécessaires au recrutement de nombreuses protéines nécessaires à la réparation, comme des recombinases (RAD51) [182] (Figure 23) ou des complexes impliqués dans la résection de l'extrémité (CtIP) [183].

Enfin, le remodelage de la chromatine aux sites de réparation est une étape nécessaire pour faciliter l'accès et la fixation des protéines. E2F1 et pRb permettent le recrutement d'histones acétyltransférase et d'enzymes de remodelage de la chromatine aux lieux de cassures afin de permettre l'accès à la machinerie de réparation [184,185].

Ce rôle de pRb/E2F1 n'a été montré que chez quelques mammifères (humain, souris) et plantes [186] (*Arabidopsis*, tabac). Cette spécificité est due à la présence d'un motif en N-terminal de E2F1 permettant une phosphorylation nécessaire au recrutement aux sites de cassures double brin. Ce motif n'est pas conservé chez d'autres espèces, notamment les invertébrés modèles comme la drosophile [178].

## C- LE COMPLEXE DREAM

Rbfl agit au sein du complexe de régulation transcriptionnelle dREAM. L'activité de ce complexe est difficile à prédire car, s'il semble majoritairement décrit comme répresseur transcriptionnel, il peut également activer l'expression des gènes. De plus, cette activité semble tissu-spécifique, puisque

certains, comme le gène *how*, peuvent être activés par le complexe dREAM dans certains tissus et réprimés dans d'autres.

Plusieurs mécanismes complémentaires peuvent expliquer la modulation de ces régulations. Un des leviers de ces régulations est la composition variable du complexe. dREAM est composé au maximum de neuf sous-unités qui s'organisent autour d'un cœur composé de protéines de la famille Mip (Myb-interacting proteins). Certaines, comme les facteurs de transcription E2F, contiennent des domaines de liaison à l'ADN qui permettent la reconnaissance de séquences cibles et la fixation du complexe sur les promoteurs associés. D'autres, comme Rbf1, sont des co-facteurs de transcription. Les sous-unités qui s'articulent autour du cœur ne sont pas toujours présentes, et la combinaison des facteurs présents influe sur l'activité du complexe. Ainsi, il semblerait que la présence de Rbf1 et dE2F2 conférerait une activité plutôt répressive, tandis que la présence du facteur de transcription Myb est associé à une activité plutôt activatrice. dREAM peut également recruter des protéines extérieures au complexe. Notamment, il est associé à des changements de conformation de la chromatine. S'il se lie préférentiellement au niveau des sites d'initiation de la transcription et à la chromatine compactée, de façon cohérente avec son rôle majoritaire de répresseur transcriptionnel, il peut interagir avec des *enhancers* en formant des boucles de chromatine ou en recrutant des protéines de liaison aux insulateurs. Il agit également sur les marques d'histones *via* le recrutement d'enzymes comme des histones désacétylases ou des méthyltransférases.

Comme les protéines de la famille RB, le complexe dREAM est particulièrement étudié pour son rôle de régulateur du cycle cellulaire. Cependant, il est également impliqué dans la régulation d'autres processus biologiques comme l'apoptose ou le développement. Comme la protéine Rbf1, les autres sous-unités du complexe dREAM sont remarquablement conservées à travers l'évolution ; ainsi que les mécanismes de régulation décrits précédemment. Dans la revue présentée ci-après dont je suis premier auteur, qui sera prochainement soumise au journal *Bioessays*, nous discutons de cette conservation et démontrons la pertinence de l'utilisation d'organismes modèles pour étudier les régulations médiées par le complexe dREAM. En effet, la conservation de ces mécanismes s'accompagne d'une moindre redondance de leurs acteurs, ce qui permet l'étude de modèles simplifiés transposables à des organismes plus complexes comme les mammifères. De plus, le complexe dREAM de mammifère est une cible d'intérêt dans les recherches de thérapies anti-cancéreuses et, dans cette optique, les études chez cet organisme se concentrent particulièrement sur son rôle dans la régulation du cycle cellulaire. Les organismes modèles non mammifères offrent une plus grande liberté d'explorer son implication dans d'autres processus. Le rôle du complexe dREAM dans l'apoptose constituait d'ailleurs un axe du projet de thèse à son origine.

## DREAM a little dREAM of DRM: model organisms and conservation of DREAM-like complexes

Marion Hoareau <sup>1,2</sup>, Aurore Rincheval-Arnold <sup>1</sup> and Isabelle Guénel <sup>1,\*</sup>

<sup>1</sup> LGBC, UVSQ, Université Paris-Saclay, 78000 Versailles, France

<sup>2</sup> Ecole Pratique des Hautes Etudes, PSL Research University, 75014 Paris, France

\*Corresponding author: isabelle.guenal@uvsq.fr (I.G.)

### Abstract

DREAM complexes are transcriptional regulators that control the expression of hundreds to thousands of target genes involved in cell cycle, quiescence, differentiation, and apoptosis. Many subunits and regulators associated with these complexes, and their plasticity, support the diversity of mechanisms by which they regulate their target genes but make their study difficult, especially in mammals. Since the discovery of the first dREAM complex (drosophila Rb, E2F, and Myb) in *Drosophila*, model organisms such as *Caenorhabditis elegans*, and plants allow diving deeper into the processes regulated by DREAM-like complexes. Here we review the conservation of DREAM complexes. We discuss model organisms' contribution to studying DREAM-mediated transcriptional regulatory mechanisms and their relevance in characterizing novel activities of DREAM complexes.

### Introduction

Cell-cycle progression is essential in all living organisms since it allows their development and growth. Therefore, the slightest dent in this well-oiled machine can have tragic consequences, such as cancer development. Guardians of the cycle progression are needed to ensure proliferation does not get out of control. One of those guardians is pRb, the first tumor suppressor identified <sup>[1]</sup>. pRb is known to regulate cell proliferation and differentiation notably via interaction with E2Fs transcription factors, which activate S phase genes. This interaction is conserved through evolution and can be found in various pluricellular eukaryotes such as mammals, invertebrates, or plants <sup>[2]</sup>. In mammals, pRb, together with p107 and p130, forms the pocket protein family. In *Drosophila*, RB family homologs are Rbf1 and Rbf2 <sup>[3]</sup>. The search for new partners for Rbf1 led scientists to identify several co-purifying proteins forming a transcription regulation complex named dREAM or MMB <sup>[4-6]</sup>. Since this discovery in early 2000, numerous genome-



wide studies revealed a broad spectrum of target genes, covering around a third of the *Drosophila* genome [7,8]. The importance of this transcription regulatory complex rapidly led to the identification of homologs in other species [9-11].

Model organisms such as *Drosophila* and nematode have considerably contributed to discovering DREAM complexes. Subsequently, the confrontation of data obtained in these models but also in mammals and plants was essential in defining the "Core complex". The evolutionary conservation of some of the activities of these complexes reinforces the idea that they can act as transcriptional activators or repressors. The large number of proteins in the complex and the diversity of transcriptional targets make its activity particularly difficult to study. Here again, the complementarity of the approaches taken and the accumulation of data in the different model organisms has made it possible to identify conserved mechanisms to understand how they regulate their targets. While in mammals, the DREAM complexes are widely described as controlling the cell cycle, working in other model organisms again highlights that these complexes may regulate other essential cellular processes. The objective of this review is not to exhaustively describe all the data reported on these complexes but to review the contribution of model organisms concerning their composition, the mechanisms they use to regulate transcription, and their role in cell fate.

### I- DREAM: A CONSERVED COMPLEX

The transcription regulation complex dREAM in *Drosophila*, is notably constituted of RB/E2F family. This observation led scientists to wonder if the structure of this complex was conserved in other organisms. It is the case in mammals, nematodes, and even plants (**Table 1**). DREAM-like complexes can take different forms, sometimes even within the same organisms, as described in *Drosophila* and mammals.

#### 1) The discovery of the dREAM complex in *Drosophila melanogaster*

Rb family has been extensively studied for its role in cell cycle progression. By interacting with the heterodimer E2F/DP (Dimerization Partner), Rbf1 control E2F-dependent transcription of cell cycle gene regulators. In *Drosophila melanogaster*, these proteins and their interactions are conserved. The Rb family is represented by two members, Rbf1 and Rbf2, whereas E2F and DP families count respectively two (dE2F1-2) and one members (dDP).

Purification of these *Drosophila* RB/E2F/DP complexes by chromatography revealed the presence of several copurifying proteins: Mip120, Mip130, RBF1, RBF2, dDP, dE2F2, Caf1p55, Mip40 and dMyb in nuclear embryo extracts [4]. These newly found complexes were called dREAM (drosophila Rb, E2F And Myb). The same year, another study focusing on Myb, a transcription factor known for its role in nuclear

activity, purified a complex named Myb-MuvB (MMB), containing the previously mentioned proteins and dLin52 [5]. A few years earlier, Mip130, Caf1p55, and Mip40 were also found as parts of the dMyb-containing complex (Mip: Myb-interacting protein) [6]. These studies aimed to find new partners for Rbf1 and dMyb.

Discovering dREAM complexes immediately raised questions about their roles. Given the transcription factor nature of some of their components (dMyb, dDP, dE2F2, Mip120, Mip130), a role in transcription was investigated. It was found that dREAM complexes bind compacted chromatin regions of polytene chromosomes which corresponds to transcriptionally silenced regions and nonacetylated histone H4 tails [4].

Interestingly, those nine subunits are not all necessary for dREAM complex action: depending on the target gene and its regulation, the composition of the complex is variable. It is organized around two « core » subunits: Mip120 and Mip130 which are essential for the complex integrity [7] (**Table 1**).

Organism	<i>Drosophila melanogaster</i>	<i>Caenorhabditis elegans</i>	<i>Homo sapiens</i>	Plants
<b>Complex</b>	<b>dREAM/ Myb-MuvB + tMAC</b>	<b>DRM</b>	<b>hDREAM/LINC /DREAM</b>	<b>DREAM</b>
<b>Protein homologs</b>	Rbf1/2	lin-35	p130/p107	RBR1
	dE2F2*	efl-1*	E2F4-5*	E2FB(+)/E2FC (-)*
	dDP*	dpl-1*	DP1*	DPA/DPB*
	Myb*	-	B-MYB*	MYB3R3 (-) / MYB3R4 (+)*
	Mip40	LIN-37	<b>LIN37</b>	-
	<b>Mip120*</b>	<b>LIN-54*</b>	<b>LIN54*</b>	<b>TCX5*</b>
	<b>Mip130*</b>	<b>LIN-9*</b>	<b>LIN9*</b>	<b>ALY2/ALY3*</b>
	Caf1p55	LIN-53	<b>RBBP4</b>	-
	Lin52	LIN-52	<b>LIN52</b>	-
-	-	-	BTE1	

**Table 1: dREAM members' homologs.** Minus signs indicate an absence of homologs. (+) and (-) indicate members that are, respectively, activators or repressors of transcription. Blue subunits are part of the MuvB core. Proteins possessing DNA-binding domains are signaled with an asterisk (\*).

## 2) DRM in *C. elegans*

In *C. elegans*, RTK- and Ras-mediated pathways induce vulva development, which follows a specific program. Mutations in those pathways are responsible for vulvaless (vul) phenotypes when downregulated and ectopic vulva (multivulva Muv) phenotypes when upregulated. Studies of these phenomena led to the

discovery of the synthetic multivulva (SynMuv) gene class, divided initially into class A and class B genes depending on the mutation they are carrying <sup>[12]</sup>. One of these class B genes, *lin-35*, has been described to encode the homolog of mammalian tumor suppressor *Rb* <sup>[13]</sup>. Moreover, homologous genes of *E2F* and *DP* coding the usual partners of pRb, are also conserved and have been described as class B SynMuv genes. They encode the EFL-1 and DPL-1 proteins which interact with LIN-35 *in vitro* <sup>[14]</sup>.

Digging further into this gene class, some other known Rbf1 partners were found: LIN-37, the homolog of Mip40, and LIN-53, the homolog of Mip120. By homology with the dREAM complex, a high molecular mass protein complex was discovered: DRM (DP, Rb, and MuvB) complex. It is composed of the « MuvB » core containing the homologs of Mip130 (LIN-9), Caf1p55 (LIN-53), Mip120 (LIN-54), Mip40 (LIN-37) and LIN-52 and the homologs of Rbf1 (LIN-35), DP (DPL-1) and E2F2 (EFL-1) <sup>[9]</sup>.

Unlike *the Drosophila melanogaster* dREAM complex, no Myb protein is involved in the DRM complex, and no Myb homolog is identified in *C. elegans*. However, the high conservation of Myb-like proteins throughout species and the conservation of Myb-binding domains of LIN-9 and LIN-52 proteins suggest that a Myb homolog was once present in *C. elegans* but lost during evolution. This hypothesis is supported by the fact that the expression of a *Drosophila* Myb in *C. elegans* causes a SynMuv phenotype <sup>[15]</sup>.

DRM complex mainly regulates transcription negatively. DNA targeting is mediated by the DNA-binding domains of LIN-54 and EFL-1/DPL-1 <sup>[16]</sup>. LIN-35 acts like a scaffold to mediate the association between MuvB and EFL-1/DPL-1, but its absence does not entirely eliminate their association with DNA. Even though their presence in the target promoters is mainly impaired, some genes are still bound by DRM. Loss of LIN-35 induces the up-regulation of numerous DRM target genes. However, it has yet to be discovered if LIN-35 is essential for repression or if this upregulation is an indirect effect due to less occupancy by other subunits <sup>[17]</sup>.

### 3) DREAM in mammals

After the discovery of *Drosophila melanogaster* and *Caenorhabditis elegans* dREAM/DRM complexes, the existence of homologs for most of the subunits described in these organisms led researchers to wonder if such complexes exist in mammals. However, a different level of complexity occurs in mammals due to the redundancy of proteins. For example, the pocket proteins family, which includes three proteins in mammals: pRb, p107, and p130, is represented by only two proteins in *D. melanogaster* (Rbf1-2) or a single one (LIN-35) in *C. elegans*. Immunoprecipitation experiments first showed that a complex was formed by the homologs of dREAM/DRM complexes <sup>[10,11]</sup>. However, not all pocket proteins family participate in the mammals' complex called DREAM (or LINC): p130 is the main one involved <sup>[10]</sup>. p107 can also be found

in the complex <sup>[11]</sup>, but its role is not entirely clear, though it seems that it might replace p130 when it is mutated. If p130 and p107 are inactive, there is no DREAM complex or repression of its targets <sup>[18]</sup>. The case of pRb seems more complex. The apparent absence of a compensation of p130/p107 double mutation and the recurrent inability to co-immunoprecipitate pRb with core proteins such as LIN54, LIN9, or LIN37 <sup>[10,19]</sup> suggest that pRb is not involved in the DREAM complex. However, some other studies were able to associate pRb with a complex purified in human cells <sup>[11,20]</sup>.

Presence of Myb protein is a divergent point between *D. melanogaster* and *C. elegans* dREAM/DRM complexes. However, as established in *Drosophila* and *C. elegans*, core proteins include LIN9, LIN37, LIN52, LIN54, and RBBP4 and form the complex called MuvB core. In mammals, "DREAM complex" refers specifically to a complex where E2F and pRb family proteins, usually p130, are added to the MuvB core whereas the complex named MMB (Myb-MuvB) is formed of the MuvB core added of Myb <sup>[21]</sup>. *Drosophila* Myb protein has three homologs in mammals: A-MYB, B-MYB, and C-MYB. Like pRb family proteins, not all three are involved in DREAM complex activity: only B-MYB is co-immunoprecipitated with Mip130/LIN9 <sup>[22]</sup>, Mip40/LIN37, and Mip120/LIN54 but not with p130 immunoprecipitations <sup>[10]</sup>. Mass spectrometry experiments also detected minor A-MYB signal in MMB complexes in human cells <sup>[10]</sup>. To sum up, in *Drosophila*, dREAM complexes Rbf1 and Myb can be found in the same complex but in mammals DREAM and MMB complexes are mutually exclusive (**Figure 1**). These two distinct complexes seem tightly linked to the progression of the cell cycle where DREAM and MMB are associated with specific cell cycle phases.

Compared to *Drosophila*, mammalian DREAM/MMB complexes bind fewer genes: a few hundred instead of one-third of the *Drosophila* genome. This lack of targets is thought to be due to DREAM/MMB being largely involved in cell cycle regulation, whereas dREAM (and DRM) complexes play roles in other cell processes, such as apoptosis or differentiation <sup>[18]</sup>. However, whatever the organism, regulations mediated by these complexes are mainly repressive.

#### 4) DREAM-like in plants

In plants, homologs of dREAM members exist. In dicotyledons, only one pRb homolog can be found: Rb-Related (RBR1, closer to pRb). In monocotyledons (maize, rice), there are two main pRb subfamilies with members involved either in negative regulation of cell cycle or endosperm development <sup>[23]</sup>. Like in other organisms, its association with E2F and DP families' proteins mediates RBR1 action. There are 6 E2F, but only three of them (E2FA, B, and C) possess a dimerization domain similar to those of other organisms <sup>[23]</sup> and two DP (DPA and B)<sup>[24]</sup>. Using pull-down and co-immunoprecipitation experiments, the interaction

study between RBR1, E2FB or C, DPA or B and homologs of the core proteins such as Mip130/LIN9 (ALY2/3) or Mip120/LIN54 (TCX5) revealed the existence of DREAM-like repressor and activator complexes. No corresponding proteins were found for Mip40/LIN37 and LIN52<sup>[23,25,26]</sup>. However, a plant-specific subunit has been identified by co-immunoprecipitation of plant DREAM complex subunits and mass spectrometry: BTE1/DCR2<sup>[27,28]</sup>.

These complexes also contained Myb homologs<sup>[23,25,26]</sup>. There are hundreds of Myb proteins in plants, but only a few are similar to mammals and *Drosophila*. Due to a triple repetition of the DNA-binding domain, this group of proteins is called R1R2R3-Myb or Myb3R. Some of them are transcription activators (Myb3R4), and some are repressors (Myb3R3 and Myb3R5)<sup>[25]</sup>. Myb3R1 was first believed to be an activator, but this activation is only done in the presence of Myb3R4. When in the presence of MYB3R3 and MYB3R5, MYB3R1 represses transcription. This observation suggests that MYB3R1 might be a switch between activated and repressed states<sup>[26]</sup>. MYB3R3 and MYB3R4 are found in those DREAM-like complexes, but never at the same time. They are associated respectively with E2FC (repressor) and E2FB (activator)<sup>[25,26]</sup>. DREAM-like complexes in plants thus seem also able to activate or repress transcription depending on the recruited subunits.

One last subunit found in DREAM-like complexes in plants is CDKA. It could regulate the complex assembly and composition through phosphorylation of the subunits<sup>[23,26]</sup>.

### 5) What about other model organisms?

In zebrafish *Danio rerio* and xenope *Xenopus tropicalis*, orthologs exist for core proteins Mip120/LIN54 and Mip130/LIN9, as well as for pRb family and Myb proteins. Despite those orthologs, no study was found on dREAM-like complexes in xenopes. Zebrafish seem to raise more interest since one study describes a role of zebrafish B-MYB in activating gene expression during G2/M phases<sup>[29]</sup>. Some authors seem to accept this as evidence of the existence of dREAM-like complexes<sup>[30,31]</sup>. However, only one proof of interaction between Myb and a core dREAM protein, Mip130/LIN9 have been reported<sup>[30]</sup>. This interaction may hint at the presence of a dREAM-like complex in zebrafish.

The *Schizosaccharomyces pombe* and *Saccharomyces cerevisiae* cases are easier since no ortholog has been found for Mip120/LIN54, Mip130/LIN9, or Rbf1. It seems safe to say that dREAM-like complexes do not exist in yeasts.

## II- A TRANSCRIPTION REGULATION COMPLEX

### 1) Are DREAMs transcription activators or repressors?

If dREAM was initially described as a repressor complex, RNA interference and chromatin immunoprecipitation (ChIP) studies showed that DREAM complexes can bind thousands of promoters and activate, repress or even not have an effect on their transcription.

The transcription regulation activity of DREAM complexes was discovered through the inactivation of some of its subunits through mutants or RNA interference techniques in *Drosophila* [4,32], *C. elegans* [16], and mammal cells [10]. When discovered in *Drosophila*, dREAM seemed to be a repressor complex mainly. However, in *Drosophila* cells, using specific RNAi expression to downregulate all members of the dREAM complex (except Caf1/p55), it has been shown that more than 600 genes were upregulated when one or more subunits were absent. Interestingly, almost as many genes were downregulated in the same conditions, demonstrating that dREAM can activate or repress transcription [7]. Similarly, in *C. elegans* embryos and germline cells, mutants of LIN-54 led to the upregulation of more than 700 genes and the down-regulation of 550 others [16]. Finally, in mammal cells, the expression of RNAi against LIN9 or LIN54 led to the inhibition of the repression of DREAM-regulated genes. However, due to the higher number of homologs in these cells, it is harder to decipher the effect of some DRM complex regulators' downregulation. For example, the loss of p130 can be compensated by p107, whose transcription is enhanced in these conditions [10].

To go further in the characterization of DREAM complexes target genes and more finely identify DREAM-binding sites, different chromatin immunoprecipitation techniques have been done. ChIP-on-chip experiments using antibodies against almost all subunits allowed the identification of more than 3500 binding peaks in *Drosophila* cells. However, only 25% of the proximal-promoter-bound of these genes show an expression change in RNAi expression context for dREAM subunits [7].

LIN-54 binds to almost 2000 locations in nematodes, with 1572 protein-coding genes. Part of those targets is conserved through evolution. 3015 orthologous gene pairs are shared by *D. melanogaster* and *C. elegans*. Around 1200 of them are bound by *Drosophila* Mip120, 650 by the nematode homolog LIN-54, and 327 by both species [16]. Among these numerous targets, DREAM complexes can regulate the transcription of their own subunits [7,10,16].

Transcriptional DREAM complexes-mediated regulation, therefore, appears complex. Far from being only a transcription repressor complex, it can also activate transcription. The presence of different DNA-binding subunits allows the targeting of thousands of promoters. This high number of targets combined with

different effects on transcription makes it very difficult to predict DREAM complexes' activity at a given promoter.

### 2) How do DREAMs regulate transcription?

If ChIP experiments could identify promoters bound by the DREAM complex and its homologs, they give little information on the mechanisms of regulation mediated by the complex. Since none of the dREAM subunits carry enzymatic activity, the ones that do not possess DNA-binding domains (Rbf1, Rbf2, Caf1p55, Mip40, Lin52) might play a role in stabilizing the complex or recruiting other partners such as chromatin-modeling enzymes. Indeed, DREAM complexes regulate transcription through multiple processes, some of which require the recruitment of partners. The multiple facets of these regulations add yet another complexity to how this complex works.

#### a) Influence of complex composition

The first and most obvious way for DREAM to regulate transcription is through the activity of its subunits. The composition of the complex is decisive for transcription regulation.

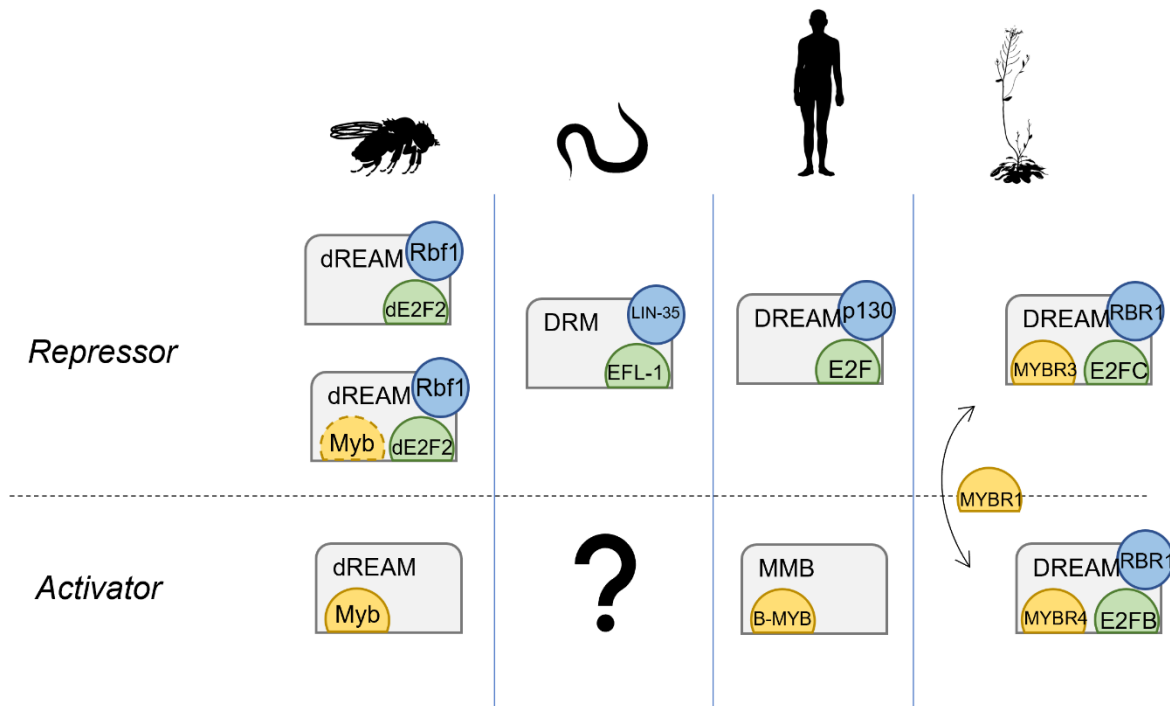
Firstly, the complex composition determines the target genes and the binding location of DREAM. DNA-binding subunits, such as LIN-54 and E2F2, have specific target sequences that allow the complex to bind to a wide range of genes. ChIP-qPCR experiments on *lin-54* worms mutants has shown that LIN-54 is involved in the recruitment of DRM at promoters. Moreover LIN-54 binding is enriched at sequences which contain binding motif for LIN-54 adjacent to E2F-binding motif <sup>[16]</sup>. E2F2 and Myb seem to be the main driving force of DREAM binding in mammals <sup>[7,10]</sup>. In *C. elegans*, where there is no Myb, DRM association with chromatin is done primarily through the DBD domains of E2F-DP and the tesmin domains of LIN-54 (Mip120) <sup>[16,17]</sup>. However, the DNA-binding domain is not the only factor of DREAM binding: chromatin association is impaired without LIN-35 (pRb) <sup>[17]</sup>.

Secondly, the complex composition determines target genes activation or repression. The subunits present in DREAM are sometimes different at promoters. RNA interference studies have been done in *Drosophila* cells to determine which subunit is necessary to regulate which targets. Mip120 and Mip130 have been identified as "core subunits" necessary for all target genes. Genes were then classified into multiple categories depending on which subunits were essential for their regulation and whether they were activated or repressed <sup>[7]</sup>.

From dREAM complex identification in *Drosophila*, Myb was considered a silent member in repression. This hypothesis was based on the generally activating role of Myb described in the literature and the absence

of de-repression in a Myb-knock-down context <sup>[5]</sup>. Georgette *et al.* classification brought to light the antagonistic effect of Rbf1-2/E2F and Myb subunits: Rbf1-2/E2F are only present in repressive complexes, and activating complexes always contain Myb. Complexes including at the same time Rbf1-2 and Myb are repressive, suggesting a dominant effect of Rbf1-2 over Myb <sup>[7]</sup>. Therefore, in *Drosophila*, three types of dREAM complexes can be found (**Figure 1**) <sup>[7,33]</sup>:

- E2F and partners (Rbf1-2, DP)-containing complexes that are free from Myb and generally repressive
- Myb-containing complexes that are free from E2F and its partners and generally activating
- E2F and Myb-containing complexes that are generally repressive.



**Figure 1 Antagonism of Rb-family and Myb-family members.** Composition of transcription repressing and transcription activating complexes in (from left to right) *Drosophila melanogaster*, *Caenorhabditis elegans*, *Homo sapiens*, and *Arabidopsis thaliana*. In mammals, p107 can also be found in DREAM complex instead of p130. Grey boxes represent the core and/or other members of the complex.

This antagonistic effect can be observed in other species (**Figure 1**). In human cells, at G0 and early G1 phase, p107/p130-DP-E2F4 and the MuvB core are associated into DREAM complex <sup>[19]</sup> and contribute to repress cell cycle genes <sup>[10,34]</sup>. During late S phase, after the release of p107/p130-DP-E2F4, B-MYB is recruited to the MuvB core to form the MMB complex<sup>[19]</sup>. The MMB complex then does not bind to E2F



target genes anymore and activates genes implicated in the G2/M phase <sup>[34]</sup>. Those activating complexes are kept during the G2/M phase and can recruit additional transcription factors <sup>[35]</sup>. After cell division, DREAM complexes, which contain RB/E2F complexes, reassemble thanks to the DYRK1A kinase. DYRK1A allows the binding of p130 to the MuvB core and phosphorylates LIN52 on serine 28, which is required for DREAM assembly. Overexpression of B-MYB or expression of LIN52-S28 disrupts DREAM assembly and promotes MMB formation <sup>[36]</sup>. Therefore, Myb and RB/E2F are not found in the same complexes in mammals.

### *b) DREAM binding sites*

#### DREAM complex binding sites on chromatin is another clue on how it affects transcription

dREAM mainly binds in transcriptionally inactive regions, which supports its repressive action. In *Drosophila melanogaster*, polytene chromosomes staining with antibodies against dREAM subunits revealed their presence on repressed chromatin. The complex staining does not colocalize with phosphorylated polymerase II, a marker of actively transcribed regions. On the contrary, Rbf1, Rbf2, and Mip130 subunits bind to nonacetylated histone H4 tails and not to acetylated ones, showing another repressed chromatin binding preference. Korenjak *et al.* proposed that dREAM complexes bind deacetylated histones to maintain them in a repressed state and protect them from further modification <sup>[4]</sup>.

Numerous genomic studies have been done to decipher dREAM complex role in transcription regulation. These studies revealed that dREAM is bound to promoter-proximal location in gene-dense regions of different organisms: 86% of dREAM-bound regions are within 1 kb of transcription start sites (TSS) in *Drosophila* cells <sup>[7]</sup>, 78% between -1000 to +100 bp from TSS in nematode embryos. In contrast in these embryos, 22 % of the binding sites left were located in inactive regions such as introns or intergenic regions <sup>[17]</sup>. The transcriptional role of dREAM and its homologs seems to be carried out through promoter binding. The promoter-proximal region is associated with a higher nucleosome density downstream from the TSS compared to genes that lack the CHR binding site or constitutively expressed genes <sup>[37]</sup>. A recent study showed that MuvB, the core of mammals' DREAM/MMB complexes, associates with +1 nucleosomes downstream of the TSS. This association is correlated with gene repression. MuvB could stabilize nucleosome position by forming a bridge between nucleosome and DNA. This mechanism could contribute to repression by inhibiting chromatin remodeling or the activity of transcription proteins such as polymerases <sup>[38]</sup>.

Finally, dREAM binding at TSS seems to play a part in regulating divergently paired genes (DPG). DPG are genes transcribed in the opposite direction with close TSS (less than 1000 bp). In *Drosophila*, they

represent around a third of genes. By interacting with insulator-binding proteins, dREAM has an enhancer-blocking activity required to repress a subset of target genes <sup>[39]</sup>.

*c) DREAMs recruit other partners to modulate transcription.*

If binding sites and composition of DREAM complexes are factors of transcription regulation, DREAMs can also recruit partners to modulate gene expression. Those partners can act directly on promoters or at a higher level by modifying the chromatin state.

L(3)MBT, a substoichiometric subunit

Despite a variable composition that can go up to nine subunits, DREAMs also recruit or interact with other proteins to regulate transcription. One of these proteins, L(3)MBT (lethal (3) malignant brain tumor), is sometimes considered an accessory subunit. It is a transcription repressor, capable of binding chromatin <sup>[40]</sup> and found to be associated with transcription regulators such as insulator-binding proteins <sup>[41]</sup> and H4K20 methylations in *Drosophila* <sup>[42]</sup>.

L(3)MBT was also identified by a mass spectrometry approach destined to find Mip120 partners in large protein complexes during the discovery of the dREAM complex <sup>[5]</sup>. This interaction was confirmed by co-immunoprecipitation and GST-pulldown experiments a few years later <sup>[7,33]</sup>.

If colocalization studies suggest L(3)MBT binding to chromosomes does not necessitate dREAM, Mip120 is still necessary for L(3)MBT correct localization on chromatin. L(3)MBT is associated with the core proteins of dREAM, Mip120, and Mip130, at some specific loci, such as the *Piwi* one in salivary glands <sup>[33]</sup>. In mammals, if no interaction has been shown between DREAM core proteins and L(3)MBTL1, the homolog of L(3)MBT, this protein nevertheless binds to E2F target sites <sup>[43]</sup>. This binding could involve some recruitment through DREAM.

Transcription factors recruitment

Even if some of the DREAM subunits are transcription factors, they are not the only ones that can intervene in the transcription regulation activities of the complex. Promoter-binding studies such as chromatin immunoprecipitation have revealed enrichments for transcription factors binding sequences. As expected, the most represented are the E2F-binding sequences, but they are not the only ones. In mammalian cells, the prediction for enriched sequence motifs identified Myb, but also CREB and NRF2-binding sequences, suggesting that these transcription factors can cooperate with the DREAM complex to regulate the transcription of their targets <sup>[10]</sup>. Likewise, Goetsch *et al.* compared in *C. elegans* the DRM peak regions

they found by ChIP to high occupancy target (HOT) regions<sup>[17]</sup>. HOT regions are genomic regions with many transcription factor binding sites<sup>[44]</sup>. They observed that around a third of their high confidence DRM peaks correspond to HOT embryonic regions, suggesting interaction with other transcription factors<sup>[17]</sup>. Those examples show potential partners for DREAM complexes. Some others have shown an interaction between DREAM/MMB and two transcription factors or co-activators: FOXM1 and YAP<sup>[21]</sup>. Since those two proteins are mammalian transcription activators and interact with Myb, it is more accurate to consider their binding to the activating B-MYB-containing complex, MMB. Indeed, after being identified in motif enrichment analysis of MuvB binding, subsequent co-immunoprecipitation confirmed the interaction of FOXM1 with it. In mammals, FOXM1 can be considered an MMB subunit, much like B-MYB: it is recruited to activate complexes, and FOXM1 binding to MMB is essential for the expression of some mitotic genes<sup>[35]</sup>. The binding mechanism of FOXM1 is not yet very clear, but some data suggest that mainly B-MYB<sup>[35,45]</sup> or LIN9<sup>[46]</sup> might be involved. This binding cannot be extrapolated to other organisms since no homolog of FOXM1 has been identified in *Drosophila* or *C. elegans*<sup>[34]</sup>.

Another example of transcription factor recruitment is YAP, a co-activator of transcription acting downstream of the Hippo pathway<sup>[47]</sup>. The pool of genes regulated by YAP includes G1/S, mitosis, and cytokinesis genes like E2F. However, YAP is known to bind enhancers and might cooperate with DREAM-bound promoters through chromatin looping<sup>[48,49]</sup>. Since mitosis and cytokinesis genes are not E2F but MuvB targets, YAP cooperation with the MMB complex was investigated. Studies show that YAP can interact with Myb and LIN9/Mip130, MMB subunits, probably through chromatin looping. MMB complex is required for YAP-mediated expression of G2/M genes<sup>[49]</sup>. Some notable targets of YAP are Myb and FOXM1 themselves<sup>[21]</sup>.

Even though the transcription factors recruited are not the same, the enrollment of external partners seems conserved through evolution.

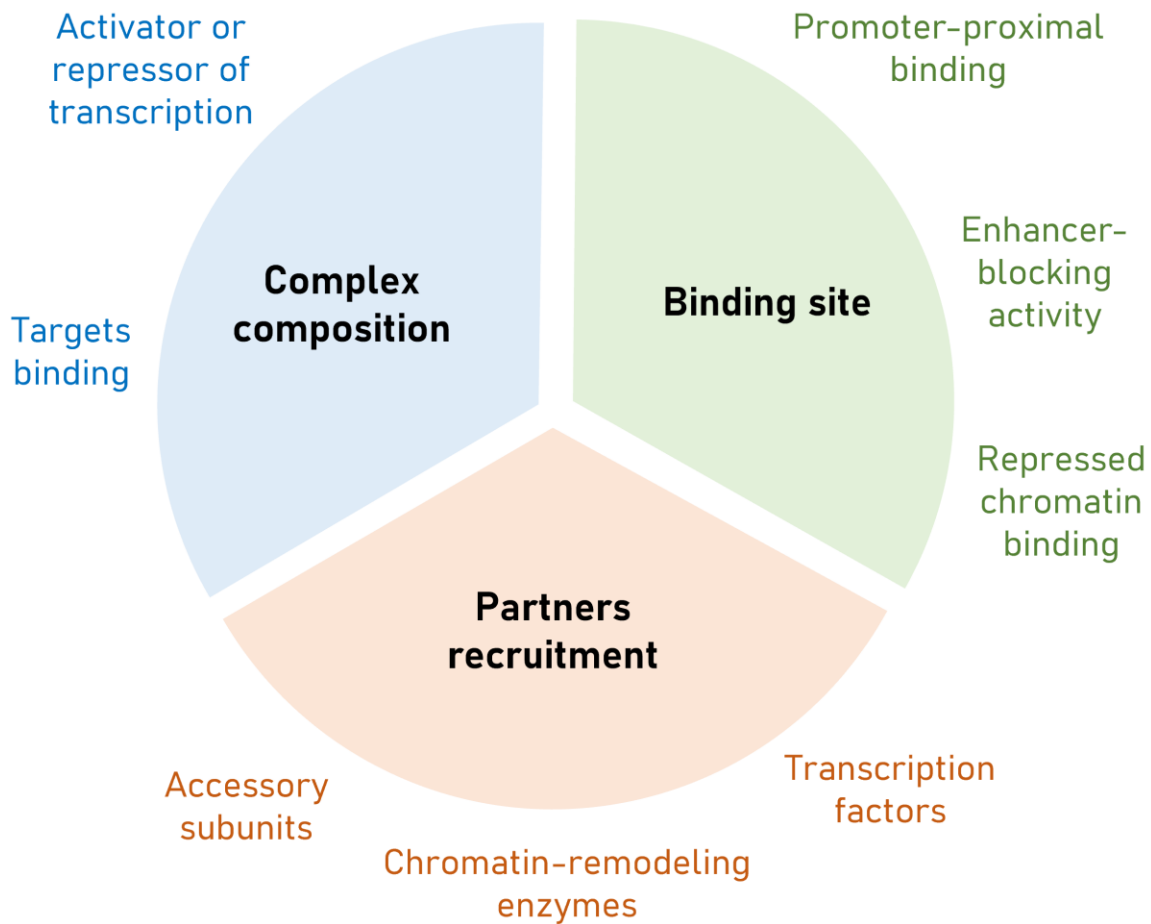
### Chromatin state and histone marks

Besides binding histone tails, the DREAM complex is able in invertebrate and in mammals to recruit some proteins to modify the chromatin state or apply marks on histones.

Insulators are some DNA elements that block the communication between an enhancer and a promoter. Dyson's lab has shown in *Drosophila* that the dREAM complex could co-operate with insulator-binding proteins. Among them, Beaf-32 is a protein, which binds frequently to TSS close to dREAM-bound regions. Moreover, CP190, a co-factor of insulator complexes is able to physically interact with dE2F2. dREAM genetically interacts with both of these factors<sup>[39]</sup>.

*lin-36* and *lin-15B* are SynMuv class B genes, which were first identified as G1 regulators in *C. elegans* <sup>[50]</sup>. Later they were identified during a RNAi screen, as regulators of DREAM target reporter gene <sup>[51]</sup>. LIN-36 and LIN-15B proteins contain a THAP domain, which is an atypical zinc-finger-DNA-binding domain. They interact with DRM through LIN-35 (pRb), at different target regions and with different mechanisms. LIN-36 mutants are known for causing cell-cycle defects, and, unsurprisingly, LIN-36 shares cell-cycle and cell-division targets with LIN-35. LIN-36 seems to function with DRM to recruit or maintain the repression-associated histone variant H2A.Z (HTZ-1 in *C.elegans*) on gene bodies. Around half of DRM targets characterized by this mark overlap with LIN-36 targets <sup>[51]</sup>. A link between the H2A.Z variant and the dREAM complex has also been found in *Drosophila* <sup>[52]</sup>. On the other hand, LIN-15B seems more specific to the germline. Mechanistically, DRM associated with LIN-15B directs the MET-2 methyltransferase to the target genes to apply repressive H3K9me2 marks <sup>[51]</sup>. LIN-15B and H3K9 methylations are also associated in DRM -mediated chromatin compaction during development <sup>[53]</sup>.

DREAM also interacts with histone deacetylases (HDAC) which remove the activating acetyl marks on histones. In *Drosophila*, Rpd3, HDAC1 homolog, has been identified as a dREAM complex partner since its discovery <sup>[5]</sup>. It could be involved in the repression of *hid* gene in apoptosis <sup>[54]</sup>. Caf1/p55 dREAM subunit is also part of the chromatin remodeling NURF (NUcleosome Remodeling Factor) complex, which suggests a possible interaction between those two complexes <sup>[5]</sup>. In mammals, RBBP4, a MuvB core subunit, interacts with HDACs or HDAC-containing complexes (NURD :NUcleosome Remodeling and Deacetylase) <sup>[34,55]</sup>. SIN3B interacts with numerous DREAM subunits in human and mouse quiescent cells and is required for target repression <sup>[56]</sup>. SIN3B is a scaffold known for assembling large repressive complexes containing HDACs <sup>[57]</sup>.



**Figure 2 DREAM complexes modulate transcription through three mechanisms:** the changing composition of the complex, its binding site on chromatin, and the recruitment of protein partners.

DREAMs are thus able to regulate transcription through multiple mechanisms. We propose to classify them into three groups (**Figure 2**):

- **Complex composition:** the DNA-binding proteins present in the complex bring DREAM complexes to their target sequences. Subunits also influence the activating or repressing action of the complex, as shown by the antagonism of Myb and Rbf1 subunits in *Drosophila*.
- **Binding site:** DREAMs can regulate transcription by binding promoters or enhancers. It has also been suggested that DREAM can maintain chromatin in a repressed state.

- Partners recruitment: DREAMs can also modulate transcription by recruiting external partners, such as accessory subunits, chromatin-remodeling enzymes or histone modifications, and other transcription factors.

Those mechanisms are conserved despite the actors involved being different.

### III- AN ESSENTIAL COMPLEX FOR BIOLOGICAL PROCESSES

Numerous genome-wide studies were led to determine the complex function after discovering dREAM complexes in *Drosophila* and its homologs in *C. elegans* and mammals. The genome-wide studies conducted on the complexes identified, through Gene Ontology analysis, processes in which DREAMs are involved.

#### 1) A crucial role in the cell cycle

In mammals, the DREAM complex binds mainly to early or late cell cycle genes <sup>[10,18]</sup>. This result is not surprising since RB family proteins, part of DREAM, are regulators of the cell cycle G1/S transition. This cell cycle role is confirmed mainly by loss-of-function studies. Depletion of LIN9/Mip130 shows that it is required for cell cycle progression and genomic stability <sup>[58]</sup>, as it leads to mitotic defects and G2/M arrest <sup>[11,19,59]</sup>. Even before the dREAM complex description, Mip proteins were found along with Myb and Caf1/p55 in a database search in *Drosophila* to identify proteins that bind to replication origin proteins, revealing a role in replication control <sup>[6]</sup>.

Mutually exclusive mammalian DREAM and MMB complexes described in the first part of our review are linked to their role in the cell cycle. During quiescence, the core associates with p130/E2F4-5 to form the DREAM complex and repress expression of some E2F-regulated genes. *B-MYB* gene be found among these targets, which explains its absence during quiescence <sup>[10,18]</sup>. The formation of the pocket protein-containing complex, DREAM, is required to enter and maintain a quiescent state <sup>[34]</sup>. It is made by the interaction of p130 (or p107) and LIN52, only possible when serine 28 of LIN52 is phosphorylated by the DYRK1A kinase <sup>[36]</sup>. In G<sub>0</sub>, DREAM is bound to E2F target genes, which are repressed. Around half of the DREAM targets are upregulated in the S phase compared to G<sub>0</sub>, which is concordant with its main repressor function <sup>[10]</sup>; When cells go out of G<sub>0</sub> to engage in the cell cycle, p130 is dissociated from E2Fs and the MuvB core. DREAM subunits are phosphorylated by cyclin-dependent kinases such as CDK2 and CDK4. This phosphorylation causes the dissociation of the complex from target promoters <sup>[11,60]</sup>. After p130/E2Fs dissociation, the MuvB core binds DNA thanks to CHR (Cell cycle genes Homology Region) sequences mainly present on LIN54 <sup>[18]</sup>. B-MYB is then able to be produced and recruited to the MuvB

core to form the MMB complex <sup>[18]</sup>. FOXM1 is then recruited <sup>[18,34]</sup> and MMB complex activate key mitotic genes like *Cyclin B1* for entry, *BUB1* for mitotic spindle checkpoint, or *Aurora-A* for spindle assembly, among others <sup>[61]</sup>. Interestingly, *Myb* and *FOXM1* are repressed DREAM targets during G0/G1 <sup>[61,62]</sup>.

The tumor suppressor p53 plays an important role in mammalian DREAM and its different forms also interact with the p53 pathway. In stress conditions as induction of DNA damage, activation of the p53 pathway induces the dissociation of B-Myb and the recruitment of p130/E2F4 to the MuvB core, forming and stabilizing the DREAM complex to promoters <sup>[63,64]</sup>. p53 plays also a crucial role in senescence, a stable cell cycle arrest state due to physiological aging or particular stress conditions <sup>[65,66]</sup>. By using a senescence bypass assay of human breast fibroblast conditionally immortalized, it was shown that respectively 49%, 16% and 22% of downregulated genes during senescence are some target of DREAM, RB-E2F and MMB-FOXM1. In this system, the expression of Myb, FOXM1, and a non-phosphorylated form of LIN52-S28 is sufficient to bypass senescence <sup>[67]</sup>.

### 2) Regulation of programmed cell death

If mammals' DREAM complex seems to be mainly involved in cell cycle regulation, *Drosophila* dREAM has a way larger spectrum of target processes. Chromatin immunoprecipitation studies revealed that around a third of all *Drosophila* genes' promoters were bound by dREAM members <sup>[7]</sup>, suggesting a role in controlling cellular processes beyond the cell cycle.

Apoptosis is one of those processes in which the dREAM complex is involved. In *Drosophila*, observation of an apoptosis dependent of Myb-led Rovani *et al.* to investigate whether or not this cell death was a direct effect or a consequence of excessive proliferation. They identified a direct induction of apoptosis by Myb in neural precursor cells at the posterior wing margin, mediated by the pro-apoptotic Grim. dE2F2 is also involved in this process. Mip130 is required for complex stabilization since Myb levels are lower in cells with a knockout mutation of Mip130 <sup>[68]</sup>. The involvement of dE2F2 and Myb is surprising, given the often admitted antagonistic role of E2F and Myb in the dREAM complex <sup>[7]</sup>. Our team found out that Rbf1 pro-apoptotic activity in the developing wing depend on a repression of the anti-apoptotic gene *buffy* and an activation of the pro-apoptotic gene *how*. Mip120 and Mip130, dREAM complex core proteins, are required for this apoptosis and participate in the *buffy's* transcriptional regulation but not in *how's* one. In contrast, Myb seems to have an antagonistic effect since its overexpression is enough to inhibit Rbf1-induced apoptosis. Myb activates *buffy* transcription and decreases the translational activation of *how* induced by Rbf1/E2F2 <sup>[69]</sup>. However, in the eye-antenna disc, the dREAM complex containing Rbf1 can inhibit

apoptosis through the repression of the pro-apoptotic *hid* <sup>[54]</sup>. dREAM can thus have opposite roles in apoptosis, which seems to depend on the tissue of interest.

dREAM complex role in apoptosis is not limited to *Drosophila*. A mutation screen aimed at identifying new programmed cell death genes in *C. elegans* revealed the role of *dpl-1*, DP homolog. Considering its partnership with pRb and E2Fs, Reddien *et al.* also investigated the role of their homologs, *lin-35* and *efl-1*, in programmed cell death. They discovered their pro-apoptotic role along with other DRM members: LIN-52 and LIN-37 (Mip40). LIN-53 (Caf1-p55) and LIN-9 (Mip130) are not involved in this process, but we can hypothesize that, as in *Drosophila*, DRM complex composition can differ depending on the target genes <sup>[70]</sup>. Another study characterized the involvement of pRb/E2F/DP homologs in apoptosis in *C. elegans*. They observed the implication of other synMuvB genes in DNA damage-induced apoptosis in unpublished data <sup>[71]</sup>.

### 3) Development and differentiation: less studied roles for DREAM

The primordial role of DREAM in the cell cycle implies a role in organisms' development. Mip130/LIN9 is required for early mouse development in mammals, as null mutants die during embryogenesis. The mitotic genes' regulator role of Mip130/LIN9 probably explains this embryonic lethality <sup>[72]</sup>. This hypothesis is supported by the fact that *C. elegans* DRM members are all synMuv genes. This family of genes was identified as genes required for the development of reproductive organs <sup>[14,73]</sup>. Korenjak *et al.* proposed in 2004 that *Drosophila* dREAM could also have a role in the reproductive system by regulating sex-associated targets whose misexpression could cause changes in cell fates <sup>[4]</sup>. This proposal is based on the fact that most dREAM members are conserved in the two organisms and that dE2F2 is known to be required to express genes important in gametogenesis <sup>[74]</sup>. In addition, mutants for dE2F2 also have fertility problems <sup>[75,76]</sup>. In *Drosophila*, a particular dREAM complex is found in the testis: the tMAC (testis Meiosis Arrest Complex). tMAC is considered a kind of dREAM complex as it is a transcription regulation complex constituted of two common subunits (Mip40 and Caf1-p55), three homologous subunits, and two others <sup>[77]</sup>. DamID data reveal that dREAM and tMAC interact in the testis: dREAM represses genes not expressed in the testis. When spermatocyte differentiation occurs, tMAC activates the ones needed by recruiting Mip40 <sup>[78]</sup>. Finally, dREAM is also involved in oogenesis. The absence of Mip120 causes an arrest of oogenesis associated with chromosome defects <sup>[79]</sup>.

Spermatocyte differentiation is not the only example of dREAM being involved in differentiation. dREAM complex is required at differentiation-related dE2F2 targets genes in *Drosophila* cells during the S phase. At these promoters, dREAM is not necessary to target genes: the disruption of dREAM using RNA



interference against Mip40 did not prevent dE2F2 from binding these promoters. Instead, dREAM seems to stabilize the complex during the S phase. It also protects Rbf1 from phosphorylation and thus keeps it in its active form <sup>[80]</sup>.

### Conclusion

DREAM complexes represent a highly conserved transcriptional regulator complex family. At first glance, they could be classic repressor complexes but digging into their regulations, they appear to have a way larger spectrum of action. Due to the richness of their targets and their numerous ways of regulating transcription through different mechanisms, DREAM-mediated regulations are hard to predict. All the interactions and partners involved in transcription modulation make DREAM complexes crucial platforms for regulating essential processes such as cell cycle and apoptosis. Thus, DREAM regulators are ideal targets for cancer apparition. Given their crucial role in cell cycle progression, Myb and FOXM1 are often overexpressed in tumors. High levels of these proteins and their targets are associated with poor prognosis in many cancers <sup>[81,82]</sup>. Studying the mechanisms of their roles in proliferation could provide clues for targets and strategies for cancer therapy.

Oddly enough, this possible anti-tumoral aspect did not raise interest in mammals' DREAM/MMB roles in apoptosis regulation. However, the *Drosophila* and *C. elegans* models are essential to investigate other possible effects of DREAMs on important biological processes in simplified models. The study of DREAM complexes is thus an example of the power of model organisms and the many possible regulations they can help unravel.

In addition to the therapeutic perspective, the vast pool of DREAMs target genes and the complexity of their regulation leaves many unanswered questions and makes room for exciting research for many years to come.

### Bibliography

1. Huang, H.-J. S., Yee, J.-K., Shew, J.-Y., Chen, P.-L., Bookstein, R., Friedmann, T., Lee, E. Y.-H. P., & Lee, W.-H. (1988). Suppression of the Neoplastic Phenotype by Replacement of the RB Gene in Human Cancer Cells. *Science*, *242*(4885), 1563–1566. <https://doi.org/10.1126/science.3201247>
2. Cao, L., Peng, B., Yao, L., Zhang, X., Sun, K., Yang, X., & Yu, L. (2010). The ancient function of RB-E2F Pathway: Insights from its evolutionary history. *Biology Direct*, *5*(1), 55. <https://doi.org/10.1186/1745-6150-5-55>
3. Du, W., Vidal, M., Xie, J. E., & Dyson, N. (1996). RBF, a novel RB-related gene that regulates E2F activity and interacts with cyclin E in *Drosophila*. *Genes & Development*, *10*(10), 1206–1218. <https://doi.org/10.1101/gad.10.10.1206>

4. Korenjak, M., Taylor-Harding, B., Binné, U. K., Satterlee, J. S., Stevaux, O., Aasland, R., White-Cooper, H., Dyson, N., & Brehm, A. (2004). Native E2F/RBF Complexes Contain Myb-Interacting Proteins and Repress Transcription of Developmentally Controlled E2F Target Genes. *Cell*, *119*(2), 181–193. <https://doi.org/10.1016/j.cell.2004.09.034>
5. Lewis, P. W. (2004). Identification of a *Drosophila* Myb-E2F2/RBF transcriptional repressor complex. *Genes & Development*, *18*(23), 2929–2940. <https://doi.org/10.1101/gad.1255204>
6. Beall, E. L., Manak, J. R., Zhou, S., Bell, M., Lipsick, J. S., & Botchan, M. R. (2002). Role for a *Drosophila* Myb-containing protein complex in site-specific DNA replication. *Nature*, *420*(6917), Article 6917. <https://doi.org/10.1038/nature01228>
7. Georgette, D., Ahn, S., MacAlpine, D. M., Cheung, E., Lewis, P. W., Beall, E. L., Bell, S. P., Speed, T., Manak, J. R., & Botchan, M. R. (2007). Genomic profiling and expression studies reveal both positive and negative activities for the *Drosophila* Myb MuvB/dREAM complex in proliferating cells. *Genes & Development*, *21*(22), 2880–2896. <https://doi.org/10.1101/gad.1600107>
8. Acharya, P., Negre, N., Johnston, J., Wei, Y., White, K. P., Henry, R. W., & Arnosti, D. N. (2012). Evidence for autoregulation and cell signaling pathway regulation from genome-wide binding of the *Drosophila* retinoblastoma protein. *G3 (Bethesda, Md.)*, *2*(11), 1459–1472. <https://doi.org/10.1534/g3.112.004424>
9. Harrison, M. M., Ceol, C. J., Lu, X., & Horvitz, H. R. (2006). Some *C. elegans* class B synthetic multivulva proteins encode a conserved LIN-35 Rb-containing complex distinct from a NuRD-like complex. *Proceedings of the National Academy of Sciences*, *103*(45), 16782–16787. <https://doi.org/10.1073/pnas.0608461103>
10. Litovchick, L., Sadasivam, S., Florens, L., Zhu, X., Swanson, S. K., Velmurugan, S., Chen, R., Washburn, M. P., Liu, X. S., & DeCaprio, J. A. (2007). Evolutionarily Conserved Multisubunit RBL2/p130 and E2F4 Protein Complex Represses Human Cell Cycle-Dependent Genes in Quiescence. *Molecular Cell*, *26*(4), 539–551. <https://doi.org/10.1016/j.molcel.2007.04.015>
11. Schmit, F., Korenjak, M., Mannefeld, M., Schmitt, K., Franke, C., von Eyss, B., Gagrlica, S., Hanel, F., Brehm, A., & Gaubatz, S. (2007). LINC, a Human Complex That is Related to pRB-Containing Complexes in Invertebrates Regulates the Expression of G<sub>2</sub>/M Genes. *Cell Cycle*, *6*(15), 1903–1913. <https://doi.org/10.4161/cc.6.15.4512>
12. Ferguson, E. L., & Horvitz, H. R. (1989). The Multivulva Phenotype of Certain *Caenorhabditis Elegans* Mutants Results from Defects in Two Functionally Redundant Pathways. *Genetics*, *123*(1), 109–121.
13. Lu, X., & Horvitz, H. R. (1998). Lin-35 and lin-53, two genes that antagonize a *C. elegans* Ras pathway, encode proteins similar to Rb and its binding protein RbAp48. *Cell*, *95*(7), 981–991. [https://doi.org/10.1016/s0092-8674\(00\)81722-5](https://doi.org/10.1016/s0092-8674(00)81722-5)
14. Ceol, C. J., & Horvitz, H. R. (2001). Dpl-1 DP and efl-1 E2F Act with lin-35 Rb to Antagonize Ras Signaling in *C. elegans* Vulval Development. *Molecular Cell*, *7*(3), 461–473. [https://doi.org/10.1016/S1097-2765\(01\)00194-0](https://doi.org/10.1016/S1097-2765(01)00194-0)
15. Vorster, P. J., Goetsch, P., Wijeratne, T. U., Guiley, K. Z., Andrejka, L., Tripathi, S., Larson, B. J., Rubin, S. M., Strome, S., & Lipsick, J. S. (2020). A long lost key opens an ancient lock: *Drosophila* Myb causes a synthetic multivulval phenotype in nematodes. *Biology Open*, *9*(5), bio051508. <https://doi.org/10.1242/bio.051508>
16. Tabuchi, T. M., Deplancke, B., Osato, N., Zhu, L. J., Barrasa, M. I., Harrison, M. M., Horvitz, H. R., Walhout, A. J. M., & Hagstrom, K. A. (2011). Chromosome-Biased Binding and Gene Regulation

- by the *Caenorhabditis elegans* DRM Complex. *PLoS Genetics*, 7(5), e1002074. <https://doi.org/10.1371/journal.pgen.1002074>
17. Goetsch, P. D., Garrigues, J. M., & Strome, S. (2017). Loss of the *Caenorhabditis elegans* pocket protein LIN-35 reveals MuvB's innate function as the repressor of DREAM target genes. *PLOS Genetics*, 13(11), e1007088. <https://doi.org/10.1371/journal.pgen.1007088>
  18. Sadasivam, S., & DeCaprio, J. A. (2013). The DREAM complex: Master coordinator of cell cycle-dependent gene expression. *Nature Reviews Cancer*, 13(8), Article 8. <https://doi.org/10.1038/nrc3556>
  19. Pilkinton, M., Sandoval, R., & Colamonici, O. R. (2007). Mammalian Mip/LIN-9 interacts with either the p107, p130/E2F4 repressor complex or B-Myb in a cell cycle-phase-dependent context distinct from the *Drosophila* dREAM complex. *Oncogene*, 26(54), 7535–7543. <https://doi.org/10.1038/sj.onc.1210562>
  20. Gargica, S., Hauser, S., Kolfshoten, I., Osterloh, L., Agami, R., & Gaubatz, S. (2004). Inhibition of oncogenic transformation by mammalian Lin-9, a pRB-associated protein. *The EMBO Journal*, 23(23), 4627–4638. <https://doi.org/10.1038/sj.emboj.7600470>
  21. Müller, G. A., Asthana, A., & Rubin, S. M. (2022). Structure and Function of MuvB Complexes. *Oncogene*, 41(21), 2909–2919. <https://doi.org/10.1038/s41388-022-02321-x>
  22. Pilkinton, M., Sandoval, R., Song, J., Ness, S. A., & Colamonici, O. R. (2007). Mip/LIN-9 regulates the expression of B-Myb and the induction of cyclin A, cyclin B, and CDK1. *The Journal of Biological Chemistry*, 282(1), 168–175. <https://doi.org/10.1074/jbc.M609924200>
  23. Desvoyes, B., & Gutierrez, C. (2020). Roles of plant retinoblastoma protein: Cell cycle and beyond. *The EMBO Journal*, 39(19). <https://doi.org/10.15252/embj.2020105802>
  24. Fischer, M., & DeCaprio, J. A. (2015). Does *Arabidopsis thaliana* DREAM of cell cycle control? *The EMBO Journal*, 34(15), 1987–1989. <https://doi.org/10.15252/embj.201592196>
  25. Kobayashi, K., Suzuki, T., Iwata, E., Nakamichi, N., Suzuki, T., Chen, P., Ohtani, M., Ishida, T., Hosoya, H., Müller, S., Leviczky, T., Pettkó-Szandtner, A., Darula, Z., Iwamoto, A., Nomoto, M., Tada, Y., Higashiyama, T., Demura, T., Doonan, J. H., ... Ito, M. (2015). Transcriptional repression by MYB 3R proteins regulates plant organ growth. *The EMBO Journal*, 34(15), 1992–2007. <https://doi.org/10.15252/embj.201490899>
  26. Magyar, Z., Bögre, L., & Ito, M. (2016). DREAMs make plant cells to cycle or to become quiescent. *Current Opinion in Plant Biology*, 34, 100–106. <https://doi.org/10.1016/j.pbi.2016.10.002>
  27. Ning, Y.-Q., Liu, N., Lan, K.-K., Su, Y.-N., Li, L., Chen, S., & He, X.-J. (2020). DREAM complex suppresses DNA methylation maintenance genes and precludes DNA hypermethylation. *Nature Plants*, 6(8), 942–956. <https://doi.org/10.1038/s41477-020-0710-7>
  28. Wang, Y., Fan, Y., Fan, D., Zhang, Y., Zhou, X., Zhang, R., Wang, Y., Sun, Y., Zhang, W., He, Y., Deng, X. W., & Zhu, D. (2022). The *Arabidopsis* DREAM complex antagonizes WDR5A to modulate histone H3K4me2/3 deposition for a subset of genome repression. *Proceedings of the National Academy of Sciences of the United States of America*, 119(27), e2206075119. <https://doi.org/10.1073/pnas.2206075119>
  29. Shepard, J. L., Amatruda, J. F., Stern, H. M., Subramanian, A., Finkelstein, D., Ziai, J., Finley, K. R., Pfaff, K. L., Hersey, C., Zhou, Y., Barut, B., Freedman, M., Lee, C., Spitsbergen, J., Neuberg, D., Weber, G., Golub, T. R., Glickman, J. N., Kutok, J. L., ... Zon, L. I. (2005). A zebrafish myb mutation causes genome instability and increased cancer susceptibility. *Proceedings of the*

- National Academy of Sciences of the United States of America*, 102(37), 13194–13199. <https://doi.org/10.1073/pnas.0506583102>
30. Yamauchi, T., Ishidao, T., Nomura, T., Shinagawa, T., Tanaka, Y., Yonemura, S., & Ishii, S. (2008). A B-Myb complex containing clathrin and filamin is required for mitotic spindle function. *The EMBO Journal*, 27(13), 1852–1862. <https://doi.org/10.1038/emboj.2008.118>
  31. Sasagawa, S., Nishimura, Y., Hayakawa, Y., Murakami, S., Ashikawa, Y., Yuge, M., Okabe, S., Kawaguchi, K., Kawase, R., & Tanaka, T. (2016). E2F4 Promotes Neuronal Regeneration and Functional Recovery after Spinal Cord Injury in Zebrafish. *Frontiers in Pharmacology*, 7, 119. <https://doi.org/10.3389/fphar.2016.00119>
  32. Wen, H., Andrejka, L., Ashton, J., Karess, R., & Lipsick, J. S. (2008). Epigenetic regulation of gene expression by Drosophila Myb and E2F2–RBF via the Myb–MuvB/dREAM complex. *Genes & Development*, 22(5), 601–614. <https://doi.org/10.1101/gad.1626308>
  33. Blanchard, D. P., Georgette, D., Antoszewski, L., & Botchan, M. R. (2014). Chromatin reader L(3)mbt requires the Myb-MuvB/DREAM transcriptional regulatory complex for chromosomal recruitment. *Proceedings of the National Academy of Sciences of the United States of America*, 111(40), E4234–4243. <https://doi.org/10.1073/pnas.1416321111>
  34. Fischer, M., & Müller, G. A. (2017). Cell cycle transcription control: DREAM/MuvB and RB-E2F complexes. *Critical Reviews in Biochemistry and Molecular Biology*, 52(6), 638–662. <https://doi.org/10.1080/10409238.2017.1360836>
  35. Sadasivam, S., Duan, S., & DeCaprio, J. A. (2012). The MuvB complex sequentially recruits B-Myb and FoxM1 to promote mitotic gene expression. *Genes & Development*, 26(5), 474–489. <https://doi.org/10.1101/gad.181933.111>
  36. Litovchick, L., Florens, L. A., Swanson, S. K., Washburn, M. P., & DeCaprio, J. A. (2011). DYRK1A protein kinase promotes quiescence and senescence through DREAM complex assembly. *Genes & Development*, 25(8), 801–813. <https://doi.org/10.1101/gad.2034211>
  37. Marceau, A. H., Felthousen, J. G., Goetsch, P. D., Iness, A. N., Lee, H.-W., Tripathi, S. M., Strome, S., Litovchick, L., & Rubin, S. M. (2016). Structural basis for LIN54 recognition of CHR elements in cell cycle-regulated promoters. *Nature Communications*, 7(1), 12301. <https://doi.org/10.1038/ncomms12301>
  38. Asthana, A., Ramanan, P., Hirschi, A., Guiley, K. Z., Wijeratne, T. U., Shelansky, R., Doody, M. J., Narasimhan, H., Boeger, H., Tripathi, S., Müller, G. A., & Rubin, S. M. (2022). The MuvB complex binds and stabilizes nucleosomes downstream of the transcription start site of cell-cycle dependent genes. *Nature Communications*, 13, 526. <https://doi.org/10.1038/s41467-022-28094-1>
  39. Korenjak, M., Kwon, E., Morris, R. T., Anderssen, E., Amzallag, A., Ramaswamy, S., & Dyson, N. J. (2014). DREAM co-operates with insulator-binding proteins and regulates expression at divergently paired genes. *Nucleic Acids Research*, 42(14), 8939–8953. <https://doi.org/10.1093/nar/gku609>
  40. Boccuni, P., MacGrogan, D., Scandura, J. M., & Nimer, S. D. (2003). The human L(3)MBT polycomb group protein is a transcriptional repressor and interacts physically and functionally with TEL (ETV6). *The Journal of Biological Chemistry*, 278(17), 15412–15420. <https://doi.org/10.1074/jbc.M300592200>
  41. Richter, C., Oktaba, K., Steinmann, J., Müller, J., & Knoblich, J. A. (2011). The tumour suppressor L(3)mbt inhibits neuroepithelial proliferation and acts on insulator elements. *Nature Cell Biology*, 13(9), 1029–1039. <https://doi.org/10.1038/ncb2306>

42. Sakaguchi, A., Joyce, E., Aoki, T., Schedl, P., & Steward, R. (2012). The histone H4 lysine 20 monomethyl mark, set by PR-Set7 and stabilized by L(3)mbt, is necessary for proper interphase chromatin organization. *PLoS One*, 7(9), e45321. <https://doi.org/10.1371/journal.pone.0045321>
43. Trojer, P., Li, G., Sims, R. J., Vaquero, A., Kalakonda, N., Boccuni, P., Lee, D., Erdjument-Bromage, H., Tempst, P., Nimer, S. D., Wang, Y.-H., & Reinberg, D. (2007). L3MBTL1, a Histone-Methylation-Dependent Chromatin Lock. *Cell*, 129(5), 915–928. <https://doi.org/10.1016/j.cell.2007.03.048>
44. Wreczycka, K., Franke, V., Uyar, B., Wurmus, R., Bulut, S., Tursun, B., & Akalin, A. (2019). HOT or not: Examining the basis of high-occupancy target regions. *Nucleic Acids Research*, 47(11), 5735–5745. <https://doi.org/10.1093/nar/gkz460>
45. Down, C. F., Millour, J., Lam, E. W.-F., & Watson, R. J. (2012). Binding of FoxM1 to G2/M gene promoters is dependent upon B-Myb. *Biochimica Et Biophysica Acta*, 1819(8), 855–862. <https://doi.org/10.1016/j.bbagr.2012.03.008>
46. Wiseman, E. F., Chen, X., Han, N., Webber, A., Ji, Z., Sharrocks, A. D., & Ang, Y. S. (2015). Dereglulation of the FOXM1 target gene network and its coregulatory partners in oesophageal adenocarcinoma. *Molecular Cancer*, 14, 69. <https://doi.org/10.1186/s12943-015-0339-8>
47. Hong, W., & Guan, K.-L. (2012). The YAP and TAZ transcription co-activators: Key downstream effectors of the mammalian Hippo pathway. *Seminars in Cell & Developmental Biology*, 23(7), 785–793. <https://doi.org/10.1016/j.semcdb.2012.05.004>
48. Zanconato, F., Forcato, M., Battilana, G., Azzolin, L., Quaranta, E., Bodega, B., Rosato, A., Bicciato, S., Cordenonsi, M., & Piccolo, S. (2015). Genome-wide association between YAP/TAZ/TEAD and AP-1 at enhancers drives oncogenic growth. *Nature Cell Biology*, 17(9), 1218–1227. <https://doi.org/10.1038/ncb3216>
49. Patschull, G., Walz, S., Gründl, M., Schwab, M., Rühl, E., Baluapuri, A., Cindric-Vranesic, A., Kneitz, S., Wolf, E., Ade, C. P., Rosenwald, A., Eyss, B. von, & Gaubatz, S. (2019). The Myb-MuvB Complex Is Required for YAP-Dependent Transcription of Mitotic Genes. *Cell Reports*, 27(12), 3533–3546.e7. <https://doi.org/10.1016/j.celrep.2019.05.071>
50. Boxem, M., & van den Heuvel, S. (2002). C. elegans class B synthetic multivulva genes act in G(1) regulation. *Current Biology: CB*, 12(11), 906–911. [https://doi.org/10.1016/s0960-9822\(02\)00844-8](https://doi.org/10.1016/s0960-9822(02)00844-8)
51. Gal, C., Carelli, F. N., Appert, A., Cerrato, C., Huang, N., Dong, Y., Murphy, J., Frapporti, A., & Ahringer, J. (2021). DREAM represses distinct targets by cooperating with different THAP domain proteins. *Cell Reports*, 37(3), 109835. <https://doi.org/10.1016/j.celrep.2021.109835>
52. DeBruhl, H., Wen, H., & Lipsick, J. S. (2013). The Complex Containing Drosophila Myb and RB/E2F2 Regulates Cytokinesis in a Histone H2Av-Dependent Manner. *Molecular and Cellular Biology*, 33(9), 1809–1818. <https://doi.org/10.1128/MCB.01401-12>
53. Costello, M. E., & Petrella, L. N. (2019). C. elegans synMuv B proteins regulate spatial and temporal chromatin compaction during development. *Development (Cambridge, England)*, 146(19), dev174383. <https://doi.org/10.1242/dev.174383>
54. Bhaskar, P. K., Surabhi, S., Tripathi, B. K., Mukherjee, A., & Mutsuddi, M. (2014). DLin52 is crucial for dE2F and dRBF mediated transcriptional regulation of pro-apoptotic gene hid. *Biochimica Et Biophysica Acta*, 1839(9), 800–812. <https://doi.org/10.1016/j.bbagr.2014.05.012>

55. Millard, C. J., Varma, N., Saleh, A., Morris, K., Watson, P. J., Bottrill, A. R., Fairall, L., Smith, C. J., & Schwabe, J. W. R. (2016). The structure of the core NuRD repression complex provides insights into its interaction with chromatin. *ELife*, *5*, e13941. <https://doi.org/10.7554/eLife.13941>
56. Bainor, A. J., Saini, S., Calderon, A., Casado-Polanco, R., Giner-Ramirez, B., Moncada, C., Cantor, D. J., Ernlund, A., Litovchick, L., & David, G. (2018). The HDAC-Associated Sin3B Protein Represses DREAM Complex Targets and Cooperates with APC/C to Promote Quiescence. *Cell Reports*, *25*(10), 2797-2807.e8. <https://doi.org/10.1016/j.celrep.2018.11.024>
57. Silverstein, R. A., & Ekwall, K. (2005). Sin3: A flexible regulator of global gene expression and genome stability. *Current Genetics*, *47*(1), 1–17. <https://doi.org/10.1007/s00294-004-0541-5>
58. Esterlechner, J., Reichert, N., Iltzsche, F., Krause, M., Finkernagel, F., & Gaubatz, S. (2013). LIN9, a subunit of the DREAM complex, regulates mitotic gene expression and proliferation of embryonic stem cells. *PloS One*, *8*(5), e62882. <https://doi.org/10.1371/journal.pone.0062882>
59. Osterloh. (n.d.). *The human synMuv-like protein LIN-9 is required for transcription of G2/M genes and for entry into mitosis*. <https://doi.org/10.1038/sj.emboj.7601478>
60. Guiley, K. Z., Liban, T. J., Felthousen, J. G., Ramanan, P., Litovchick, L., & Rubin, S. M. (2015). Structural mechanisms of DREAM complex assembly and regulation. *Genes & Development*, *29*(9), 961–974. <https://doi.org/10.1101/gad.257568.114>
61. Fischer, M., Grossmann, P., Padi, M., & DeCaprio, J. A. (2016). Integration of TP53, DREAM, MMB-FOXM1 and RB-E2F target gene analyses identifies cell cycle gene regulatory networks. *Nucleic Acids Research*, *44*(13), 6070–6086. <https://doi.org/10.1093/nar/gkw523>
62. Müller, G. A., Stangner, K., Schmitt, T., Wintsche, A., & Engeland, K. (2017). Timing of transcription during the cell cycle: Protein complexes binding to E2F, E2F/CLE, CDE/CHR, or CHR promoter elements define early and late cell cycle gene expression. *Oncotarget*, *8*(58), 97736–97748. <https://doi.org/10.18632/oncotarget.10888>
63. Mannefeld, M., Klassen, E., & Gaubatz, S. (2009). B-MYB is required for recovery from the DNA damage-induced G2 checkpoint in p53 mutant cells. *Cancer Research*, *69*(9), 4073–4080. <https://doi.org/10.1158/0008-5472.CAN-08-4156>
64. Quaas, M., Müller, G. A., & Engeland, K. (2012). P53 can repress transcription of cell cycle genes through a p21(WAF1/CIP1)-dependent switch from MMB to DREAM protein complex binding at CHR promoter elements. *Cell Cycle (Georgetown, Tex.)*, *11*(24), 4661–4672. <https://doi.org/10.4161/cc.22917>
65. Hayflick, L., & Moorhead, P. S. (1961). The serial cultivation of human diploid cell strains. *Experimental Cell Research*, *25*, 585–621. [https://doi.org/10.1016/0014-4827\(61\)90192-6](https://doi.org/10.1016/0014-4827(61)90192-6)
66. Mijit, M., Caracciolo, V., Melillo, A., Amicarelli, F., & Giordano, A. (2020). Role of p53 in the Regulation of Cellular Senescence. *Biomolecules*, *10*(3), 420. <https://doi.org/10.3390/biom10030420>
67. Kumari, R., Hummerich, H., Shen, X., Fischer, M., Litovchick, L., Mitnacht, S., DeCaprio, J. A., & Jat, P. S. (2021). Simultaneous expression of MMB-FOXM1 complex components enables efficient bypass of senescence. *Scientific Reports*, *11*(1), 21506. <https://doi.org/10.1038/s41598-021-01012-z>
68. Rovani, M. K., Brachmann, C. B., Ramsay, G., & Katzen, A. L. (2012). The dREAM/Myb-MuvB complex and Grim are key regulators of the programmed death of neural precursor cells at the *Drosophila* posterior wing margin. *Developmental Biology*, *372*(1), 88–102. <https://doi.org/10.1016/j.ydbio.2012.08.022>

69. Clavier, A., Baillet, A., Rincheval-Arnold, A., Coléno-Costes, A., Lasbleiz, C., Mignotte, B., & Guéna, I. (2014). The pro-apoptotic activity of *Drosophila* Rbf1 involves dE2F2-dependent downregulation of *diap1* and *buffy* mRNA. *Cell Death & Disease*, *5*, e1405. <https://doi.org/10.1038/cddis.2014.372>
70. Reddien, P. W., Andersen, E. C., Huang, M. C., & Horvitz, H. R. (2007). DPL-1 DP, LIN-35 Rb and EFL-1 E2F Act With the MCD-1 Zinc-Finger Protein to Promote Programmed Cell Death in *Caenorhabditis elegans*. *Genetics*, *175*(4), 1719–1733. <https://doi.org/10.1534/genetics.106.068148>
71. Schertel, C., & Conradt, B. (2007). *C. elegans* orthologs of components of the RB tumor suppressor complex have distinct pro-apoptotic functions. *Development*, *134*(20), 3691–3701. <https://doi.org/10.1242/dev.004606>
72. Reichert, N., Wurster, S., Ulrich, T., Schmitt, K., Hauser, S., Probst, L., Götz, R., Ceteci, F., Moll, R., Rapp, U., & Gaubatz, S. (2010). Lin9, a subunit of the mammalian DREAM complex, is essential for embryonic development, for survival of adult mice, and for tumor suppression. *Molecular and Cellular Biology*, *30*(12), 2896–2908. <https://doi.org/10.1128/MCB.00028-10>
73. Fay, D. S., & Han, M. (2000). The synthetic multivulval genes of *C. elegans*: Functional redundancy, Ras-antagonism, and cell fate determination. *Genesis (New York, N.Y.: 2000)*, *26*(4), 279–284. [https://doi.org/10.1002/\(sici\)1526-968x\(200004\)26:4<279::aid-gene100>3.0.co;2-c](https://doi.org/10.1002/(sici)1526-968x(200004)26:4<279::aid-gene100>3.0.co;2-c)
74. Dimova, D. K., Stevaux, O., Frolov, M. V., & Dyson, N. J. (2003). Cell cycle-dependent and cell cycle-independent control of transcription by the *Drosophila* E2F/RB pathway. *Genes & Development*, *17*(18), 2308–2320. <https://doi.org/10.1101/gad.1116703>
75. Cayirlioglu, P., Bonnette, P. C., Dickson, M. R., & Duronio, R. J. (2001). *Drosophila* E2f2 promotes the conversion from genomic DNA replication to gene amplification in ovarian follicle cells. *Development (Cambridge, England)*, *128*(24), 5085–5098. <https://doi.org/10.1242/dev.128.24.5085>
76. Frolov, M. V., Huen, D. S., Stevaux, O., Dimova, D., Balczarek-Strang, K., Elsdon, M., & Dyson, N. J. (2001). Functional antagonism between E2F family members. *Genes & Development*, *15*(16), 2146–2160. <https://doi.org/10.1101/gad.903901>
77. Doggett, K., Jiang, J., Aleti, G., & White-Cooper, H. (2011). Wake-up-call, a *lin-52* paralogue, and Always early, a *lin-9* homologue physically interact, but have opposing functions in regulating testis-specific gene expression. *Developmental Biology*, *355*(2), 381–393. <https://doi.org/10.1016/j.ydbio.2011.04.030>
78. Laktionov, P. P., Maksimov, D. A., Romanov, S. E., Antoshina, P. A., Posukh, O. V., White-Cooper, H., Koryakov, D. E., & Belyakin, S. N. (2018). Genome-wide analysis of gene regulation mechanisms during *Drosophila* spermatogenesis. *Epigenetics & Chromatin*, *11*(1), 14. <https://doi.org/10.1186/s13072-018-0183-3>
79. Cheng, M.-H., Andrejka, L., Vorster, P. J., Hinman, A., & Lipsick, J. S. (2017). The *Drosophila* LIN54 homolog Mip120 controls two aspects of oogenesis. *Biology Open*, *6*(7), 967–978. <https://doi.org/10.1242/bio.025825>
80. Lee, H., Ragusano, L., Martinez, A., Gill, J., & Dimova, D. K. (2012). A dual role for the dREAM/MMB complex in the regulation of differentiation-specific E2F/RB target genes. *Molecular and Cellular Biology*, *32*(11), 2110–2120. <https://doi.org/10.1128/MCB.06314-11>
81. Calvisi, D. F., Simile, M. M., Ladu, S., Frau, M., Evert, M., Tomasi, M. L., Demartis, M. I., Daino, L., Seddaiu, M. A., Brozzetti, S., Feo, F., & Pascale, R. M. (2011). Activation of v-Myb avian

- myeloblastosis viral oncogene homolog-like2 (MYBL2)-LIN9 complex contributes to human hepatocarcinogenesis and identifies a subset of hepatocellular carcinoma with mutant p53. *Hepatology (Baltimore, Md.)*, 53(4), 1226–1236. <https://doi.org/10.1002/hep.24174>
82. Thorner, A. R., Hoadley, K. A., Parker, J. S., Winkel, S., Millikan, R. C., & Perou, C. M. (2009). In vitro and in vivo analysis of B-Myb in basal-like breast cancer. *Oncogene*, 28(5), 742–751. <https://doi.org/10.1038/onc.2008.430>



## D- PROBLEMATIQUE DE LA THESE

Rbfl, l'homologue du suppresseur de tumeur pRb, est à la fois capable d'induire l'apoptose dans des tissus en développement (disque imaginal d'aile) et de protéger de l'apoptose dans des tissus non prolifératifs (neurones) [158].

L'action de Rbfl passe par la régulation transcriptionnelle de gènes cibles, qui peuvent être activés ou réprimés. Rbfl ne possédant pas de domaine de liaison à l'ADN, cette modulation se fait par interaction avec de nombreuses protéines régulatrices de la transcription, comme des protéines modificatrices des histones, des facteurs de remodelage de la chromatine ou encore des co-facteurs de transcription.

Si les protéines de la famille RB ont été très étudiées pour leur rôle dans le cycle cellulaire, leur implication dans la régulation de l'apoptose est bien moins investiguée. Ce rôle, pourtant indispensable, est complexe dans la mesure où Rbfl peut être pro- ou anti-apoptotique, activateur ou répresseur.

Le but initial de ce projet était d'identifier les cibles de Rbfl dans le disque imaginal d'aile et dans le disque imaginal d'œil (ou les neurones) afin de déterminer des facteurs régulés différemment pouvant expliquer son rôle antagoniste vis-à-vis de l'apoptose. En effet, cet antagonisme peut s'expliquer par un ensemble différent de gènes régulés et/ou un ensemble de gènes communs régulés de façon opposée dans les deux tissus. L'utilisation d'une approche combinée ChIP-seq/RNA-seq devait nous permettre de caractériser les cibles directes de Rbfl : d'une part, l'immunoprécipitation de chromatine permet d'identifier les promoteurs sur lesquels sont fixés Rbfl ; d'autre part, le RNA-seq permet de déterminer quels gènes sont différentiellement transcrits. Le croisement des données de ces deux approches peut révéler les gènes dont les promoteurs liés par Rbfl et la transcription est modulée. Nous souhaitons également, à plus long terme, une fois ces cibles identifiées et validées, déterminer quelles sous-unités du complexe dREAM sont également présentes à ces promoteurs. En effet, comme présenté dans la revue dont je suis premier auteur de la partie précédente, la composition du complexe dREAM régit son activité activatrice ou répressive, et ces résultats nous auraient permis de comprendre le rôle de dREAM et de ses sous-unités dans le contrôle de l'apoptose.

Les approches de biologie moléculaires telles que celles que nous avons mises en place nécessitent une grande quantité de matériel biologique. Par exemple, on considère qu'une expérience de ChIP nécessite 400 disques imaginaires d'aile par échantillon [187], nombre qui peut augmenter si nous nous intéressons à une sous-population du disque d'aile. Cela représentait plusieurs milliers de disques pour notre approche. La méthode classique de récupération des disques imaginaires chez la larve de drosophile est la dissection. La dissection de plusieurs milliers de disques est en général réalisée par plusieurs chercheurs sur un temps long, ce qui génère de la variabilité dans le traitement des disques.

Pour nous affranchir de cette tâche fastidieuse et de la variabilité induite, j'ai mis au point un protocole de récupération en masse de disques imaginaires. Celui-ci est basé sur la récupération des organes par broyage des larves et différentes étapes de sédimentation pour séparer les disques imaginaires d'aile du reste des organes. Cette mise au point a fait l'objet d'un article méthodologie contenant les détails de ce protocole ainsi que les contrôles qualité effectués sur les disques obtenus, permettant d'affirmer que ceux-ci ont une qualité comparable aux disques imaginaires disséqués. Parmi ces contrôles, une quantification de l'apoptose a été réalisée, suivant un protocole publié que j'ai contribué à élaborer et qui est décrit dans une publication dont je suis seconde autrice, jointe en annexe I de la thèse. Ce protocole, qui permet de quantifier des marquages apoptotiques sur tissus, inclut une macro permettant une analyse des résultats en aveugle. Dans le cadre de cette publication, j'ai participé à l'acquisition des images utilisées pour les mises au point ainsi qu'aux tests des différentes versions de la macro décrite dans cet article. La comparaison de mes résultats avec ceux de deux autres expérimentateurs a notamment permis de montrer que dans ce protocole d'analyse d'images, le biais expérimentateur était négligeable [188]. Le protocole d'isolement des disques imaginaires est présenté dans un article publié qui fait l'objet de la première partie du chapitre Résultats. L'optimisation technique de ce protocole ayant été plus longue qu'anticipé, nous avons choisi de recentrer le projet sur le disque imaginaire d'aile.

Nous avons également mis au point et réalisé une expérience de ChIP-seq sur disques imaginaires d'aile. Malheureusement, nous nous sommes heurtées à de nombreuses difficultés lors de ces mises au point qui nous ont forcées, avec la crise de la COVID-19, à limiter le nombre d'échantillons à séquencer. L'analyse de ces données de séquençage a été réalisée sur la plateforme en ligne Galaxy, mais les résultats de cette analyse ne sont pas exploitables.

Pour le RNA-seq, nous avons mis au point un tri cellulaire en cytométrie de flux avec la collaboration de la plateforme d'imagerie et de cytométrie de l'Université Versailles -Saint-Quentin-en-Yvelines. Ce tri nous a permis de réaliser les extractions d'ARN sur la sous-population du disque imaginaire d'aile dans laquelle s'exprime le pilote d'intérêt. Lors de la création des banques, la plateforme de séquençage a fait face à un incident technique qui a altéré la qualité de nos données en augmentant considérablement le taux de duplicats de PCR. Nous avons choisi d'analyser tout de même ces données, cette fois-ci en réalisant nous-mêmes les traitements bio-informatiques, en retirant ou non les duplicats de PCR. Si malheureusement, ces données n'ont pas non plus pu être exploitées, elles ont été l'occasion de me former aux analyses de transcriptomique.

En parallèle de ces expériences, nous nous sommes également intéressées aux moyens de valider les cibles potentiellement identifiées. Si un test de l'activité pro-apoptotique de Rbf1 est utilisé en routine au laboratoire, nous n'avons pas de moyens déjà mis au point pour tester son activité anti-apoptotique. Nous avons donc mis en place un test basé sur le sauvetage par Rbf1 de la létalité pupale induite par la surexpression d'un pro-apoptotique dans les neurones.

# **MATERIELS ET METHODES**

Ce chapitre Matériels et Méthodes ne concerne que les sections Résultats II-, III- et IV-. La section Résultats I- étant un article publié, la partie Matériels et Méthodes correspondante est incluse dans la publication.

### Souches de drosophile

Les drosophiles ont été élevées sur milieu standard. Les souches utilisées sont présentées ci-après par ordre alphabétique.

Désignation	Génotype	Provenance
Balanceur lié	$\frac{Sp; MKRS}{SM5; TM6B}$	LGBC
Debcl (D3)	$\frac{Sp}{CyO - GFP}; \frac{UAS - Debcl(D3)}{TM6, Tb^1, Sb^1}$	LGBC
Dmp53	$\frac{Sp}{CyO - GFP}; \frac{UAS - dp53}{UAS - dp53}$	Bertrand Mollereau, LBMC
eGFP	$\frac{+ eGFP}{+ eGFP}$	LGBC
elav	$\frac{elav C155 - Gal4}{X ou Y}; \frac{+ +}{+ +}$	Bloomington Stock Center BL-458
GMR	$\frac{GMR MM - Gal4}{CyO - GFP}; \frac{+}{+}$	Marianne Malartre, I2BC
GUS-p53	$\frac{w^{1118}}{w^{1118}}; \frac{P[w^+, GUS - p53]}{P[w^+, GUS - p53]}; \frac{+}{+}$	Bloomington Stock Center BL-6584
HA-Rbf1 2Mα	$\frac{w^*}{w^*}; \frac{P\{w^+, UAS - HA - RBF1\}_{2MA\alpha}}{CyO, act5C - GFP}; \frac{+}{+}$	LGBC
LacZ	$\frac{w^{1118}}{w^{1118}}; \frac{+ P\{w^+, UAS - LacZ.NZ\}J312}{+}; \frac{P\{w^+, UAS - LacZ.NZ\}J312}{+}$	Bloomington Stock Center BL-3956
mCD8-GFP II	$\frac{w^*}{w^*}; \frac{P\{UAS - mCD8 :: GFP\} attP40}{P\{UAS - mCD8 :: GFP\} attP40}; \frac{+}{+}$	Bloomington Stock Center BL-32186
mCD8-GFP III	$\frac{w^*}{w^*}; \frac{+ P\{yw, UAS - CD8 :: GFP\} attP2}{+}; \frac{P\{yw, UAS - CD8 :: GFP\} attP2}{+}$	Bloomington Stock Center BL-32185
Rbf1	$\frac{w^*}{w^*}; \frac{+ P\{w^+, UAS - RBF1\}}{+}; \frac{TM6, Tb^1, Sb^1}{+}$	LGBC
Rbf1-HA 1M	$\frac{w^*}{w^*}; \frac{+ UAS - RBF1 - HA_{1M}}{+}; \frac{TM3, Sb^1}{+}$	LGBC
vg	$\frac{vg - Gal4}{vg - Gal4}; \frac{+}{+}$	Laboratoire Joel Silber, Institut Jacques Monod
vg>eGFP	$\frac{w^*}{w^*}; \frac{P[w^+, vg - Gal4]}{P[w^+, vg - Gal4]}; \frac{P[w^+, UAS - eGFP]}{P[w^+, UAS - eGFP]}$	LGBC

wi	$\frac{w^{1118}}{w^{1118}}; \frac{+}{+}; \frac{+}{+}$	Laboratoire François Rouyer, NeuroPSI
----	---	--

### Croisements

Sauf indication contraires, les croisements ont été réalisés de la manière suivante. Des femelles vierges apportant les pilotes ont été croisées sur milieu standard avec des mâles apportant le gène d'intérêt sous contrôle de séquences UAS. Les croisements ont été placés à 25°C et les individus ont été changés de tubes tous les deux jours.

### Comptage des encoches à la marge de l'aile

Les ailes d'individus surexprimant *LacZ*, *rbf1*, *HA-rbf1* ou *rbf1-HA* sous contrôle du pilote *vestigial*, obtenus par croisement, ont été classées et comptées selon le nombre d'encoches présentes à la marge de l'aile selon l'échelle suivante : 0 = « sauvage », 1 = « faible », 2 à 3 = « intermédiaire », 4 ou plus = « fort ». Pour chaque génotype, plus de cent ailes ont été comptées.

### Observation des phénotypes d'œil

Des individus *GMR>Debc1* (D3) ; *GMR>Debc1* (D3), *Rbf1* ; *GMR>p53* ; *GMR>p53*, *Rbf1* ou *elav>GUS-p53* obtenus par croisements, ainsi que des individus *vg-Gal4* ont été séparés selon le sexe et le génotype. Les individus d'intérêt ont été sectionnés à la jonction tête/thorax ou au niveau du thorax et les têtes ont été déposées, de face ou de profil, sur du papier adhésif. Les photographies ont été prises grâce à la caméra de la loupe binoculaire MZFLIII (Leica) au grossissement x16.

### Test de suppression de la létalité pupale

Des femelles vierges « Balanceur lié » ont été croisées avec des mâles *eGFP*, *Rbf1*, *Rbf1-HA 1M*, *HA-Rbf1 2MA $\alpha$*  et *GUS-p53*, ainsi que des mâles *GUS-p53 + Rbf1* étiqueté ou non issus de croisements entre la souche *GUS-p53* et les différentes souches *Rbf1*. Sept femelles vierges *elav* ont été croisées avec trois mâles des descendance des huit croisements précédents. Des individus *elav>eGFP* ; *elav>Rbf1* ; *elav>HA-Rbf1* ; *elav >GUS-p53* ; *elav>GUS-p53*, *Rbf1* ; *elav>GUS-p53*, *HA-Rbf1* et *elav>GUS-p53* ont été obtenus par croisements. Ils sont tous été balancés sur *SM5:TM6B*. Ils ont été séparés et comptés selon le sexe et le génotype d'intérêt jusqu'à épuisement des naissances. Les individus d'intérêt ont été sectionnés au niveau du thorax et la partie antérieure a été déposée de profil sur du papier adhésif. Les photographies ont été prises grâce au logiciel Infinity Capture associé à la caméra Infinity 3 (Teledyne Luminera) montée sur la loupe binoculaire SZX16 (Olympus) au grossissement x16. La différence statistique est estimée par des tests de Chi-2 d'homogénéité, le seuil de significativité est placé à  $p = 0,05$ .

### Immunoprécipitation de chromatine (ChIP)

Les disques imaginaires d'aile de larves L3 ont été récupérés par le protocole décrit dans le chapitre Résultats, section I sans l'étape de dissociation. Ils ont ensuite été fixés dans 525  $\mu$ L de formaldéhyde (1,8 % en Ringer) pendant 10 minutes sur roue à température ambiante. Le pontage protéines/ADN a été stoppé par l'ajout de 50  $\mu$ L de glycine 1,25 M. Les disques ont été lavés au Ringer 1X trois fois puis conservés à  $-80^{\circ}\text{C}$  au moins 20 minutes. La lyse cellulaire a été effectuée en tampon de lyse (140 mM NaCl, 10 mM Tris-HCl pH8, 1 mM EDTA, 1 % Triton X-100, 0,1 % sodium deoxycholate) supplémenté à 0,3 % de SDS. La sonication a été réalisée par 16 cycles 30s ON/30s OFF dans un sonicateur Bioruptor (Diagenode) permettant l'obtention de fragments de 200 à 300 paires de bases. La chromatine a été centrifugée à 10 000  $\times g$  pendant 15 minutes à  $4^{\circ}\text{C}$ . Le surnageant a été dilué au tiers pour obtenir une concentration de SDS de 0,1 %. Au moins 20  $\mu$ L ont été prélevés et constituent la fraction « input ». Pour chaque IP, 20  $\mu$ L de billes ont été lavées dans du tampon de lyse avec anti-protéases (Roche complete EDTA-free protease inhibitor cocktail). 2  $\mu$ g d'anticorps anti-HA ChIP grade (ab9110, Abcam) ou 5  $\mu$ g d'IgG dans 300  $\mu$ L de tampon de lyse + anti-protéases ont été ajoutés sur les billes Dynabeads (Dynabeads Protein G, 10003D, Invitrogen), Adembeads (Bio-Adembeads Protein AG, 04630, Ademtech) ou MagnaChIP (MagnaChIP Protein G, 16-662, Sigma-Aldrich). Seul du tampon de lyse + anti-protéases a été ajouté sur la résine pré-couplée HA (Monoclonal anti-HA agarose mouse, HA-7 clone (Sigma)). Ces suspensions ont été incubées 2h sur roue à  $4^{\circ}\text{C}$ . Après prélèvement de 10 % du volume constituant la fraction « input », la chromatine restante est divisée de façon équitable entre chaque condition puis déposée sur les billes dont le surnageant a été préalablement retiré. Le volume est complété à 300  $\mu$ L avec du tampon de lyse + anti-protéases. Après incubation sur roue à  $4^{\circ}\text{C}$  sur la nuit, une fraction du surnageant des billes est récupérée et constitue la fraction non liée (FNL). Les billes sont lavées 5 minutes sur roue à  $4^{\circ}\text{C}$  deux fois dans du tampon de lyse, deux fois dans du tampon de lyse avec 500 mM de NaCl, une fois dans du tampon LiCl (10 mM Tris-HCl pH8, 0,25 M LiCl, 0,5 % NP40, 0,5 % sodium deoxycholate, 1 mM EDTA) et une fois dans du TE (10 mM Tris HCl pH8, 1 mM EDTA), elles constituent la fraction immunoprécipitée (IP). Les billes sont resuspendues dans 100  $\mu$ L de solution TE additionnée de 10  $\mu$ L de SDS 10% pour une concentration finale d'environ 1 % et de 6  $\mu$ L de solution de protéinase K à 10 mg/mL, soit 60  $\mu$ g (Proteinase K from *Tritirachium album*, P2308, Sigma-Aldrich). Les fractions FNL et input sont également supplémentées de 1 % SDS et 60  $\mu$ g de protéinase K. Chaque fraction est incubée au ThermoMixer® (Eppendorf) à  $65^{\circ}\text{C}$  1300 rpm sur la nuit. L'ADN est purifié le lendemain avec le kit « Macherey-Nagel PCR Clean-up », selon le protocole « high-SDS samples » du fabricant. Alternativement, les fractions « input », « FNL » et « IP » ont été purifiées grâce au kit I-Pure v2 (Diagenode) selon le protocole du fabricant. Chaque élution est faite en deux fois avec 25  $\mu$ L d'eau stérile RNase/DNase free (UltraPure™ Distilled Water DNase/RNase free, Invitrogen), pour un

## MATERIELS ET METHODES

volume final de 50  $\mu$ L. L'input est dosé grâce au Quantus™ Fluorometer (Promega) et au kit Quantifluor One dsDNA (E4871, Promega).

### qPCR

Les expériences de PCR quantitative ont été réalisées sur des appareils CFX 96 ou 384 (Bio-Rad) en utilisant du SYBR GREEN (ITaq Universal SYBR Green Supermix, 1725124, Bio-Rad) et 2,5  $\mu$ L d'ADN. Les amorces utilisées sont présentées dans le tableau ci-dessous, par ordre alphabétique.

Gène	Amorce sens	Amorce anti-sens	Localisation
<i>Buffy</i>	CAAACATTGGCTCGTTGTTG	CCCTGCTTACCCATCCATTA	Promoteur
<i>how</i>	ATCGGTAAAGTGGCCTTGTG	TCTCTCACCTCGCTCATTT	Promoteur
<i>PCNA</i>	ACATCCCTATCCCGCTCATT	CGTGCCTCGAACATATTGAA	Promoteur /début CDS
<i>RpL32</i> <i>/ rp49</i>	TACAGGCCCAAGATCGTGAAG	GACGCACTCTGTTGTCGATAC	Corps du gène
<i>slp-1</i>	CGCATTAGCTACCCGTCTCT	TGGTTTCTGATTCTCACACACGAC	Promoteur

### Immunofluorescence

La récupération des disques imaginaux d'ailes s'est faite par dissection de larves de stade 3 dans du PBS 1X. Les disques ont été fixés en formaldéhyde 3,7 % pendant 20 minutes à température ambiante et lavés trois fois en PBS 1X, 0,3 % Tween-20 sur roue à température ambiante. Ils ont ensuite été incubés sur la nuit avec une dilution 1/800 d'anticorps anti-HA (lapin, 3724S, Cell Signaling Technology) et 1/500 d'anticorps anti-GFP (souris, 11814460001, Roche) dans une solution 1X PBS, 0,3 % Tween-20, 2 % BSA. Le lendemain, après trois rinçages en 1X PBS 0,3 % Tween-20, les disques sont incubés 2 heures dans une dilution 1/400 d'anticorps secondaires anti-lapin couplés Alexa 568 (Alexa-fluor-568-conjugated goat anti-rabbit-IgG (H+L) antibody, 11011, Molecular Probes, Thermofisher) et anti-souris couplés Alexa 488 (Alexa-fluor-488-conjugated goat anti-rabbit-IgG (H+L) antibody, 11001, Molecular Probes, Thermofisher). Enfin, les disques sont montés entre lame et lamelle dans du milieu de montage ProLong Diamond (A-11011, Invitrogen) et les images ont été acquises grâce au microscope confocal Leica SP8 (Leica Camera) avec excitation à 488 et 568 nm.

### ChIP-seq

La ChIP a été réalisée de la façon décrite dans le paragraphe dédié. La banque, le séquençage ainsi que le contrôle qualité ont été réalisés par la plateforme de séquençage de l'Institut de Biologie Intégrative de la Cellule (I2BC, Gif-sur-Yvette). Le reste de l'analyse a été réalisé sur la plateforme en ligne Galaxy, les programmes suivants ont été utilisés.

<b>Etape</b>	<b>Intitulé du programme</b>	<b>Paramètres</b>
<b>Alignement</b>	Bowtie2 – map reads against reference genome	Reference genome > built-in genome > D. melanogaster Aug. 2014 (dm6)
<b>Filtre des données</b>	Filter SAM or BAM, output SAM or BAM – filter on FLAG MAPQ RG LN or by region	Header > Include header Minimum MAPQ quality score > 20
<b>Peak calling</b>	MAC2 callpeak – Call peaks from alignment results	Control File > No Format of Input Files > Paired-End BAM Effective genome size > D. melanogaster (1.2e8)
<b>Intersection des coordonnées des pics avec le génome</b>	Intersect the intervals of two datasets	Return > overlapping intervals For at least > 1 bp

Les pics générés sont visualisés grâce à IGV (version 2.11.2).

### RNA-seq

Les disques imaginaires d'aile de larves L3 ont été récupérés par le protocole décrit dans le chapitre Résultats, section I « Manuscrit 2 : Mass purification protocol for drosophila melanogaster wing imaginal discs: an alternative to dissection to obtain large numbers of disc cells » avec dissociation. Les cellules ont ensuite été analysées en utilisant le cytomètre en flux BD FACSAria™ III (BD Biosciences) équipé d'un laser 488 nm et le logiciel associé BD FACSDiva™ (version 9.0.1, BD Biosciences). Les débris et doublets de cellules ont été écartés sur des critères de taille (forward scatter, FSC). Elles ont ensuite été triées selon leur fluorescence GFP.

Les cellules triées ont été culotées par centrifugation 4000  $xg$  pendant 3 minutes à 4°C, puis resuspendues dans 100  $\mu L$  de TRIzol (15596018, Invitrogen) et congelées à -20°C. Une fois la quantité de cellules nécessaire accumulée, les échantillons sont décongelés sur glace et rassemblés en un seul tube par génotype (dans la limite de 1 mL), complété à 1 mL de TRIzol. Après agitation par vortex et incubation de 5 minutes à température ambiante, 200  $\mu L$  de chloroforme glacé sont ajoutés. L'ensemble est centrifugé à 12000  $xg$  pendant 15 min à 4°C et un volume V de la phase aqueuse est récupéré. Un volume V d'isopropanol glacé ainsi que du LiCl pour une concentration finale de 120 mM. La précipitation est faite à -20°C sur le week-end. Après précipitation, les échantillons sont centrifugés à 21 000  $xg$  pendant 1h30 à 2h à 4°C. Le surnageant est éliminé et le culot est rincé trois fois par ajout de 500  $\mu L$  d'éthanol 70 % glacé et centrifugation 7500  $xg$  5 min 4°C. Le surnageant est éliminé une dernière fois et le culot mis à sécher à température ambiante. Après séchage complet, le culot est resuspendu dans 25  $\mu L$  d'eau stérile RNase/DNase free (UltraPure™ Distilled Water



## MATERIELS ET METHODES

DNase/RNase free, Invitrogen)et dosé au Quantus™ Fluorometer (Promega) et au kit Quantifluor RNA System (E3310, Promega) .

Les banques et le séquençage ont été effectués par la société Fasteris (Plan-les-Ouates, Suisse). L'analyse a été réalisée sous Linux sur le serveur Fidji de la plateforme de génomique de l'UVSQ, puis sous Rstudio. Le détail des programmes utilisés est disponible dans le chapitre « Résultats ». Les pics générés sont visualisés grâce à IGV (version 2.11.2).

# RESULTATS

## I- Mise au point d'un protocole d'isolation en masse de disques imaginaires d'ailes

Afin d'identifier les cibles directes de Rbf1, ce projet se base sur une approche combinée de deux techniques de biologie moléculaire : l'immunoprécipitation de chromatine suivie de séquençage (ChIP-seq) et le séquençage des ARN (RNA-seq). Une des premières problématiques à résoudre a été celle de la récupération du matériel biologique nécessaire. En effet, ces approches nécessitent une quantité élevée d'acides nucléiques qui peut être difficile à rassembler, en particulier lorsqu'on s'intéresse à des sous-populations cellulaires. Dans cette thèse, nous souhaitons nous intéresser à la fois à deux populations : le disque imaginal d'œil (dans lequel Rbf1 est anti-apoptotique) et le domaine *vestigial* du disque imaginal d'aile (dans lequel Rbf1 est pro-apoptotique). Dans les deux cas, le nombre de cellules d'intérêt est limité et ces cellules ne sont pas directement accessibles, accumuler suffisamment de matériel pour des approches telles que le ChIP-seq et le RNA-seq peut donc se révéler fastidieux. On considère par exemple qu'une expérience de ChIP nécessite environ 400 disques imaginaires d'aile par échantillon, pour des conditions dans lesquelles le disque entier est utilisé, le facteur d'intérêt est présent en concentration suffisante, et l'anticorps utilisé pour l'immunoprécipitation présente une bonne efficacité [187]. Si dans notre cas la concentration de Rbf1 n'est pas un problème puisqu'il est surexprimé et étiqueté, le fait de ne s'intéresser qu'à une sous-population cellulaire augmente encore le nombre de disques nécessaires.

La méthode classique de récupération des disques imaginaires de la larve de drosophile est la dissection manuelle. En poursuivant l'exemple du disque imaginal d'aile, le domaine *vestigial*, qui est notre domaine d'intérêt, représente environ 40 % des cellules du disque dans les conditions contrôle. Lorsqu'on y surexprime Rbf1, pro-apoptotique dans ce tissu, le domaine *vestigial* ne représente plus que 10 à 15 % des cellules. Ainsi un échantillon témoin de ChIP nécessiterait un millier de disques imaginaires d'aile, et un échantillon en condition de surexpression de Rbf1 environ 4000. En prenant en compte les réplicats associés à chacune de ces conditions, cela représente plusieurs milliers de larves à disséquer, chacune contenant deux disques imaginaires d'aile. Dans notre laboratoire, une doctorante expérimentée peut disséquer environ 100 disques d'aile par heure, soit 50 larves. Une tâche d'une telle ampleur nécessiterait donc des centaines d'heures de dissection. Pour de grands nombres de disques, les dissections mobilisent en général plusieurs chercheurs. Cette méthode nécessitant minutie et concentration, le temps élevé et la participation de différents expérimentateurs introduisent de la variabilité dans le traitement des disques. Cette variabilité est à la fois due aux habitudes de dissection de chacun et à leurs productivité et précision, pouvant varier selon divers paramètres comme la fatigue ou l'expérience. Nous souhaitons donc développer un protocole permettant de récupérer de grandes quantités de disques imaginaires tout en s'affranchissant de la variabilité induite par l'expérimentateur.

Dans les années 1960-1970, deux protocoles proposent des méthodes d'isolement en masse de disques imaginaux de drosophile. Ceux-ci s'appuient sur des gradients de densité dans lesquels migrent les organes de larves broyées grossièrement. Ces protocoles permettent la récupération de fractions pures de disques imaginaux de morphologie normale [189,190], pouvant pour certains être transplantés et se développer normalement en structures adultes [190]. Cependant, ces protocoles ayant été publiés il y a plus de cinquante ans, ils étaient impossibles à reproduire. En effet, certaines méthodes ne sont pas assez détaillées ou certains outils ne sont plus disponibles. Plus récemment, en 2014, un autre protocole publié, inspiré de ceux de 1965 et 1971, permet également d'isoler les disques imaginaux d'aile. Cependant, une étape clé de ce protocole est le tri des organes par leur taille et le patron d'expression de la GFP par un Biosorter®, un matériel très coûteux non disponible dans la plupart des laboratoires [191]. A partir de ces différents protocoles, nous avons développé notre propre protocole, pouvant être utilisé dans n'importe quel laboratoire. Nous avons donc réalisé de nombreux tests pour mettre au point les grandes étapes, à savoir le broyage des larves et la séparation des disques imaginaux d'aile par sédimentation et gradient en Ficoll. Afin de s'assurer de la qualité relative de ces disques par rapport à ceux récupérés par dissection, nous avons également réalisé des contrôles qualité. Parmi eux, une quantification du niveau d'apoptose induit par la dissection et ce protocole. Cette quantification a été réalisée selon un protocole mis au point et publié par le laboratoire [188] et dont j'ai participé à l'élaboration. Cet article, dont je suis second auteur, est présenté en Annexe I.

Ce protocole permet de récupérer des fractions enrichies en disques imaginaux d'ailes à hauteur de 50 %. Il permet, selon la quantité de larves de stade L3 utilisée comme matériel de départ, la récupération de plusieurs centaines de disques imaginaux d'aile avec un rendement d'environ 11,5 %. Le contaminant principal de ces fractions étant les glandes salivaires, une étape de filtration a été ajoutée afin de recueillir uniquement les cellules de disques imaginaux, permettant d'obtenir à la fin des fractions cellulaires pures pour les expériences nécessitant une dissociation des disques imaginaux. Ce résultat est obtenu en deux à trois heures et est réalisable entièrement par un seul expérimentateur. Il permet ainsi de s'affranchir de la variabilité induite par la dissection, en particulier lorsque plusieurs chercheurs y participent. De plus, la durée de ce protocole ne dépend que peu du nombre de larves utilisé comme matériel de départ. Un nombre plus élevé de larves permettra donc la récupération d'un nombre plus élevé de disques imaginaux, dans un délai raisonnable par rapport au temps nécessaire à la dissection. Enfin, ce protocole permet également un traitement plus homogène des différents génotypes, puisque ceux-ci peuvent être traités en parallèle par le même expérimentateur plutôt que les uns à la suite des autres.

L'ensemble de ces mises au point et contrôles qualité, ainsi que le protocole final, sont compilés dans l'article suivant [192].

Protocol

# Mass Purification Protocol for *Drosophila melanogaster* Wing Imaginal Discs: An Alternative to Dissection to Obtain Large Numbers of Disc Cells

Marion Hoareau <sup>1,2</sup>, Juliette de Noiron <sup>1,2</sup>, Jessie Colin <sup>1,2,\*</sup>  and Isabelle Guénel <sup>1,\*</sup> <sup>1</sup> LGBC, UVSQ, Université Paris-Saclay, 78000 Versailles, France<sup>2</sup> Ecole Pratique des Hautes Etudes, PSL Research University, 75014 Paris, France

\* Correspondence: jessie.colin@ephe.psl.eu (J.C.); isabelle.guenal@uvsq.fr (I.G.)

† Current address: Unité Biologie des ARN des Pathogènes Fongiques, Institut Pasteur, 75015 Paris, France.

**Simple Summary:** *Drosophila melanogaster*, also known as the fruit fly, is a widely used organism model, especially for genetic studies or as a model for pathologies. The *Drosophila* genome is well known and conserved within humans, thus allowing biologists to obtain numerous mutants and transgenic flies from the fruit flies. Gene function studies at the cellular and molecular levels are often performed using extracts of larval tissues. Due to their small size, it is difficult to dissect substantial amounts of these tissues for performing genomic or proteomic experiments. This paper develops a simple method to purify larval tissues en masse. This protocol preserves tissue integrity in the same way as manual dissection; the protocol is achievable by individual researchers and allows the purification of different samples simultaneously.

**Abstract:** *Drosophila melanogaster* imaginal discs are larval internal structures that become the external organs of the adult. They have been used to study numerous developmental processes for more than fifty years. Dissecting these imaginal discs for collection is challenging, as the size of third-instar larvae organs is typically less than 1 mm. Certain experimental applications of the organs require many cells, which requires researchers to spend several hours dissecting them. This paper proposes an alternative to dissection in the form of a mass enrichment protocol. The protocol enables the recovery of many wing imaginal discs by grinding large quantities of third-instar larvae and separating the organs using filtration and a density gradient. The wing imaginal discs collected with this protocol in less than three hours are as well preserved as those collected by dissection. The dissociation and filtration of the extract allow the isolation of a large amount of wing imaginal disc cells.

**Keywords:** *Drosophila melanogaster*; larva; mass purification; isolation; wing imaginal discs; density gradient



**Citation:** Hoareau, M.; de Noiron, J.; Colin, J.; Guénel, I. Mass Purification Protocol for *Drosophila melanogaster* Wing Imaginal Discs: An Alternative to Dissection to Obtain Large Numbers of Disc Cells. *Biology* **2022**, *11*, 1384. <https://doi.org/10.3390/biology11101384>

Received: 1 September 2022

Accepted: 19 September 2022

Published: 22 September 2022

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

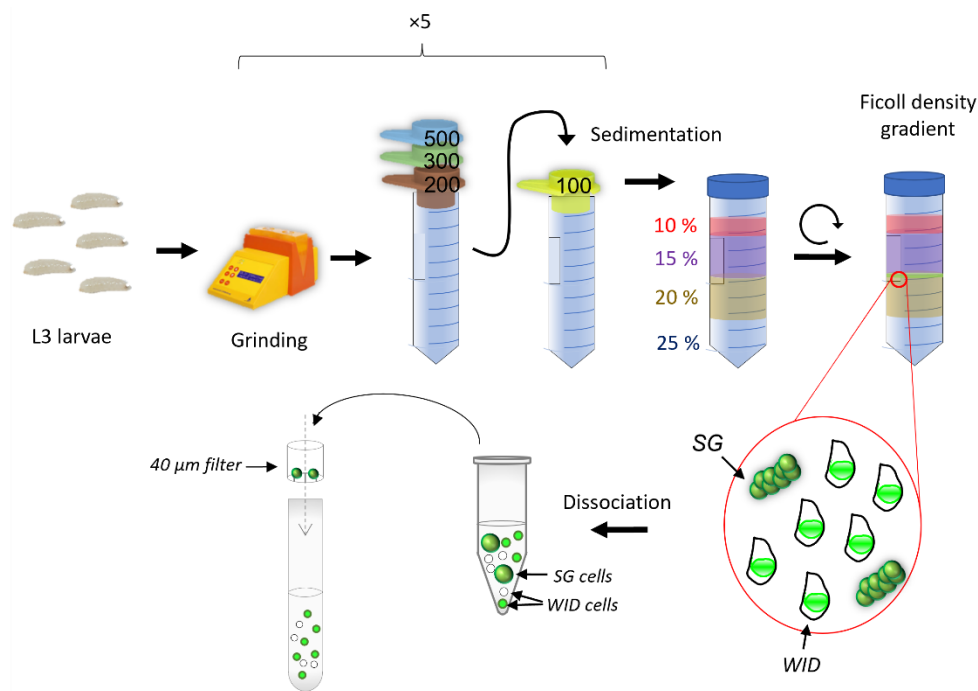
## 1. Introduction

In Holometabola, imaginal discs are larval internal epithelial structures that are precursors of external organs in the adult, such as wings or legs. These structures are widely used in the model organism *Drosophila melanogaster* for developmental studies. The structures are accessible, and most genetic tools engineered for this species are designed to be used in these organs, including drivers to use the UAS-GAL4 system [1] or mitotic clones [2] enabling constructs, among others. Many genes and signaling cascades involved in developmental control are conserved between mammals and *Drosophila melanogaster* [3], making imaginal discs an interesting model to study processes involved in both development [4,5] and human diseases (see for example [6,7]). Such studies led to significant discoveries in recent decades, including homeotic genes [8] and the Hippo pathway [9]. However, although *Drosophila* wing imaginal disc has proven to be a great model for genetic or cell biology

approaches, its use is more difficult in biochemistry or molecular biology experiments that need large amounts of starting material. For instance, a chromatin immunoprecipitation (ChIP) requires 400 wing imaginal discs per sample [10], and this number can be even higher if the factor of interest has a low concentration or is expressed only in a subpopulation of the disc cells. Manual dissection is the classical way to collect imaginal discs. Thus, doing experiments such as a ChIP with associated corresponding controls rapidly leads to the need for dissecting hundreds of larvae. Dissection is time-consuming and raises reproducibility issues, as several scientists may be involved. Subtle variations may occur throughout the dissection process, imaginal disc storage, or crosslinking step. Molecular biology experiments thus necessitate an alternative method to isolate imaginal discs en masse, in a reproducible manner, and by a single experimenter in a limited amount of time.

Since imaginal discs such as wing ones are precursors of non-essential organs, they are instrumental in studying cell death as it is possible to induce cell death during the development in these structures without affecting the survival of the specimen [11]. Our team is interested in apoptosis and tissue homeostasis. Our favorite model organ is the wing imaginal disc [12,13]. However, as our projects necessitated a large number of discs, we decided to invest in developing a mass wing imaginal discs enrichment protocol. This protocol would allow us to harvest a significantly greater sample size of wing imaginal discs than previously achieved for larger-scale experiments. A few articles describe such protocols. The first two were published in the 1960 and 1970s [14,15] and offer mass isolation methods that allow the recovery of hundreds of imaginal discs using density gradients. Methods related to the large-scale collection of fruit fly larval tissues described during the 1960–1970s [14,15] cannot be easily reproduced due to the lack of documentation of methods and the unavailability of some tools. More recently, Marty et al. [16] presented a new version of this protocol, but their goal was only to roughly separate organs and not to obtain pure imaginal disc fractions. In this study, the authors used a Biosorter<sup>®</sup> (Union Biometrica Inc., Holliston, MA, USA) to isolate wing imaginal discs according to their shape, size, or Green Fluorescent Protein (GFP) pattern. However, this device is expensive and not commonly available. Therefore, we wanted to develop a mass enrichment protocol routinely usable in any laboratory.

We propose here an imaginal disc mass enrichment protocol without manual dissection that allows the recovery of a large number of discs in an optimized time range, doable entirely by one experimenter, using density gradients and sedimentation; see the overall protocol in Figure 1. This method allows the experimenter to recover about 11.5% of the wing imaginal discs input within two to three hours in an enriched fraction (about 50%) that can be easily dissociated and filtrated to recover the wing imaginal disc cells. The protocol is optimized for wing imaginal discs, but its adaptation to other larvae organs such as eye-antenna imaginal discs or brains could be possible after different set-ups, especially the nature of the Ficoll gradient layers.



**Figure 1.** Overview of the mass purification protocol. L3 larvae obtained from synchronized egg-laying are collected en masse using a wash bottle. The larvae were initially ground using a GentleMACS™ (Miltenyi Biotec, Bergisch Gladbach, North Rhine-Westphalia, Germany). The material was then filtered through a series of strainers with decreasing mesh size, allowing the selection of elements between 100 μm and 200 μm. This material is resuspended in 10% (*w/v*) Ficoll and loaded on top of a 15:20:25% (*w/v*) Ficoll gradient. After centrifugation, wing imaginal discs (WID) are found at the 15:20% (*w/v*) interface along with some salivary gland (SG) pieces. After dissociation, filtering the cells through a 40 μm filter is enough to separate salivary gland cells from disc cells.

## 2. Materials and Methods

### 2.1. Fly Stocks

Flies were raised at 25 °C on a standard medium. The *vg-GAL4* strain is a generous gift from Joel Silber (Institut Jacques Monod, Université Paris Cité, France). The *UAS-mCD8-GFP* was obtained from the Bloomington Drosophila Stock Center (BL-32185).

### 2.2. Protocol for Mass Enrichment of Wing Imaginal Discs

Third-instar larvae were collected by flushing the side of their rearing tubes. The larvae were then ground using the GentleMACS™ device (Miltenyi Biotec, Bergisch Gladbach, Germany) and filtered through a series of sieves. The resulting material was loaded in 10% Ficoll solution (*w/v*) on top of a 15:20:25% gradient. After centrifugation, the 15:20% interface containing the enriched wing imaginal discs was collected and rehydrated in Ringer 1X (Supplementary File S1). After dissociation, filtration on 40 μm retained the cells of the salivary glands and allowed the recovery of only imaginal disc cells.

### 2.3. Immunostaining and Images Acquisition

For the “dissected” condition, wing imaginal discs were dissected from third-instar larvae in 1X Phosphate Buffered Saline (PBS), pH 7.6. Discs obtained by mass enrichment or dissection were fixed with 3.7% formaldehyde in 1X PBS for 20 min at room temperature and washed three times for 10 min in PBST (1X PBS, 0.3% Triton X-100).

Apoptosis was analyzed as previously described in de Noiron et al. [17]. In brief, the discs were blocked for 1 h in PBST-BSA (1X PBS, 0.3% Tween 20, 2% Bovin Serum Albumin) and incubated overnight with 1:100 dilution of anti-cleaved *Drosophila* Dcp-1

(Asp216, Cell Signaling Technology) at 4 °C. The following day, after 3 washes in PBST, wing discs were incubated for two hours with Alexa-568-coupled anti-rabbit secondary antibody (A-11011, Invitrogen) diluted to 1:400 in PBST. Finally, wing discs were mounted in ProLong Diamond (Invitrogen, Waltham, MA, USA), and images were acquired using a Leica SP8 confocal microscope (Leica Camera, Wetzlar, Germany) at 568 nm. At least 30 wing imaginal discs are analyzed for each condition. Image analysis was performed on Fiji with the macro described in de Noiron\_2021 [17], whose main steps are median filter application, stacking (Z-project), threshold determination, and signal quantification.

For tissue structure analysis, wing imaginal discs were incubated with phalloidin-ATTO 665 (Sigma-Aldrich, St. Louis, MO, USA) 1 µg/mL, and Hoechst 33,342 (Thermo Fisher Scientific, Waltham, MA, USA) 1 µg/mL, in PBST-BSA for 1 h at room temperature. Larvae were then washed twice for 5 min in PBST. Finally, wing discs were mounted in Citifluor™ (Biovalley, Basel, Switzerland) and observed with a Leica SP8 upright confocal microscope.

#### 2.4. Flow Cytometry Analysis

Third-instar larvae expressing *UAS-mCD8-GFP* under the control of the *vestigial-GAL4* driver were used in [18]. Wing imaginal discs recovered by either dissection or mass enrichment protocol were dissociated using a 1:50 dilution (final concentration 0.1% *w/v*) of 5% (*w/v*) protease (Sigma-Aldrich P8811) in 1X Ringer incubated for 20 min at 25 °C under gentle shaking (300 rpm). Dissociation was completed by gently passing the cells through a P1000 tip. After pelleting, cells were resuspended in 1X Ringer and filtered through a 40 µm sieve to retain salivary glands cells and eventual debris. Cells were then analyzed using the BD FACSAria™ III (BD Biosciences, Franklin Lakes, NJ, USA) equipped with a 488 nm laser line and the BD FACSDiva™ software (9.0.1, BD Biosciences, Franklin Lakes, NJ, USA). Cells were selected through forward scatter (FSC) to keep out debris and cell clusters. They were then sorted based on whether or not they expressed GFP.

For the cell death assessment by viability dye, 10 third-instar larvae expressing *UAS-mCD8-GFP* under the control of the *vestigial-GAL4* driver were used for each group. Wing imaginal discs recovered by either dissection in 1X PBS, dissection in 1X Ringer, or mass enrichment protocol in 1X Ringer were dissociated using a 0.1% (*w/v*) solution of protease (Sigma-Aldrich P8811) in a 1X PBS or 1X Ringer and incubated for 20 min at 25 °C under gentle shaking (300 rpm). Dissociation was completed by gently passing the cells through a P1000 tip. After pelleting, the cells were incubated for 3 min on ice in viability dye solution (Fixable Viability Dye eFluor™ 780, eBioscience™, Thermo Fisher Scientific, Waltham, MA, USA) diluted to 1:1000 in 1X PBS or 1X Ringer. After a wash in 1X PBS or 1X Ringer, cells were resuspended in 1X PBS or 1X Ringer and filtered through a 40 µm sieve. The fluorescence of the dye was then measured using the BD FACSAria™ III equipped with a 635 nm laser line and the BD FACSDiva™ software. Cells were selected through FSC, excluding debris and cell clusters. Data analysis was performed with FlowJo software (10.8.1 version, FlowJo, Ashland, OR, USA). Graphs and statistical analysis were performed using RStudio (ggplot2 and ggpubr packages, RStudio, Boston, MA, USA) with a significance threshold of 5%.

### 3. Results

#### 3.1. Grinding and Buffer Parameters to Release Internal Organs without Damaging Them

Few protocols of mass enrichment for imaginal discs have previously been described. They are all based on the protocol developed by Fristrom and Mitchell [14], but their use needs some modernizing. This protocol requires dismantling larvae to release internal organs later isolated by differential sedimentation. As we mainly work on wing imaginal discs, we focused on this organ and chose to monitor its enrichment process with the help of fluorescent wing discs. To this end, we used larvae expressing GFP under the control of the *vestigial* driver, a driver allowing expression in a wide band at the dorso-ventral frontier of wing and haltere wing discs. Like many other wing imaginal disc drivers [1,19],



it also displays leaky expression in the salivary glands, which is not an issue as salivary glands are easily distinguishable from wing imaginal discs.

### 3.1.1. Grinding Larvae: A Fine-Tuning between Breaking the Cuticle and Keeping the Organs Undamaged

The first step of this protocol consists of dismantling the larvae without damaging the imaginal discs. Protocols from Mitchell and Cohen laboratories used either a meat grinder that is in all likelihood not sold anymore [14] or a custom-made device specifically developed in this laboratory for this particular usage [15]. Our goal was to use a commercial tissue dissociator to ensure the reproducibility of the larvae grinding. Thus, we tried two devices. The first was a bead beater like the ones usually available in most laboratories, in our case, the Precellys<sup>®</sup> (Bertin Technologies, Montigny-le-Bretonneux, France) from Bertin. The second was the GentleMACS<sup>™</sup> from Miltenyi, which was successfully used for such applications by the Basler lab [16]. We expected it to perform a gentle grinding, as it was developed for tissue dissociation rather than whole lysis. As these devices are usually dedicated to mammalian tissue lysis or dissociation, their ability to grind *Drosophila* larvae is not documented. The grinding process was visually monitored by assessing the integrity and location (released from the larvae or not) of fluorescent discs at the end of the process.

The development tests were mainly carried out using 2.5 mL of larvae, representing around 600 individuals. This larvae amount is collected by flushing the side of six vials. Adult flies were transferred to a fresh medium every 24 h to ensure that all offspring were the same age and obtained in sufficient quantity. We used 12 vials containing 15 virgin females and 6 males to collect the same amount (2.5 mL) of larvae from a cross where the larvae of interest represented only 50% of the progeny. These values must be adapted according to the viability of the individuals of interest.

For the assays with the Precellys<sup>®</sup>, we used 2 mL tubes containing either 2.8 mm diameter beads or a mix of 1.4 and 2.8 mm beads with different grinding cycles. We found that 2.8 mm beads alone were more efficient in breaking cuticles and that the recovered imaginal discs were in better condition than using the bead mix. However, the yield was low since most discs remained attached to the larvae (data not shown). Furthermore, considering the quick mass enrichment protocol we wanted to set up, we did not further pursue the set-up with this device as it can only handle a limited number of larvae (100–120/tube) at once, and the recovery of the material after grinding was challenging to handle—these drawbacks are incompatible with mass enrichment.

The GentleMACS<sup>™</sup> uses tubes with spinning helix-like structures and comes with several pre-registered programs of different speeds and/or duration. Two types of tubes are available; the C tubes offer a gentler grinding and are dedicated to tissue dissociation into individual cells, whereas M tubes offer stronger grinding and are recommended for complete cell lysis as required for biomolecule extraction. We first used the “liver 1” programs (48 rotations per round, 15 s) that were previously used for such purpose [16] or “liver 2” programs (78 rpr, 24 s) with the two kinds of tubes (Supplementary Figure S1). In our hands, these programs were too gentle to detach the imaginal discs efficiently, and upon advice from the technical support from Miltenyi Biotec, we tested the “brain1” programs (116 rpr, 36 s) or “brain” 2 programs (100 rpr, 30 s), the second of which turned out to be much more appropriate. We observed that with the M tubes, larvae were ground either not enough or too much, resulting in damaged imaginal discs. By contrast, the C tubes allowed a dismantling of the larvae that released intact imaginal discs (Supplementary Figure S1). We used the GentleMACS<sup>™</sup> with the “brain 2” program and the C tubes based on these tests. This grinding step is repeated five times until most, if not all, internal organs are released from the larvae.

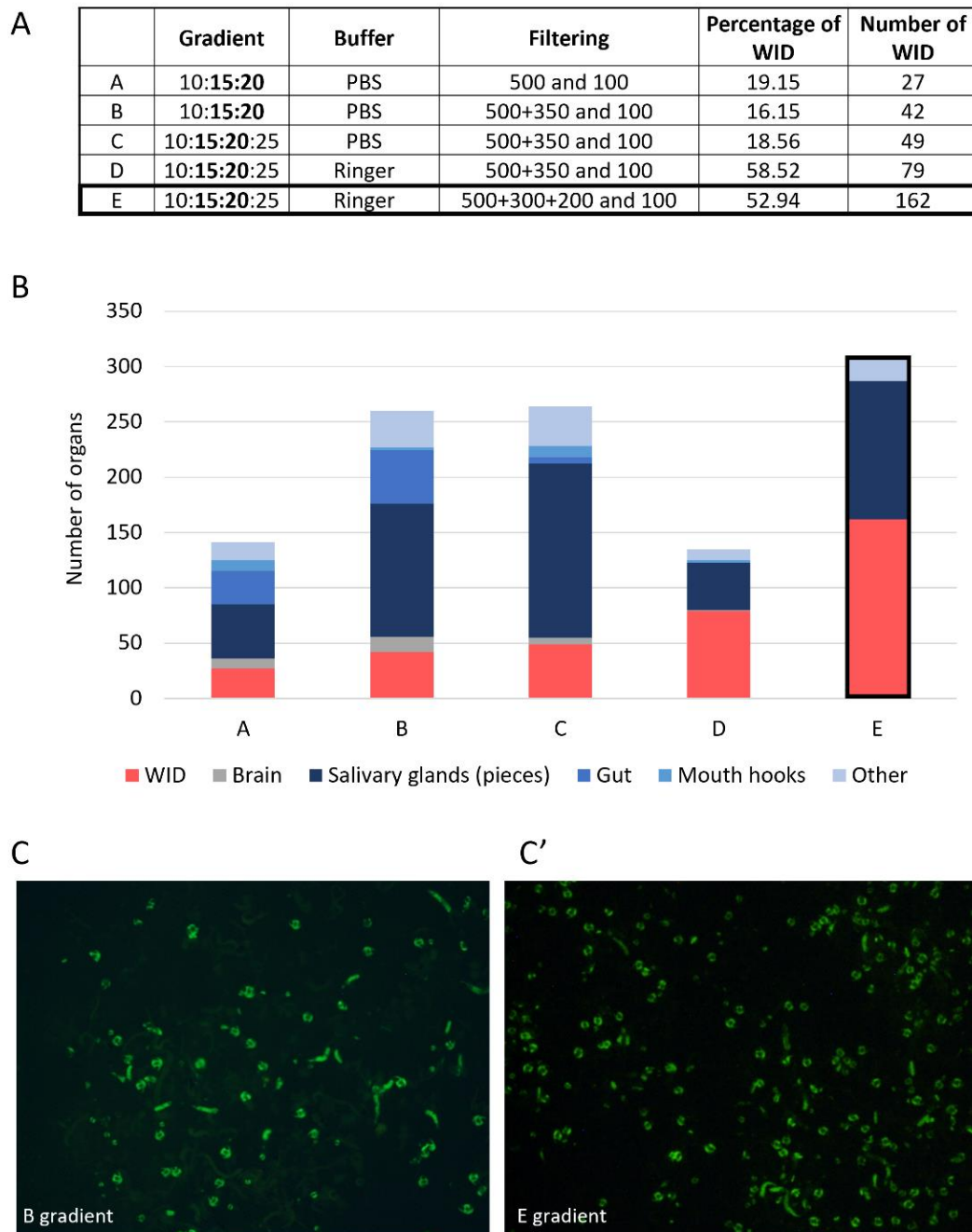
### 3.1.2. Filtration and Ficoll Gradient: Steps to Separate Organs of Interest from the Others

The second step of the protocol consists of isolating imaginal discs from the other organs. As per the protocol from Marty [16], we first loaded the ground material, without

further treatment, directly on a Ficoll gradient. However, this was not satisfying, since it led to a gradient saturation. We then added a step of rough separation of the organs based on their size. To this end, we used a series of filters with different meshing. We first tried a simple filtration step with only a 100  $\mu\text{m}$  strainer that kept wing imaginal discs and let us eliminate little pieces of organs (such as gut pieces). However, the ground fractions also contain many contaminants of larger size and/or density, such as pieces of cuticles or mouth hooks. Given the amount of material loaded on the filter, it was still rapidly clogged with these tissues. We thus decided to add a 500  $\mu\text{m}$  filter before the 100  $\mu\text{m}$  one to eliminate large structures, such as cuticles or tiny unground larvae, but the output of this 500  $\mu\text{m}$  was still not clean enough. After numerous tries with different strainer sizes, we combined of 500  $\mu\text{m}$ , 300  $\mu\text{m}$ , and 200  $\mu\text{m}$  in serial filtration (Figure 2). This filter combination retains all empty cuticles, larvae that are not fully grinded, and most mouth hooks while letting imaginal discs go through. This filtrate then undergoes 100  $\mu\text{m}$  filtration that retains imaginal discs but allows small-sized contaminants such as fat body and small gut pieces to be removed. The material recovered from the 100  $\mu\text{m}$  strainer contains imaginal discs and structures with the same size range, such as brains, proventriculus, “string-like” structures such as salivary glands and gut pieces, and some fat body pieces.

Still, to avoid gradient saturation, we added a sedimentation step. Indeed, this was a means to further improve the purity as gut and fat body pieces are much less dense than wing imaginal discs. To assess the efficiency of the sedimentation and purification steps, we needed a way to visually localize wing imaginal discs in the tubes. To do so, we took advantage of LacZ expression in those discs and added a fixation-less staining step of the beta-galactosidase [20] (Supplementary Figure S2). Finally, sedimentation with low-speed centrifugation allowed cleaning out this fraction.

Once the input for the Ficoll gradient separation was cleared from most unwanted larval elements, we worked on the enrichment step. We tried many gradient configurations (number, concentration, and volume of layers). We present here only a subset of them representing the milestones of the development of the method (Figure 2). The most recently published protocol uses a Ficoll gradient prepared with PBS, which allows the recovery of the imaginal discs at the 16:25% interface [16]. The oldest method comprises two gradients, the first placing imaginal discs at the interface with 14:19% of a Ringer-prepared Ficoll gradient and the second allowing further separation of this fraction using a continuous 14–24% gradient [14]. The latest should provide the purest imaginal discs fractions, whereas the most recent gives much less pure preparation, but this was not a problem for the authors since they further sorted the enriched discs using a BioSorter. We wanted our method to stand in between, being the easiest and shortest possible but still favoring purity over yield. We then decided to improve the quality of the output fraction using a single gradient. We started with PBS-prepared Ficoll gradients, as used in the single gradient protocol. As in previous experiments, discs were found in the interphase in the 14–25% range. We expected a disc enrichment in the 15:20% interface, with 15 and 20% layers becoming the basic set-up of our method. We soon noticed that adding a layer with more diluted Ficoll (i.e., 10%) improved the separation. Adding another layer of 25% Ficoll at the bottom of the gradient further improved the enrichment of imaginal discs. Our final protocol corresponds to the “E” gradient in Figure 2A. It comprises a Ficoll gradient with a first 10% layer in which the ground material is resuspended. This layer is then loaded on top of a 15:20:25% gradient. Using filter tips coated with fat larvae tissue prevents the loss of wing imaginal discs sticking to the tips. This improvement and precise pipetting at the interphase of the gradient significantly increased the protocol yield compared to the one in Figure 2A “E”. Indeed, in six enrichment experiments from 2.5 mL of third instar larvae (about 600 larvae), we obtained an average of 138 wing imaginal discs per preparation, corresponding to a yield of 11.5%.



**Figure 2.** Overview of the improvement in the yield and purity during Ficoll gradient set-up. The “E” gradient is used in this protocol final version. (A) Table showing 5 examples of gradient configuration (Ficoll gradient in %, filtering in  $\mu\text{m}$  (pore size)) and their outcome in proportion and number of wing imaginal discs (WID). Inputs were 5 mL of larvae. (B) Count of the different organs obtained in the output fractions. The “Other” group includes other imaginal discs (mainly legs and eye-antennae), proventriculus, and pieces of cuticles. (C) Examples of fractions obtained with the B (C) and E (C’) gradients in GFP-exciting light,  $\times 18$ . All pseudocircular GFP-marked objects are wing imaginal discs with *vg*-driven GFP expression.

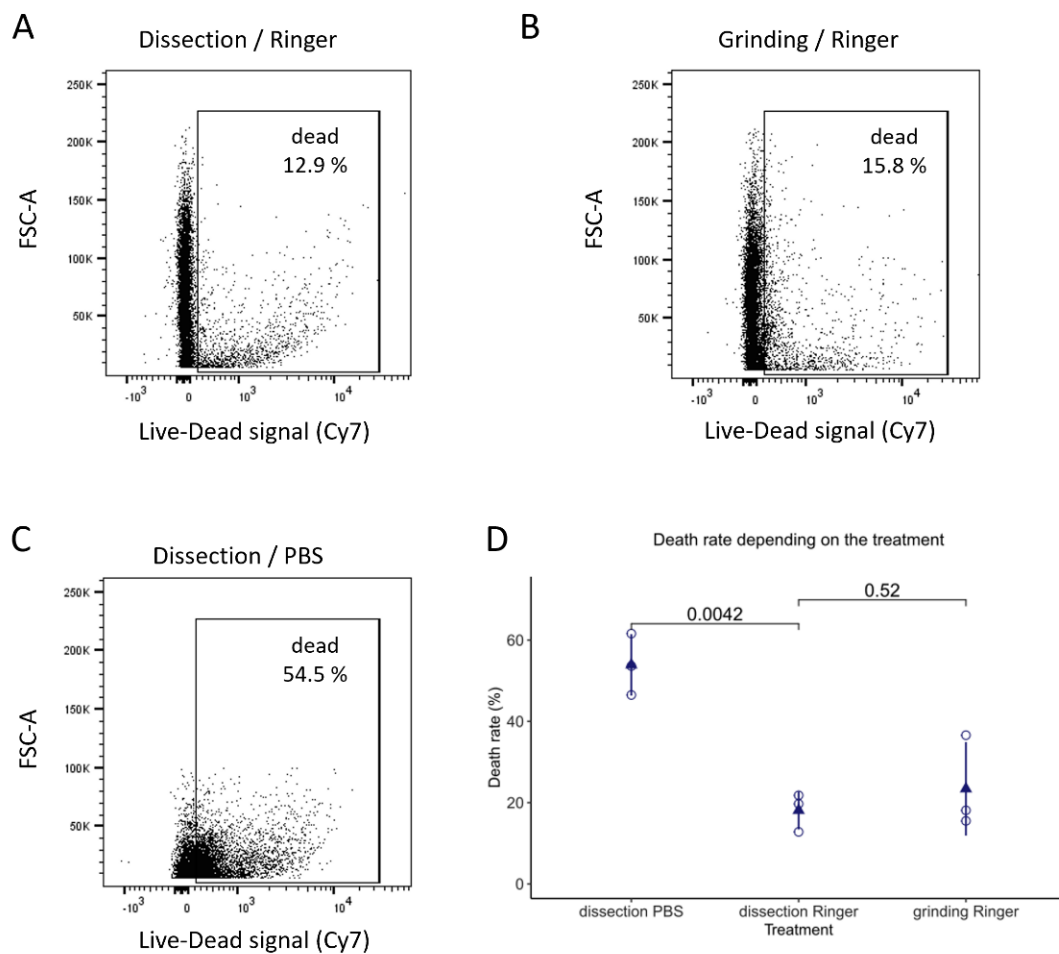
After centrifugation, wing imaginal discs, along with eye and antenna discs and salivary gland pieces, were found at the 15:20% interface. Some more wing imaginal discs could be recovered at the 20:25% interface, but we chose not to recover them because the interface also contains the vast majority of salivary glands and mouth hooks. For our applications, cells from imaginal discs were easy to separate from salivary gland cells, which are the primary contaminants in this case.

### 3.1.3. Finding the Best Buffer for Isolating and Maintaining Intact Imaginal Discs

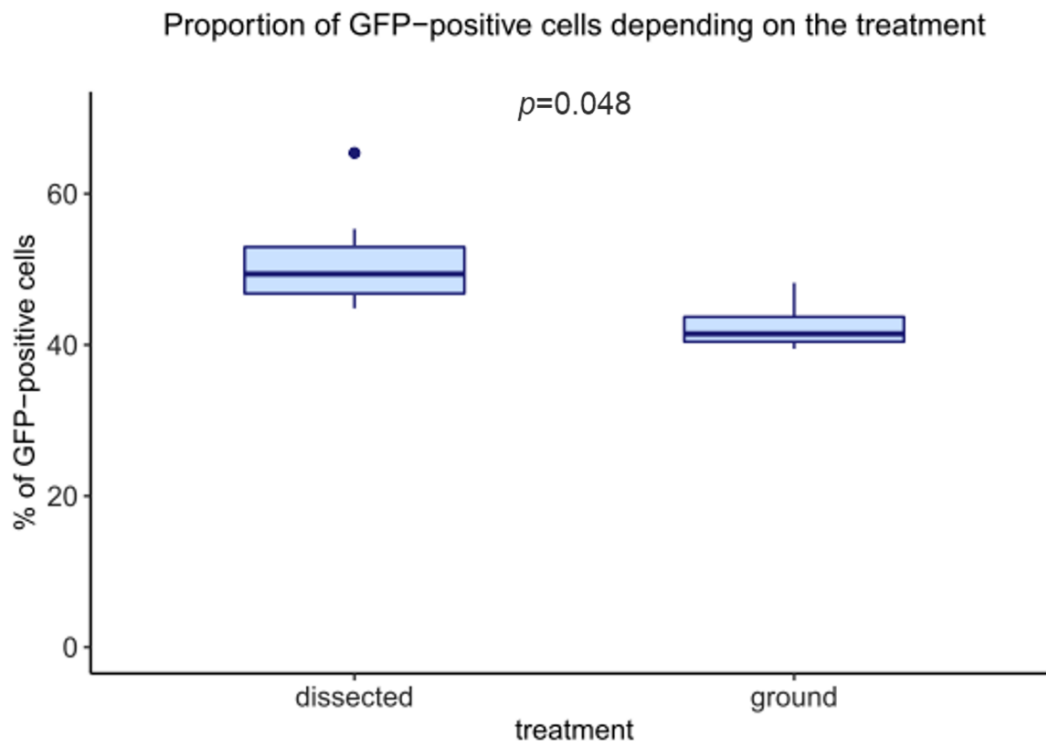
An important difference between the protocols of Marty et al. 2014 [16] and Fristrom and Mitchell 1965 [14] is the buffer used throughout the enrichment process. The original article by Fristrom and Mitchell [14] used a Ringer solution [21], whereas the most recently used solution was PBS. We started setting up the protocol using PBS, the cheapest and most commonly used buffer. However, after Ficoll gradient enrichment, the recovered discs looked dehydrated (data not shown). Since discs recovered in Fristrom and Mitchell's article [14], for which Ringer solution was used, could be successfully transplanted, we considered switching to this buffer. After replacing PBS with Ringer solution during manual dissections, we observed that discs had a better appearance. We performed a viability assay using a dye that enters damaged cells with a permeabilized membrane to assess whether imaginal disc cells were indeed in more physiological conditions in a Ringer solution than in PBS.

We used this assay to determine whether one of the buffers, PBS or Ringer, is more damaging than the other. Confirming our observation, PBS is a much more damaging agent as cells obtained by dissection in PBS were more than twice as frequently labeled compared to cells treated in Ringer solution (Figure 3D). In addition, it could be envisioned that the mass enrichment process may be more damaging than dissection. Grinding could exert more mechanical stress than dissection, while using the Ficoll polymer may induce osmotic stress. The analysis of the impact of the recovery method showed that the mass enrichment process is not more damaging for the cells than dissection in this assay (Figure 3, compare A and B and see D). To further ensure that the mass enrichment process did not induce ectopic apoptosis, we performed apoptotic cell labeling using an antibody against activated caspases (anti-cleaved Dcp-1) [22]. This staining showed that mass enrichment does not induce more apoptosis than classical manual dissection (Supplementary Figure S3). The slight difference between those two conditions could be explained by the difference in the buffer used (PBS for dissection, Ringer for grinding). Finally, switching PBS to the Ringer solution improved the quality of the material recovered and the enrichment process (compare Figure 2A protocol "C" with Figure 2A protocol "D" and "E") as it decreased the number of contaminants while increasing the number of discs recovered.

We continued the quality control of the material recovered by mass enrichment by analyzing the number of GFP-positive cells using flow cytometry. This analysis was carried out on larvae expressing the GFP in the wing disc under the control of the *vg-GAL4* driver. This driver induces GFP expression in a wide band at the dorso-ventral frontier, representing roughly 40% of the cells. Visual inspection of the GFP pattern readily allows the recognition of the *vg* pattern, confirming that discs do not undergo such stress that would lead to GFP extinction. Since salivary gland cells are bigger than wing imaginal disc cells [23], it is easy to discard them; after dissociation, passing the cell suspension through a 40 µm filter is enough to retain the bigger cells. After dissociation, filtration, and flow cytometry analysis of either mass-enriched or dissected wing imaginal discs, the percentage of GFP-positive cells was measured (Figure 4).



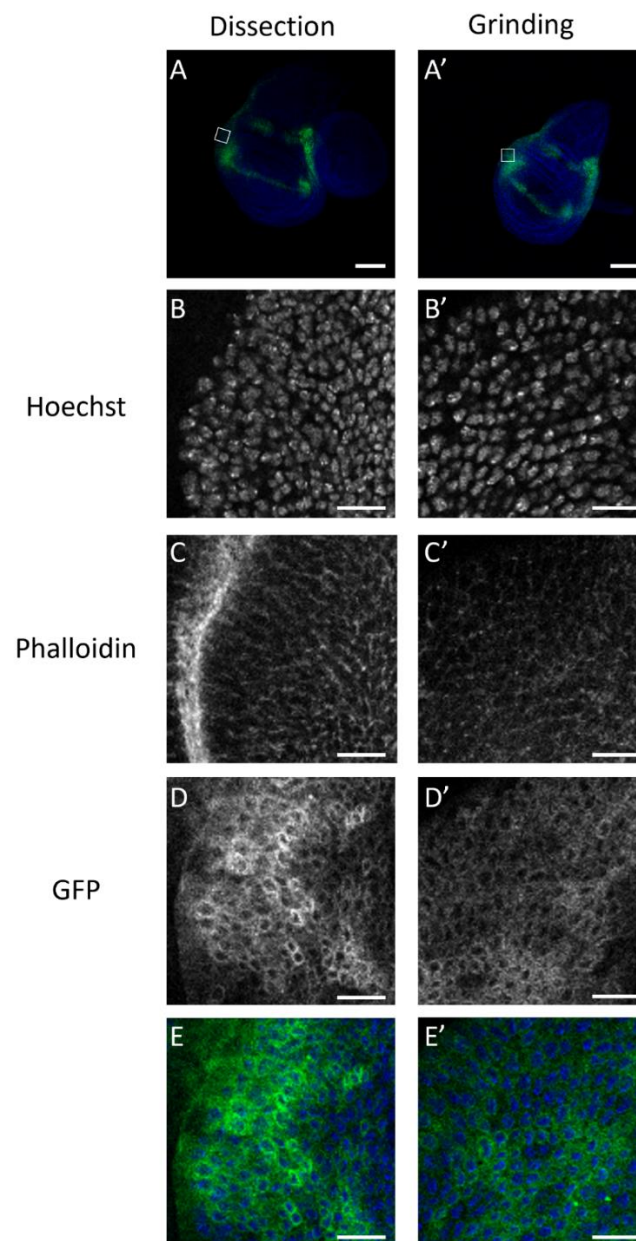
**Figure 3.** Effects of buffer and disc recovery method on cell viability. Wing imaginal discs obtained by either dissection (A,C) or grinding (B) in 1X Ringer (A,B) or 1X PBS (C) were dissociated, filtered through a 40  $\mu\text{m}$  sieve, and stained with a viability dye (Fixable Viability Dye eFluor™ 780) that enters damaged cells with a permeabilized membrane. The dye intensity was measured by flow cytometry (635 nm) to assess the death rate (expressed in % of the total number of individual cells). (A–C) Cytograms correspond to FSC = cell size versus Live-Dead signal (Cy7). (D) The experiment was performed three times, and each point on the dot plot represents the death rate for one experiment. The triangular point represents the mean for each data set. Statistical significance was determined by one-way ANOVA using Bonferroni’s correction for multiple comparison testing.



**Figure 4.** Proportion of GFP-positive cells in wing imaginal discs recovered by mass-enrichment protocol or dissection. *vg* > GFP expressing wing discs were recovered by either dissection or using our mass enrichment protocol, were dissociated, and the number of GFP-positive cells was measured by flow cytometry. A Wilcoxon test was performed for comparison ( $n = 12$ ).

Mass enrichment of discs led to a slight decrease in the proportion of GFP-positive cells compared with dissected discs (from 47.7% to 42.7%), but this decrease was expected as ground samples were not 100% pure. This fraction contains other organs (such as other discs) whose cells do not express GFP, thus increasing the GFP-negative fraction. This limited decrease indicates that the wing imaginal disc fraction recovered by mass enrichment is relatively pure. This result confirms our visual assessment and is supported by the fact that after dissociation and filtration, whatever the protocol used, the percentage of cells recovered compared to the total number of events counted by cytometry is not significantly different (Figure 4).

The tissue structure was also explored using fluorescent staining of *vg-GAL4*, mCD8-GFP, wing imaginal discs to stress the quality of the wing imaginal discs recovered by the mass enrichment protocol. Hoechst and phalloidin-ATTO 655 stained the nucleus and the plasma membrane, respectively. mCD8GFP protein is known to localize at the plasma membrane. The experiment was performed in parallel on wing discs isolated by dissection or mass enrichment protocol. The results in Figure 5 show that the wing disc cells obtained by the mass enrichment protocol have a normal structure. The results in Figure 5 show that the wing disc cells obtained by the mass enrichment protocol have a normal structure. The slight differences in intensity or morphology observed for staining within a single image (e.g., Figure 5C) or between images obtained under the two conditions (dissection or mass enrichment protocol) can be explained by the fact that the imaginal disc is a folded pseudo-columnar epithelium. Thus, the image plane is not at the same apical-basal level for all cells. The results do not reveal any significant difference in the quality of the discs recovered by both methods.

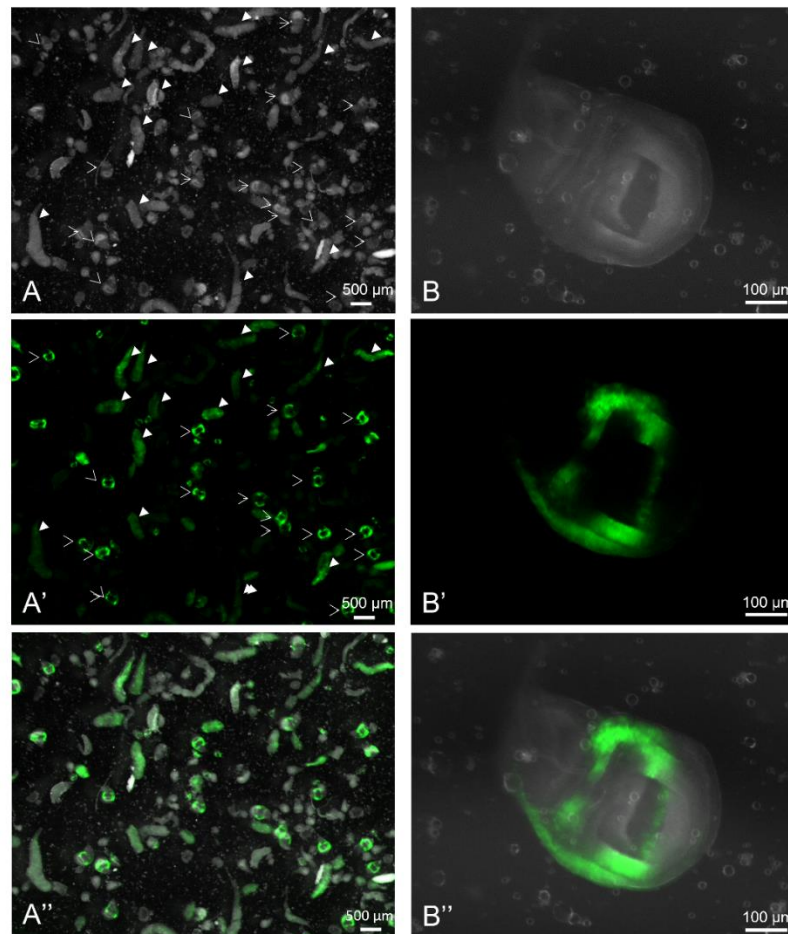


**Figure 5.** Cell structure in wing imaginal discs recovered by dissection or mass-enrichment protocol. Cell structure was assessed in wing imaginal discs recovered by dissection (**left** column) or mass-enrichment protocol (“grinding”, **right** column). Confocal images were acquired with 10X objective, and only a slice of this pseudocolumnar epithelium is shown (**A,A'**). Images (**B–E'**) correspond to the same wing imaginal discs and were acquired with a 63X objective. They show a zoomed-in view of the white boxes in (**A,A'**). Nuclei were visualized using Hoechst (**B,B'**), plasma membrane using phalloidin ATTO655 (**C,C'**). GFP is coupled to mCD8 protein, whose expression is driven by the *vg-GAL4* driver (**D,D'**). Representative overlays of nuclei staining (blue) and GFP (green) are shown in whole discs (**A,A'**) or in an enlargement of the wing imaginal disc posterior zone. Images E and E' are an overlay of ((**B,D,B'**) and (**D'**)), respectively. Scale bars correspond to 100  $\mu\text{m}$  (**A,A'**) or 10  $\mu\text{m}$  (**B–E'**).

### 3.2. Final Protocol (See Supplementary File S1 for Details)

Flushing the side of the tubes with water allows the recovery of L3 larvae from synchronized egg-laying. After rinsing in water to remove the growth medium, larvae are transferred in GentleMACS C tubes and ground in Ringer solution. The ground material

is filtered through several strainers of decreasing meshes to separate organs by size; this mainly removes cuticles and fat bodies. These steps are repeated five times to ensure the good separation of organs without damaging them. The filtered organ suspension (100–200  $\mu\text{m}$ ) then undergoes a sedimentation step to remove gut pieces, fat body, and some other non-discs organs. Finally, a Ficoll density gradient allows the separation of imaginal discs from other organs. Salivary gland pieces constitute the primary contaminant (Figures 1 and 6). Wing imaginal cells can be isolated and separated from salivary gland cells using a dissociation and a filtration step.



**Figure 6.** Yield and purity of the mass enrichment protocol. Overview of the resulting material obtained at the end of the procedure under white light (A,B) or GFP-exciting light (A',B'). (B–B'') show wing imaginal discs (the posterior compartment is on the bottom) corresponding to a 6.4 times magnification of (A–A''). Empty arrowheads point to wing imaginal discs, and filled arrowheads point to salivary glands. Procedure (protocol E) was performed on larvae expressing GFP under the control of the *vestigial* driver.

#### 4. Discussion

The increasing use of global approaches to study gene expression or protein regulation justifies the development of mass purification techniques for *Drosophila* imaginal discs. By studying protocols from the 1960–1970s, we set up a protocol that allows recovery of third-instar larvae wing imaginal discs without manual dissection. To do so, we had to set up every step, from the grinding of larvae to the density gradient (Figure 2 and Supplementary Figure S1). During our protocol set-up, a compromise was made between purity and quantity. Therefore, the selected protocol allows obtaining fractions consisting of 50% of wing imaginal discs (Figure 2). There are two types of contaminants: salivary



glands, the most abundant ones, and rare other imaginal discs (mainly eye-antenna discs). However, if pure fractions are needed, wing imaginal discs can be easily picked out with pliers. Moreover, if the objective is to recover the cellular fraction corresponding to the discs, a dissociation step followed by filtration allows removing the cells of the salivary glands.

Indeed, since they are much bigger than those of imaginal discs, passing the solution through a 40  $\mu\text{m}$  strainer is enough to separate them from the wing imaginal disc cells of interest. The advantages and disadvantages of the protocol presented in this article are summarized in Supplementary Figures S4 and S5, which compare the pros and cons of some existing protocols.

The protocol we selected has a yield of 11.5%, which is low compared to manual dissection yield, considered around 85%, by taking into account the organs that might be lost or damaged during the process. However, this mass enrichment protocol presents advantages over manual dissection.

- (i) We used around 600 larvae as an input to obtain, on average, 140 wing imaginal discs. Depending on the researcher, this amount can be dissected in about two to three hours. However, if more wing imaginal discs are needed, the dissection time increases proportionally. One of the advantages of the mass enrichment protocol is that it can process several samples simultaneously by using parallel filtration montages and doing the filtration and centrifugations (sedimentation and Ficoll density gradient) at the same time (Figure 1). Therefore, its duration is only slightly impacted and should not take more than three hours.
- (ii) When high amounts of organs are needed, multiple researchers usually dissect them together. As each researcher has their own method and ease of dissection, which may be influenced by the time spent dissecting, this may pose reproducibility problems. The enrichment protocol is standardized and can be performed entirely by one researcher, reducing reproducibility issues, which constitutes another advantage over manual dissection.
- (iii) Since dissection allows only one sample to be processed at a time, this induces a delay in the treatment of each condition. Our protocol does not present this disadvantage, as several samples can be treated almost simultaneously. Thus, it allows a more homogenous treatment between conditions.

As mentioned in Section 3.1, this protocol requires upstream preparation: for all the individuals to be at the same stage, we recommend transferring the adults to fresh tubes every 24 h. As a reference, we obtain 2.5 mL of larvae (around 600 L3 larvae) with six vials of stable stock lines. When high amounts of material or numerous conditions are needed, these preparation steps can be facilitated using larger vials or even bottles. According to offspring viability, the number of necessary vials is up to the researcher's decision. While the amount of work necessary to produce enough larvae seems large compared to that required for dissection, this technique allows the larvae to be harvested simultaneously, avoiding the treatment delay between each larva. Pooling all the larvae from a larger number of tubes also decreases the batch effect of each vial.

If larvae of interest are obtained by crossing and represent only 50% of the progeny, the number of vials must be doubled to obtain an equal amount of wing imaginal discs. Depending on the parents' genotypes, only a fraction of the offspring may be of interest in certain conditions. Collecting these individuals based on their specific phenotypes (Tb, CyOGFP, etc.) is possible, but this would add a time-consuming step. In this case, using a fluorescent protein expressed in the population of interest presents two advantages. Firstly, it allows the identification of discs of interest at the end of the mass enrichment without selecting the larvae before starting the protocol. Secondly, when a large amount of disc is required to isolate cells of interest, the dissociation and filtration step allows obtaining wing imaginal cells that can be sorted by flow cytometry. Depending on the experimental design, other solutions can be considered, such as tagged proteins of interest and magnetic cell sorting [24].

Interestingly, this protocol can also be adapted to isolate other *Drosophila* organs by modifying the filtration or the Ficoll gradient steps using the indications detailed in this specific part. Moreover, to facilitate the tracking of organs throughout the gradient, we adapted a beta-galactosidase staining protocol without fixation to color the organs blue and have a visual clue of their localization [20] (Supplementary Figure S2). This staining can be applied to help isolate other organs.

Older protocols were able to graft the collected organs, which became normal adult structures [14]. We did not test this ability. However, by using a cell viability assay (Figure 3) and controlling the apoptosis level (Supplementary Figure S3), we showed that this protocol does not induce more damage than the dissection of wing imaginal discs. The mass enrichment protocol is standardized and achievable quickly, allowing a more homogenous treatment of wing imaginal discs. In the case of experiments that necessitate the analysis of a large number of imaginal discs, this constitutes a considerable advantage of a mass purification protocol. Remarkably, our analysis of cell viability strongly emphasizes the importance of cell buffer (Figure 3) and shows that cell survival is higher in Ringer buffer than in the PBS buffer, yet is commonly used for dissections.

As another quality control, we compared the expression of GFP in wing imaginal discs in which the vestigial driver drove GFP expression. Visual examination showed no notable difference in GFP levels between discs (see Figures 5 and 6 for GFP-expressing wing imaginal discs recovered through the mass enrichment protocol). The percentage of GFP-positive cells estimated through flow cytometry corresponding to discs recovered by mass enrichment protocol and dissection is quite similar. The slight observed difference can be easily explained by the presence of a small number of other organs, such as eye-antennae discs, which “dilute” the GFP-positive cells when using a mass enrichment protocol. Nevertheless, GFP analysis shows that obtained discs are homogenous and that the purification protocol is reproducible. In many approaches, the cells of interest represent only a subpopulation of the cells in the imaginal disc. In this case, obtaining a sufficient number of cells is difficult to achieve by manual dissection. One of the interests of this approach is to allow a large quantity of a subpopulation of imaginal disc cells to be rapidly obtained by coupling the mass enrichment protocol to cell sorting by flow cytometry.

## 5. Conclusions

In summary, this paper shows that this mass enrichment protocol allows the rapid purification of a large quantity of wing imaginal discs of equivalent quality to manually dissected discs for multiple samples simultaneously and by a single investigator. Wing imaginal discs recovered by mass enrichment allow the isolation of wing imaginal disc cells suitable for numerous applications such as cytometry analyses (Figures 3 and 4), transcriptomics, and proteomics.

**Supplementary Materials:** The following supporting information can be downloaded at <https://www.mdpi.com/article/10.3390/biology11101384/s1>, Figure S1: Overview of the GentleMACS™ programs tested; Figure S2: Example of a Ficoll gradient with  $\beta$ -galactosidase-stained wing imaginal discs; Figure S3: Apoptosis staining in the wing discs according to the disc enrichment method; Figure S4: Comparison of wing imaginal discs isolation protocols; Figure S5: Summary table of pros and cons of available imaginal discs isolation protocols; File S1: Mass enrichment protocol.

**Author Contributions:** Conceptualization, J.C. and I.G.; methodology, M.H., J.C. and I.G.; validation, M.H., J.C. and I.G.; formal analysis, M.H., J.C. and I.G.; investigation, M.H., J.d.N., J.C. and I.G.; writing—original draft preparation, M.H.; writing—review and editing, M.H., J.C. and I.G.; supervision, J.C. and I.G.; project administration, I.G.; funding acquisition, I.G. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by UVSQ, EPHE, and private funding collected by the UVSQ Foundation.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** The data that support the findings of this study are available from the corresponding authors, I.G. and J.C., upon reasonable request.

**Acknowledgments:** We thank Simon Glatigny for helpful discussions concerning cytometry experiments, Christine Wintz for technical assistance, Aurore Rincheval-Arnold and Mélanie Fages for their help in confocal experiments, and Sébastien Gaumer for his help in statistical analysis and his critical reading of the manuscript. Image acquisition, image analysis, and cytometry experiments were performed at the CYMAGES imaging facility, part of UVSQ-Paris Saclay University. The authors also greatly acknowledge Anne-Laure Raveu from the CYMAGES facility.

**Conflicts of Interest:** The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

## References

1. Brand, A.H.; Perrimon, N. Targeted Gene Expression as a Means of Altering Cell Fates and Generating Dominant Phenotypes. *Development* **1993**, *118*, 401–415. [[CrossRef](#)] [[PubMed](#)]
2. Wu, J.S.; Luo, L. A Protocol for Mosaic Analysis with a Repressible Cell Marker (MARCM) in *Drosophila*. *Nat. Protoc.* **2006**, *1*, 2583–2589. [[CrossRef](#)]
3. Celniker, S.E.; Rubin, G.M. The *Drosophila melanogaster* Genome. *Annu. Rev. Genom. Hum. Genet.* **2003**, *4*, 89–117. Available online: <https://www.annualreviews.org/doi/10.1146/annurev.genom.4.070802.110323> (accessed on 22 July 2022). [[CrossRef](#)] [[PubMed](#)]
4. Beira, J.V.; Paro, R. The Legacy of *Drosophila* Imaginal Discs. *Chromosoma* **2016**, *125*, 573. [[CrossRef](#)] [[PubMed](#)]
5. Gou, J.; Stotsky, J.A.; Othmer, H.G. Growth Control in the *Drosophila* Wing Disk. *WIREs Syst. Biol. Med.* **2020**, *12*, e1478. [[CrossRef](#)]
6. Denton, D.; O’Keefe, L.; Kumar, S. Chapter Fifteen—*Drosophila* as a Model to Understand Autophagy Deregulation in Human Disorders. In *Progress in Molecular Biology and Translational Science*; Martinez, A.B., Galluzzi, L., Eds.; Autophagy in Health and Disease; Academic Press: Cambridge, MA, USA, 2020; Volume 172, pp. 375–409.
7. Morata, G.; Calleja, M. Cell Competition and Tumorigenesis in the Imaginal Discs of *Drosophila*. *Semin. Cancer Biol.* **2020**, *63*, 19–26. [[CrossRef](#)] [[PubMed](#)]
8. Lewis, E.B. A Gene Complex Controlling Segmentation in *Drosophila*. *Nature* **1978**, *276*, 565–570. [[CrossRef](#)]
9. Justice, R.W.; Zilian, O.; Woods, D.F.; Noll, M.; Bryant, P.J. The *Drosophila* Tumor Suppressor Gene Warts Encodes a Homolog of Human Myotonic Dystrophy Kinase and Is Required for the Control of Cell Shape and Proliferation. *Genes Dev.* **1995**, *9*, 534–546. [[CrossRef](#)]
10. Loubiere, V.; Delest, A.; Schuettengruber, B.; Martinez, A.-M.; Cavalli, G. Chromatin Immunoprecipitation Experiments from Whole *Drosophila* Embryos or Larval Imaginal Discs. *Bio-Protocol* **2017**, *7*, e2327. [[CrossRef](#)]
11. Diaz-Garcia, S.; Ahmed, S.; Baonza, A. Analysis of the Function of Apoptosis during Imaginal Wing Disc Regeneration in *Drosophila melanogaster*. *PLoS ONE* **2016**, *11*, e0165554. [[CrossRef](#)]
12. Clavier, A.; Baillet, A.; Rincheval-Arnold, A.; Coléno-Costes, A.; Lasbleiz, C.; Mignotte, B.; Guénel, I. The Pro-Apoptotic Activity of *Drosophila* Rbf1 Involves DE2F2-Dependent Downregulation of Diap1 and Buffy mRNA. *Cell Death Dis.* **2014**, *5*, e1405. [[CrossRef](#)] [[PubMed](#)]
13. Clavier, A.; Rincheval-Arnold, A.; Baillet, A.; Mignotte, B.; Guénel, I. Two Different Specific JNK Activators Are Required to Trigger Apoptosis or Compensatory Proliferation in Response to Rbf1 in *Drosophila*. *Cell Cycle* **2016**, *15*, 283–294. [[CrossRef](#)] [[PubMed](#)]
14. Fristrom, J.W.; Mitchell, H.K. The Preparative Isolation of Imaginal Discs from Larvae of *Drosophila melanogaster*. *J. Cell Biol.* **1965**, *27*, 445–448. [[CrossRef](#)] [[PubMed](#)]
15. Zweidler, A.; Cohen, L.H. Large-Scale Isolation and Fractionation of Organs of *Drosophila melanogaster* Larvae. *J. Cell Biol.* **1971**, *51*, 240–248. [[CrossRef](#)] [[PubMed](#)]
16. Marty, F.; Rockel-Bauer, C.; Simigdala, N.; Brunner, E.; Basler, K. Large-Scale Imaginal Disc Sorting: A Protocol for “Omics”-Approaches. *Methods* **2014**, *68*, 260–264. [[CrossRef](#)]
17. de Noiron, J.; Hoareau, M.; Colin, J.; Guénel, I. Apoptosis Quantification in Tissue: Development of a Semi-Automatic Protocol and Assessment of Critical Steps of Image Processing. *Biomolecules* **2021**, *11*, 1523. [[CrossRef](#)]
18. Pfeiffer, B.D.; Ngo, T.-T.B.; Hibbard, K.L.; Murphy, C.; Jenett, A.; Truman, J.W.; Rubin, G.M. Refinement of Tools for Targeted Gene Expression in *Drosophila*. *Genetics* **2010**, *186*, 735–755. [[CrossRef](#)]
19. Hrdlicka, L.; Gibson, M.; Kiger, A.; Micchelli, C.; Schober, M.; Schöck, F.; Perrimon, N. Analysis of Twenty-Four Gal4 Lines in *Drosophila melanogaster*. *Genesis* **2002**, *34*, 51–57. [[CrossRef](#)]
20. Tomizawa, M.; Tsumaki, K.; Sone, M. Characterization of the Activity of  $\beta$ -Galactosidase from *Escherichia coli* and *Drosophila melanogaster* in Fixed and Non-Fixed *Drosophila* Tissues. *Biochim. Open* **2016**, *3*, 1–7. [[CrossRef](#)]
21. Ephrussi, B.; Beadle, G.W. A Technique of Transplantation for *Drosophila*. *Am. Nat.* **1936**, *70*, 218–225. [[CrossRef](#)]

22. Sarkissian, T.; Timmons, A.; Arya, R.; Abdelwahid, E.; White, K. Detecting Apoptosis in Drosophila Tissues and Cells. *Methods* **2014**, *68*, 89–96. [[CrossRef](#)] [[PubMed](#)]
23. Brun, S.; Rincheval-Arnold, A.; Colin, J.; Risler, Y.; Mignotte, B.; GuénaI, I. The Myb-Related Gene Stonewall Induces Both Hyperplasia and Cell Death in Drosophila: Rescue of Fly Lethality by Coexpression of Apoptosis Inducers. *Cell Death Differ.* **2006**, *13*, 1752–1762. [[CrossRef](#)] [[PubMed](#)]
24. Iyer, E.P.R.; Iyer, S.C.; Sulkowski, M.J.; Cox, D.N. Isolation and Purification of Drosophil Peripheral Neurons by Magnetic Bead Sorting. *J. Vis. Exp. JoVE* **2009**, *34*, e1599. [[CrossRef](#)]

## **Materials and Reagents**

- A. Larvae recovery and grinding
  1. *Drosophila* third instar larvae in their rearing tubes
  2. Water
  3. 1X Ringer (3 mM NaCl, 182 mM KCl, 46 mM CaCl<sub>2</sub>)
  4. GentleMACS™ C tubes (Miltenyi, [130-093-237](tel:130-093-237))
  
- B. Filtration, sedimentation, and gradient
  1. 50 mL polypropylene tubes (Corning-Falcon, 352070)
  2. 25 mL pipettes (Corning-Falcon, 356535)
  3. 10 mL pipettes (Corning-Falcon, 357551)
  4. 1X Ringer (3 mM NaCl, 182 mM KCl, 46 mM CaCl<sub>2</sub>)
  5. pluriStrainers 500 µm, 300 µm, 200 µm, 100 µm (PluriSelect, 43-50500, 43-50300, 43-50200, 43-50100)
  6. Funnel (PluriSelect, 42-50000)
  7. Ficoll PM400 30% (w/v), 25% (w/v), 20% (w/v), 15% (w/v) solutions prepared in 1X Ringer solution from (Thermofisher 11590724)
  
- C. Rehydration and pelleting
  1. 10 mL sterile pipettes (Corning-Falcon, 357551)
  2. 1X Ringer (3 mM NaCl, 182 mM KCl, 46 mM CaCl<sub>2</sub>)
  3. SmartStrainers 70 µm (Miltenyi, [130-098-462](tel:130-098-462))
  4. Sterile filter tips 1000 µL (Starlab, S1126-7810)
  5. 1.5 mL tubes (Eppendorf, 0030120086)
  
- D. Dissociation and filtration
  1. 1.5 mL tubes (Eppendorf, 0030120086)
  2. 1X Ringer (3 mM NaCl, 182 mM KCl, 46 mM CaCl<sub>2</sub>)
  3. Sterile filter tips 1000 µL (Starlab, S1126-7810)
  4. Protease 0.1% solution diluted from in 1X Ringer from (Sigma, p8811)
  5. PluriStrainer Mini 40 µm (PluriSelect 43-10040)

## **Equipment**

1. Wash bottles
2. gentleMACS™ Dissociator (Miltenyi)
3. Calibration checked pipettes 20-200 µL, 100-1000 µL
4. Eppendorf 5810R refrigerated centrifuge with swing bucket rotor (A-4-62) or equivalent with settings for acceleration and deceleration
5. Eppendorf ThermoMixer® or other heating and shaking device

## **Procedure**

- A. Larvae collection and grinding
  1. Flush the sides of the tubes with a wash bottle filled with water to collect larvae in an empty beaker.

*Note. To ensure that all larvae are third instar, transfer the parents to a fresh medium every 24 hours and take care not to overcrowd the tubes. See Note 1 for the number of larvae needed.*
  2. Rinse the larvae with water to get rid of the remaining medium residues.
  3. Transfer the larvae to a 500 µm strainer to eliminate the water.
  4. Transfer the larvae to a GentleMACS C tube containing around 2.5 mL of 1X Ringer.

*Note. For optimal grinding, do not put over the 2.5 mL graduation of larvae per tube, otherwise, all larvae may not be correctly ground without adding grinding cycles and therefore, adding time to the duration of the protocol.*

5. Complete the tube with 1X Ringer up to the 10 mL graduation.
6. Grind larvae using the "brain\_02\_02" program of the GentleMACS™. (100 rpm, 30 sec)  
*Note. Centrifuge quickly the C tube to pellet the larvae debris that could be stuck in the cap helix*

B. Filtration, sedimentation, and gradient

1. Filter the ground material through an assembly of 500, 300, and 200  $\mu\text{m}$  strainers (see Fig. 1) and recover the filtrate in a 50 mL tube (named "1", which contains the imaginal discs released).

*Note. Flush the helix with a wash bottle filled with 1X Ringer to be sure to recover most material.*

2. Put back in the C tube what did not pass the 500  $\mu\text{m}$  sieve and add 1X Ringer up to 10 mL and repeat steps A6-B1 until all larvae are properly disrupted (usually 5 cycles are sufficient)

Filtrate the content of the "1" tube through a 100  $\mu\text{m}$  strainer and recover what is retained in the strainer (size > 100  $\mu\text{m}$ ) in a fresh 50 mL conical tube named "2" by turning it upside down on top of the tube *Note. In the first round of filtration, the liquid content that passes 100  $\mu\text{m}$  is rich in lipids. It can be further filtered through a 30  $\mu\text{m}$  strainer to get a solution very useful for pipette coating thus avoiding the organs to stick to the pipettes during downstream steps.*

*Note. If the "2" tube is full, pass its content through the 100  $\mu\text{m}$  filter to discard some liquid.*

3. Centrifuge the "2" tube (containing the 100-200  $\mu\text{m}$  material) at 200 xg 1 min 4°C.
4. Discard the supernatant and resuspend the pellet in 20 mL of 1X Ringer.
5. Centrifuge at 50 xg 1 min 4°C.
6. Discard the supernatant and resuspend the pellet in 20 mL of 1X Ringer.
7. Centrifuge at 10 xg 1 min 4°C.
8. Wait for 20 seconds.
9. Discard the supernatant and resuspend the pellet in 20 mL of 1X Ringer.
10. Repeat steps B7-B8
11. Discard the supernatant and resuspend the pellet in 7.5 mL of 1X Ringer.
12. Add 2.5 mL 30 % Ficoll (w/v) for a 7.5 % Ficoll final concentration.
13. Mix well and centrifuge at 50 xg 5 min 4°C (acc. 9, dec. 3)
14. Discard the supernatant and resuspend the pellet in 4 mL 1X Ringer.
15. Add 2 mL Ficoll 30% (w/v) for 10 % Ficoll final concentration and mix well. Store at +4°C while preparing the gradient.

*Note. If there is a lot of material, increase the volume of the 10 % layer (Ringer + Ficoll) otherwise the layer will be too heavy for the gradient and will go under the 15 % layer.*

16. In a new 50 mL tube, add 9 mL of 25 % Ficoll (w/v).
17. Centrifuge 3220 xg, 3 min, 4°C (acc. 3, decc. 1)
18. Carefully pour the 20% (w/v) layer (14 mL) on top of the 25% (w/v) layer.
19. Carefully pour the 15% (w/v) layer (14 mL) on top of the 20% (w/v) layer.
20. Coat the pipette with the solution kept in step B2.
21. Carefully pour the layer containing the material (from step B15).
22. Centrifuge 80 xg, 20 min, 4°C (acc. 3, decc. 1)

### C. Rehydration and pelleting

1. Coat the pipette with the solution kept in step B2.
2. Carefully pipet the interface between the 15 and 20% (w/v) layers (2<sup>nd</sup> and 3<sup>rd</sup> layers).
3. Collect the interface in a 70  $\mu$ m strainer.
4. Rinse with 1X Ringer.
5. Place the strainer with the material in a box containing a large volume of 1X Ringer (about 20 mL, depending on the size of the box) so that the material in the strainer is fully covered with buffer.
6. Leave the material to rehydrate for at least 20 min at RT.

*Note: you can check at this point how many wing imaginal discs you recovered with a stereomicroscope.*

7. Transfer the material to a 1.5 mL microtube.

*Note. To avoid losing material sticking to the tips, flush the material out of the strainer with some Ringer into the microtube rather than pipetting it.*

8. Centrifuge 4000  $xg$ , 3 min, 4°C: wing imaginal discs are in the pellet.

*Note: Stop here if you need undissociated wing imaginal discs.*

### D. Dissociation

1. Discard supernatant from the C8 step.
2. Resuspend the pellet in 1 mL of 0,1%.protease (w/v)
3. Incubate for 20 min at 25°C in a Thermomixer with gentle shaking (300 rpm).
4. Complete the dissociation with gentle ups and downs with a 1000  $\mu$ L tip.
5. Centrifuge 4000  $xg$ , 3 min, 4°C: wing imaginal disc cells are in the pellet.
6. Resuspend the pellet in Ringer 1X.
7. Filter the solution through a 40- $\mu$ m filter: wing imaginal disc cells pass through, and salivary gland cells stay in the filter.

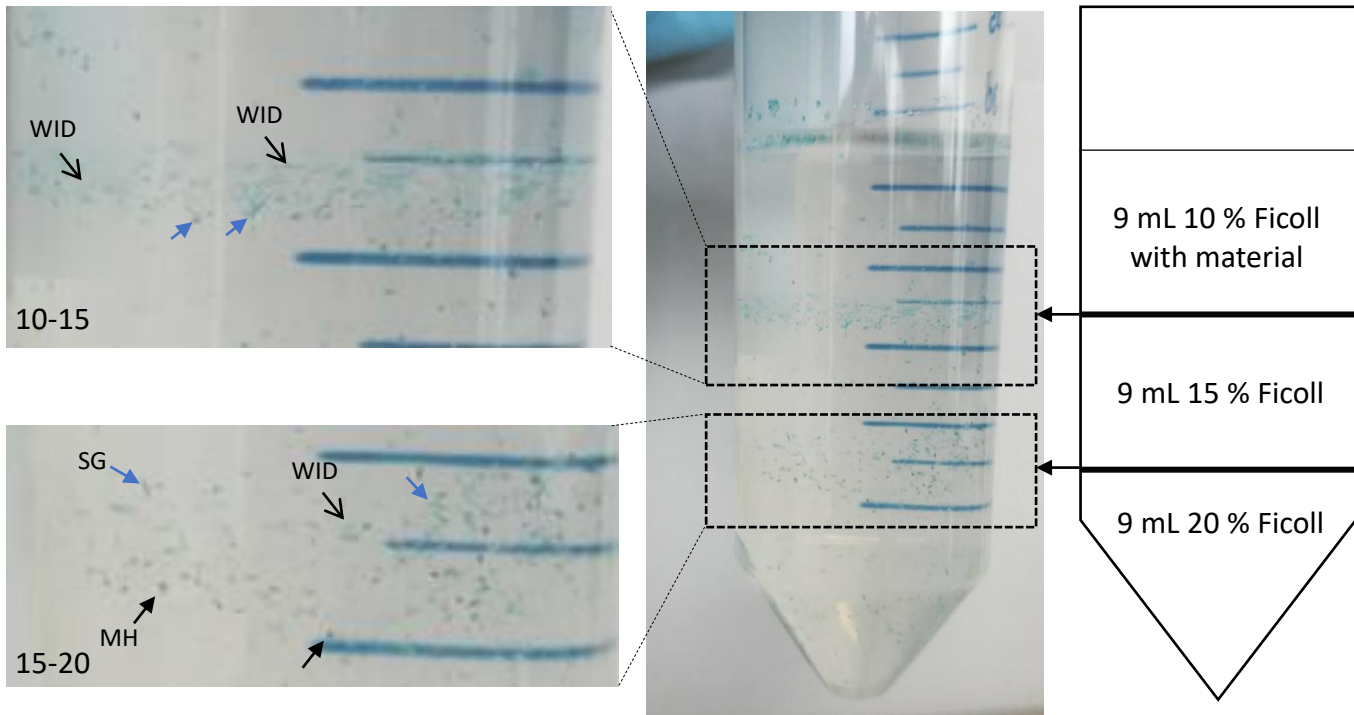
Note 1: It takes around 600 larvae (= 2,5 mL) to get 100 to 150 wing imaginal discs.

Test		Larvae disrupted?	Intact discs?
M	« liver » (1 or 2)	-	NT
M	« liver » (1 or 2) + « brain » (1)	+	-
M	« brain » (2)	++	-
C	« liver » (2)	-	NT
C	« liver » (1) + « brain » (2)	+	-
C	« brain » (2)	++	++

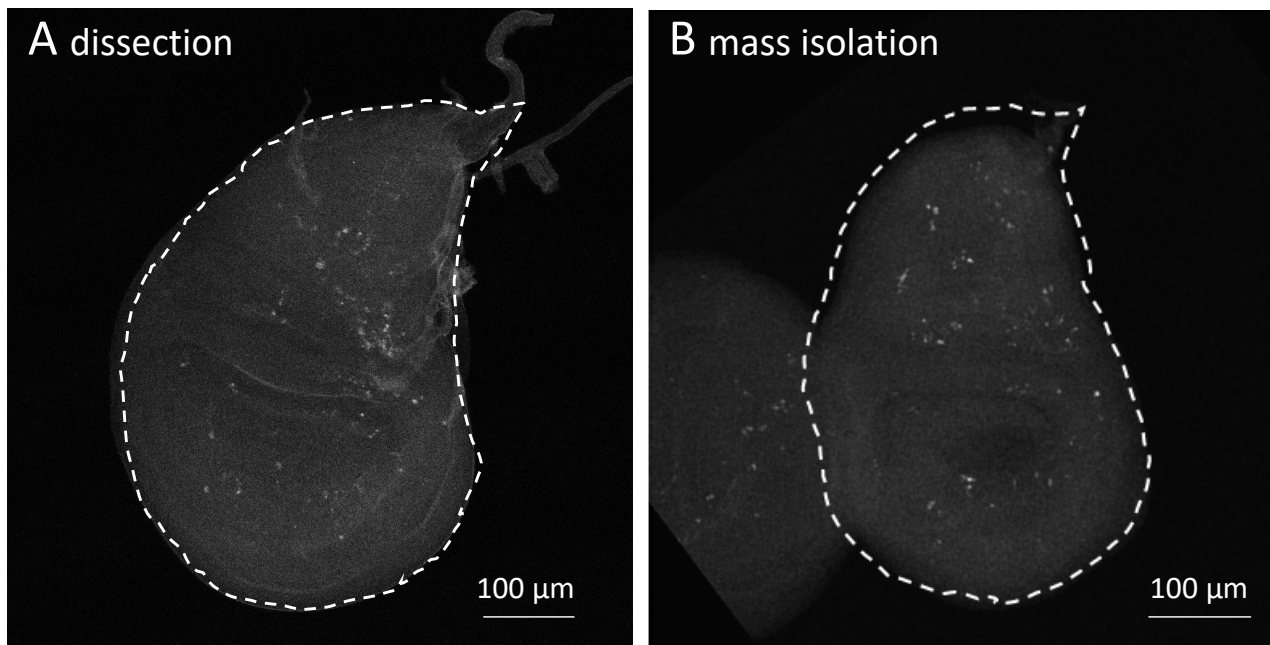
Liver: 1(48 rpr, 15 s) or 2(78 rpr, 24 s) / Brain: 1(116 rpr, 36 s) or 2(100 rpr, 30 s)

**Supplementary Figure S1: Overview of the GentleMACS™ programs tested.** The GentleMACS™ (Miltenyi) device can be used with either C or M kinds of tubes and display various pre-registered programs. We tested several combinations to identify conditions efficient enough to open larvae and detach discs without damaging these organs. NT : not tested, ++ : successful and optimal try, + : successful but not optimal try, - : unsuccessful try

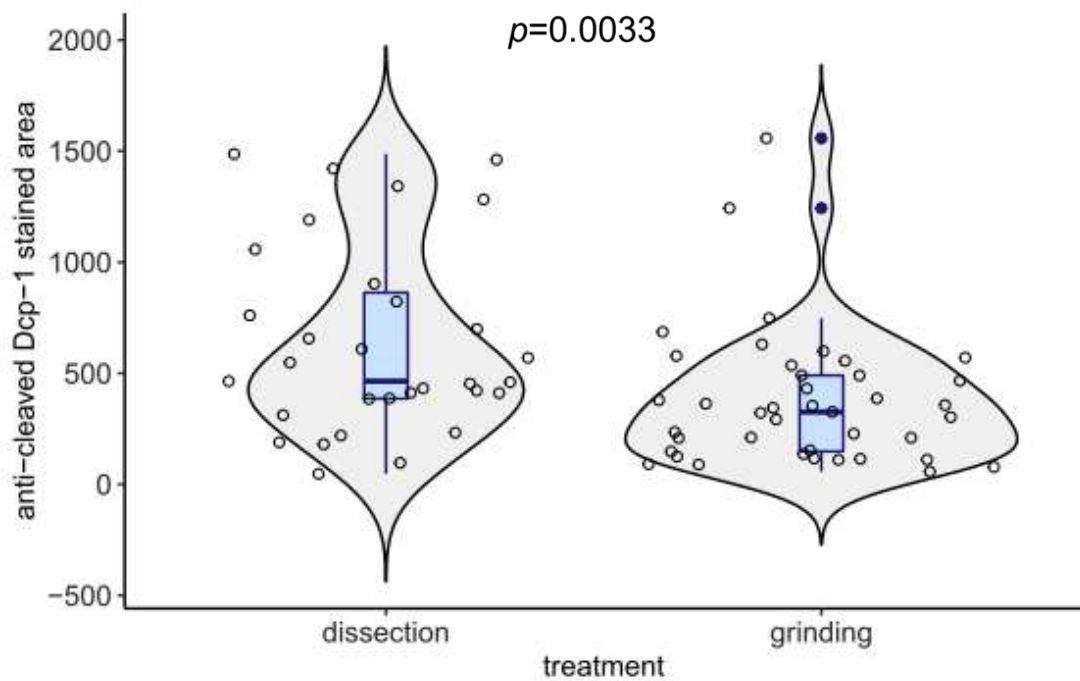




**Supplementary Figure S2: Example of a Ficoll gradient with  $\beta$ -galactosidase-stained wing imaginal discs.** Ficoll gradient after migration with  $\beta$ -galactosidase-stained wing imaginal discs. Inserts show magnification of the 10-15 and 15-20 interphases, blue arrows point to salivary glands (SG), and black arrows point to mouth hooks (MH). Smaller blue structures are wing imaginal discs (WID).



**C** Anti-cleaved Dcp-1 stained area depending on the treatment



**Supplementary Figure S3: Apoptosis staining in the wing discs according to the disc enrichment method.**

Protocol	Buffer	Grinding method	Filtration	Ficoll gradient	Yield	Purity	Quality
Fristrom and Mitchell, 1965	Ringer	Meat grinder	1.43 mm, 0.5 mm, n°5 silk sifting cloth	1st gradient 14:19%, 3 min 450 xg (estimated speed*) 2 <sup>nd</sup> 14% layer on top of linear gradient (14 to 24)	10 to 20%	95% containing leg, wing and eye-antenna imaginal discs	Transplantable
Zweidler and Cohen, 1971	Organ medium (5 mm disodium glycerophosphate, 10 mm KH <sub>2</sub> PO <sub>4</sub> , 30 mm KCl, 10 MM MgCl <sub>2</sub> , 3 rum CaCl <sub>2</sub> , 162 RIM sucrose pH 6.8)	Hand made automatic dissection device	Nylon net	Linear gradient 2 to 5%, no centrifugation but homemade sedimentation device	Unclear (« From 500 g of larvae, about 0.4 g of imaginal discs are obtained. »)	Unclear (« Contamination, as estimated by microscope observation, is about 20% by mass »)	Unclear, mentions that the technique has been used to study nuclear proteins and nuclear acids
Marty <i>et al.</i> , 2014	PBS	GentleMACS™ program m_liver_01_01	500 µm	16:25, 1h 4332 xg (estimated speed*)	1000 to 1500 wing discs for 5 mL of larvae. According to our calculations : 42 to 62%	Low after Ficoll gradient (enriched in wing imaginal discs) but sorted with Biosorter afterwards	ChIP-on-chip done (see Schertel <i>et al.</i> , 2015)
Hoareau <i>et al.</i> , 2022	Ringer	GentleMACS™ program m_brain_02_01	500/300/200 µm + 100 µm	10:15:20:25%, 20 min 80 xg	11.5%	50% directly after gradient due to salivary glands, but easy removal	Normal GFP expression and death rate

**Supplementary Figure S4:** Comparison of wing imaginal discs isolation protocols.

Protocol	Pros	Cons
Fristrom and Mitchell, 1965	Very pure Different kind of imaginal discs Good quality	Different kind of imaginal discs Meat grinder not available 2 Ficoll gradient including one continuous
Zweidler and Cohen, 1971	No crushing (automated dissection-like device with needles)	Lots of homemade devices difficult to reproduce Unclear on yield and purity
Marty <i>et al.</i> , 2014	High amount of discs Minimalist protocol	Use of PBS Need of a BioSorter
Hoareau <i>et al.</i> , 2022	Pretty high amount of discs Possibility to treat simultaneous samples Pure cells after dissociation and filtration Good quality	Yield 50 % purity right after gradient

**Supplementary Figure S5:** Summary table of pros and cons of available imaginal discs isolation protocols.

## II- Outils de validation des cibles identifiées

A la suite de la mise au point du protocole d'isolement en masse de disques imaginaux, il a fallu considérer la validation des cibles que nous espérons identifier en ChIP-seq et RNA-seq. En effet, si les analyses bio-informatiques permettent d'identifier des gènes régulés, ces régulations doivent être validées. Ce type d'analyse génère souvent des listes de plusieurs centaines de gènes, nous souhaitons réduire les cibles d'intérêt à celles correspondant aux problématiques du laboratoire, et notamment l'apoptose. Dans notre cas, nous souhaitons déterminer si les gènes identifiés comme cibles directes de Rbfl dans le disque imaginal d'aile et sélectionnés sur la base de leur fonction sont i) effectivement des cibles de Rbfl ii) capables de moduler la voie d'apoptose induite par Rbfl. Pour le premier point, des expériences de RT-PCR quantitative (RT-qPCR) ont été envisagées. Pour le second, nous avons besoin de tests permettant d'évaluer l'effet de ces gènes sur l'activité de Rbfl dans chacun des tissus utilisés : un premier permettant de déterminer si la cible est impliquée dans l'activité pro-apoptotique de Rbfl dans le disque d'aile, et un autre pour évaluer son éventuelle implication dans l'activité anti-apoptotique de Rbfl dans le disque d'œil. La mise en place de ces deux outils font l'objet de ce chapitre.

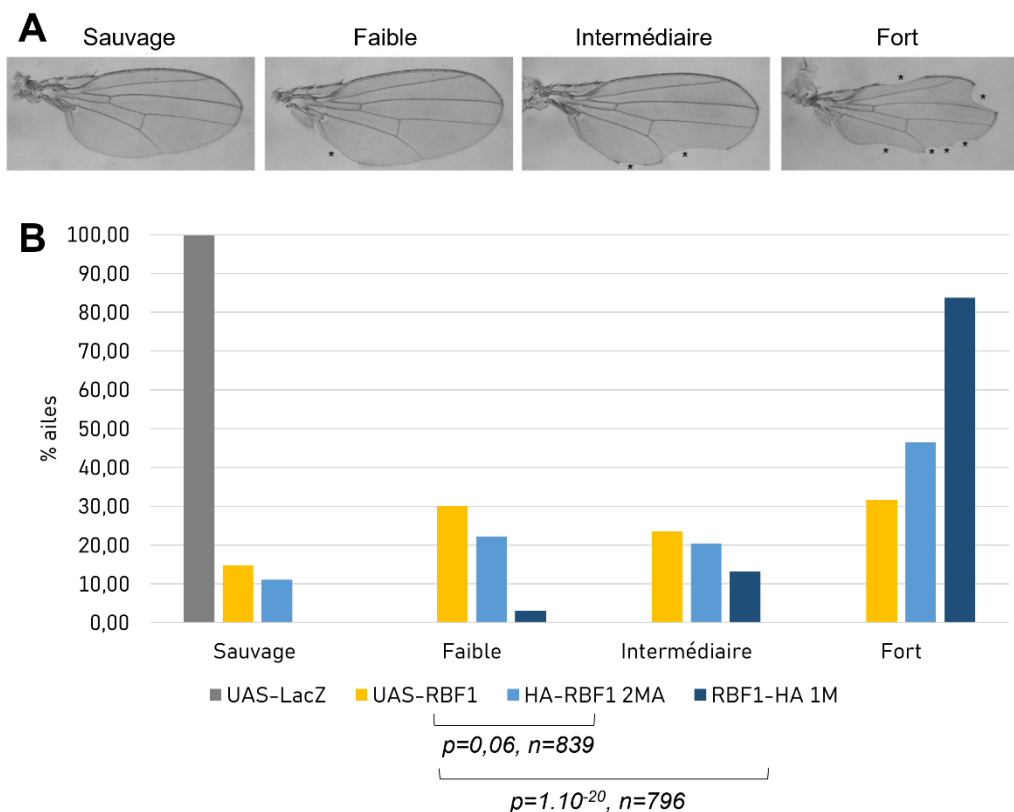
### 1. Activité pro-apoptotique dans l'aile

Le laboratoire travaillant sur l'apoptose induite par Rbfl, un test permettant d'évaluer si un acteur intervient directement ou indirectement dans le processus d'apoptose de Rbfl était déjà disponible et utilisé en routine. Celui-ci repose sur la capacité de Rbfl à induire l'apparition d'encoches à la marge de l'aile lorsqu'il est surexprimé sous contrôle du pilote *vestigial-Gal4* (*vg-Gal4*) [158]. Ces encoches sont une conséquence de l'apoptose induite au cours du développement dans le domaine *vestigial* du disque imaginal d'aile, visible par marquage TUNEL. Le nombre d'encoches présentes dans l'aile adulte est proportionnel à la quantité d'apoptose induite : plus il y a d'apoptose, plus le nombre d'encoches est grand. Ainsi, une échelle de force des phénotypes d'aile a été définie pour décrire la quantité d'apoptose (Figure 24A).

L'apoptose induite par Rbfl présente un profil caractéristique de répartition des phénotypes d'aile à travers les différents niveaux de sévérité (Figure 24B, jaune). En utilisant une surexpression de *rbfl* dans des contextes de déplétion (mutant, ARN interférent) ou au contraire de surexpression d'un gène d'intérêt impliqué dans la voie d'apoptose médiée par Rbfl, une modification de ce profil caractéristique est observée. Ainsi, l'inhibition d'une protéine X nécessaire à l'apoptose entraînera la diminution de celle-ci, ce qui se traduira par des phénotypes moins sévères et un décalage du profil vers la gauche. A l'inverse, une co-surexpression de Rbfl et de la protéine X induira une augmentation de l'apoptose et des phénotypes sévères, et donc un décalage du profil vers la droite. La surexpression ou la déplétion d'une protéine Y non impliquée dans l'apoptose médiée par Rbfl ne devrait avoir

aucun impact sur le profil de répartition des phénotypes d'aile. Ainsi, nous souhaitons soumettre des lignées mutantes ou des ARN interférents contre les cibles identifiées à ce test afin de déterminer leur implication dans l'apoptose induite par *Rbf1*.

Dans la pratique, l'utilisation de ce test repose donc sur le comptage des encoches à la marge des ailes d'individus issus de différents croisements réalisés puis repiqués le même jour, avec le même nombre de mâles et de femelles vierges. Ces comptages ont lieu à des jours consécutifs et à des heures fixes, jusqu'à épuisement des nouveaux adultes. De préférence, les ailes d'au moins une centaine d'individus sont comptées, en séparant mâles et femelles.



**Figure 24 : Effet de l'étiquette HA sur la distribution des phénotypes d'ailes encochées.** (A) Classement des phénotypes selon quatre catégories (sauvage, faible, intermédiaire, fort) selon le nombre d'encoches, signalées par des astérisques, présentes à la marge de l'aile. Adapté de Clavier *et al.*, 2014. (B) Répartition des phénotypes d'ailes lors de la surexpression de *rbf1*, étiqueté (nuances de bleu) ou non (jaune), par le pilote *vestigial-Gal4*. On observe un décalage de la répartition des phénotypes vers la droite correspondant à une aggravation significative des phénotypes induits par les *Rbf1* étiquetés en C-terminal (bleu foncé) par rapport au *Rbf1* de référence non étiqueté (jaune). Test de Wilcoxon.

J'ai également utilisé ce test pour caractériser l'activité pro-apoptotique de différentes souches permettant la surexpression de *Rbf1* étiqueté. Le laboratoire dispose de lignées transgéniques permettant l'expression de *Rbf1* étiqueté HA en N-terminal ou en C-terminal. Ces lignées ayant été générées par insertion aléatoire de *P[UAS-Rbf1-HA]* ou *P[UAS-HA-Rbf1]*, nous ne connaissons pas leur site d'insertion et donc leur niveau d'expression. Comme nous souhaitons utiliser cette étiquette HA dans les expériences de ChIP, nous souhaitons nous assurer que sa position ne perturbe pas les

interactions protéine-protéine indispensables à sa fonction de co-facteur de transcription. Pour cela, nous avons caractérisé la force des phénotypes d'ailes induit par ces souches afin d'en sélectionner une ayant le profil le plus proche possible de celui induit par Rbf1 non étiqueté. Parmi les cinq lignées disponibles au laboratoire et que nous avons sélectionnées, trois ont été éliminées car induisant des phénotypes d'ailes sauvages ou très faibles. Les deux restantes, Rbf1-HA 1M et HA-Rbf1 2MA $\alpha$ , induisent des phénotypes d'ailes encochées. Avec la souche Rbf1-HA 1M, le profil est très décalé vers la droite avec une forte majorité de phénotypes forts ; tandis qu'avec la HA-Rbf1 2MA $\alpha$ , le profil est similaire au Rbf1 non étiqueté, avec toutefois des phénotypes légèrement plus forts (Figure 24B). C'est cette dernière souche que nous avons choisi d'utiliser pour les expériences d'immunoprécipitation.

### 2. Activité anti-apoptotique : de l'œil aux neurones

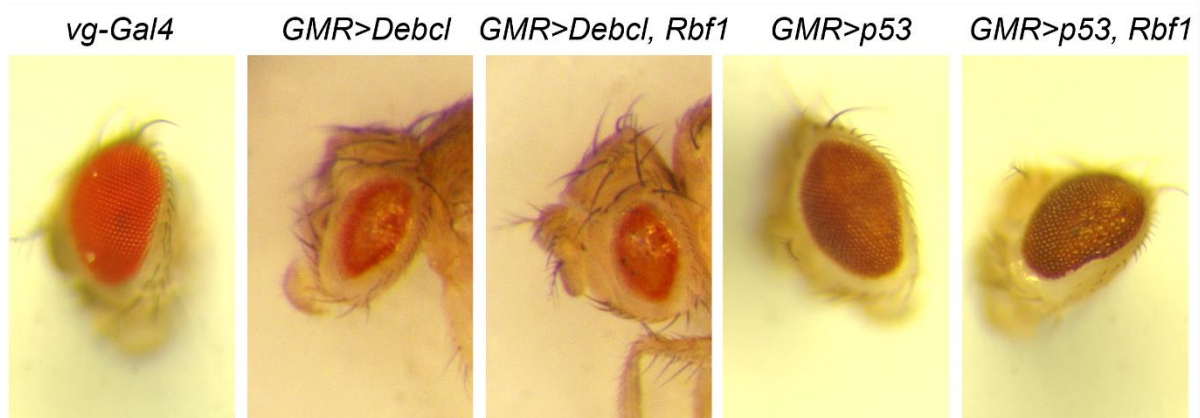
L'activité anti-apoptotique de Rbf1 dans le disque imaginal d'œil, bien que découverte par le laboratoire, est bien moins étudiée que son activité pro-apoptotique. Afin de pouvoir déterminer si les cibles identifiées sont impliquées dans cette activité, nous avons souhaité mettre en place un test phénotypique simple, semblable à celui utilisé dans l'aile. Si les phénotypes d'encoches à la marge de l'aile sont faciles à observer et à quantifier, les phénotypes d'œil de drosophile sont bien plus compliqués à évaluer. En effet plusieurs paramètres entrent en jeu : taille, rugosité, apparence des ommatidies et des soies ou encore couleur. La mise au point d'un outil de criblage de la suppression de l'apoptose induite par un pro-apoptotique tel que p53, facile d'utilisation a donc été l'objet d'une partie de mon projet de thèse.

#### a. L'œil, un modèle difficilement utilisable

Nous sommes dans un premier temps parties des résultats décrits dans l'article Milet *et al.* paru en 2010 [158]. Ceux-ci démontrent l'activité anti-apoptotique de Rbf1. Pour rappel, la surexpression de *rbf1* dans le disque imaginal d'œil n'induit pas d'apoptose. Au contraire, lorsque Rbf1 et Dmp53, l'homologue du pro-apoptotique p53, sont co-surexprimés dans le disque imaginal d'œil ou dans les neurones, une réduction de l'œil adulte moins importante que celle induite par la surexpression de Dmp53 seul est observée. Dans cet article, deux pilotes ainsi que deux protéines pro-apoptotiques sont testées. D'une part, le pilote *elav-Gal4*, qui permet une expression pan-neuronale, est utilisé pour diriger l'expression de *Dmp53*. D'autre part, *Debcl*, membre pro-apoptotique de la famille Bcl-2, est exprimé sous contrôle du pilote *GMR-Gal4*, dans la région postérieure au sillon morphogénétique du disque d'œil.

Pour commencer, notre intérêt se portant spécifiquement sur le disque imaginal d'œil, nous avons utilisé le pilote *GMR-Gal4*. En effet, le pilote *elav-Gal4* permettant une expression dans les neurones, il n'est pas spécifique de notre tissu d'intérêt. Nous avons donc testé l'effet de l'induction de deux pro-

apoptotiques dans ce tissu : *p53* et *Debcl*. Pour le premier, une seule souche était disponible au laboratoire. Pour le second gène, trois souches étaient à notre disposition : une première portant le transgène *UAS-Debcl* sur le chromosome 2, surnommée « D2 » ; une deuxième portant le même transgène sur le chromosome 3, surnommée « D3 » ; et une troisième contenant ces deux inserts, surnommée « D2 D3 ». Le choix de la souche à utiliser s'est fait selon deux paramètres : l'intensité des phénotypes induits par ces souches et leur facilité d'utilisation. En effet, à cette étape, nous souhaitons juste évaluer s'il était possible d'observer des phénotypes liés à l'apoptose ainsi qu'une suppression au moins partielle du phénotype associé. Nous avons donc besoin uniquement de phénotypes visibles et non très forts. L'expérience du laboratoire ainsi que des tests rapides nous ont permis d'éliminer la souche dite « D2 » qui n'induit que des phénotypes d'apoptose très faibles dont la suppression serait difficilement détectable. La souche dite « D2 D3 » possédant deux inserts sur deux chromosomes différents, est plus difficile à utiliser car il faut pouvoir suivre leur co-ségrégation dans les descendants d'intérêt qui sont de ce fait en nombre plus restreint. Nous avons donc choisi d'utiliser la souche permettant la surexpression de *Debcl* « D3 ».



**Figure 25 : Effet sur la morphologie de l'œil de la surexpression de gènes pro-apoptotiques avec ou sans co-surexpression de *Rbf1* par le pilote *GMR*.** La surexpression des gènes pro-apoptotiques *p53* et *Debcl* induisent de l'apoptose dans le domaine GMR du disque d'œil, qui se traduit par un phénotype d'œil à l'âge adulte, manifesté par une réduction de la taille de l'œil ainsi qu'une dépigmentation et un aspect rugueux au centre de l'œil par rapport à un œil sauvage (*vg-Gal4*). La co-surexpression de *Rbf1* (*GMR > p53, Rbf1* et *GMR > Debcl, Rbf1*) ne semble pas supprimer ce phénotype. Grossissement x16.

La surexpression de *Debcl* dans le domaine GMR induit une réduction de la taille de l'œil adulte, qui apparaît moins large qu'un œil d'adulte sauvage. On observe également une dépigmentation, surtout au milieu, ainsi qu'un aspect rugueux. Les ommatidies semblent avoir perdu leur organisation. Ces phénotypes sont plus légers lors de la surexpression de *Dmp53* : l'œil est moins réduit et la désorganisation et dépigmentation ne sont que faibles et visibles sur une bande fine au milieu de l'œil (Figure 25). Pour les deux génotypes, peu d'individus ont pu être observés et les phénotypes peuvent varier d'un individu à l'autre.

Lorsqu'on co-sureprime *Rbf1* avec ces pro-apoptotiques, les phénotypes restent très semblables. Il semblerait que la zone désorganisée soit plus limitée en co-surexpression de *Rbf1* et *Debcl* qu'avec



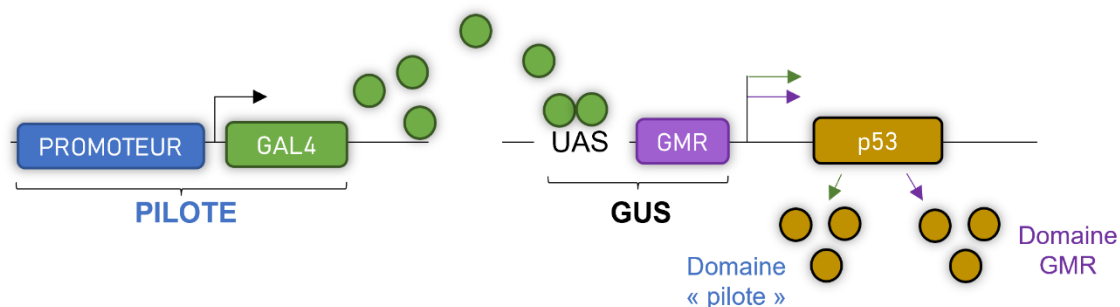
Debel seul, mais cette observation n'est valable que pour quelques individus et ne peut être faite avec certitude (Figure 25). Il n'est donc pas possible de mettre en évidence de suppression, même partielle, du phénotype par *Rbf1*. Il semble donc que l'effet anti-apoptotique de *Rbf1* est trop limité dans ces conditions pour générer une modification visible du phénotype. Comme évoqué en début de chapitre, la caractérisation des phénotypes d'œil est difficile car elle nécessite la prise en compte de plusieurs paramètres, eux-mêmes parfois complexes à observer. Dans notre cas, la description de la morphologie de l'œil nécessite la prise en compte de la taille, de la forme, de la rugosité et de l'organisation apparente des ommatidies. De plus, la surexpression d'un gène pro-apoptotique, même dans une zone limitée comme l'œil, est généralement associée à de la létalité. Cela limite les effectifs disponibles pour l'étude, sachant que la variabilité des phénotypes nécessiterait au contraire des effectifs importants pour que l'analyse soit robuste.

Le pilote *GMR-Gal4* ne semble donc pas être approprié pour notre objectif, qui est de trouver un test de suppression phénotypique facilement utilisable pour un crible génétique. Ici, les phénotypes, par leur multiplicité et leur variabilité, sont difficiles à caractériser précisément, et l'effet de la surexpression de *rbf1* semble trop faible pour observer une différence en contexte de surexpression d'un gène pro-apoptotique seul ou de co-surexpression avec *rbf1*.

L'étape suivante a été de tester si le pilote *elav-Gal4*, qui permet une surexpression dans tous les neurones, était plus adapté à notre crible. Malheureusement, nous n'avons observé aucun phénotype sur l'œil adulte en contexte de surexpression de *p53* par rapport à des individus contrôles (données non présentées). Devant ces résultats surprenants, nous avons souhaité vérifier si notre souche *elav-Gal4* permettait bien une surexpression dans les cellules attendues grâce à la surexpression de la GFP. Nous avons observé une fluorescence dans les cellules correspondantes. La souche *p53*, quant à elle, avait déjà prouvé sa capacité à être responsable de mort cellulaire lors des expériences avec le pilote *GMR-Gal4*. Les deux souches utilisées semblent donc bien posséder les transgènes d'intérêt, mais l'apoptose induite dans les neurones par la surexpression de *p53* avec le pilote *elav* ne résulte pas en des phénotypes visibles chez l'adulte.

### b. Autre souche, autre phénotype

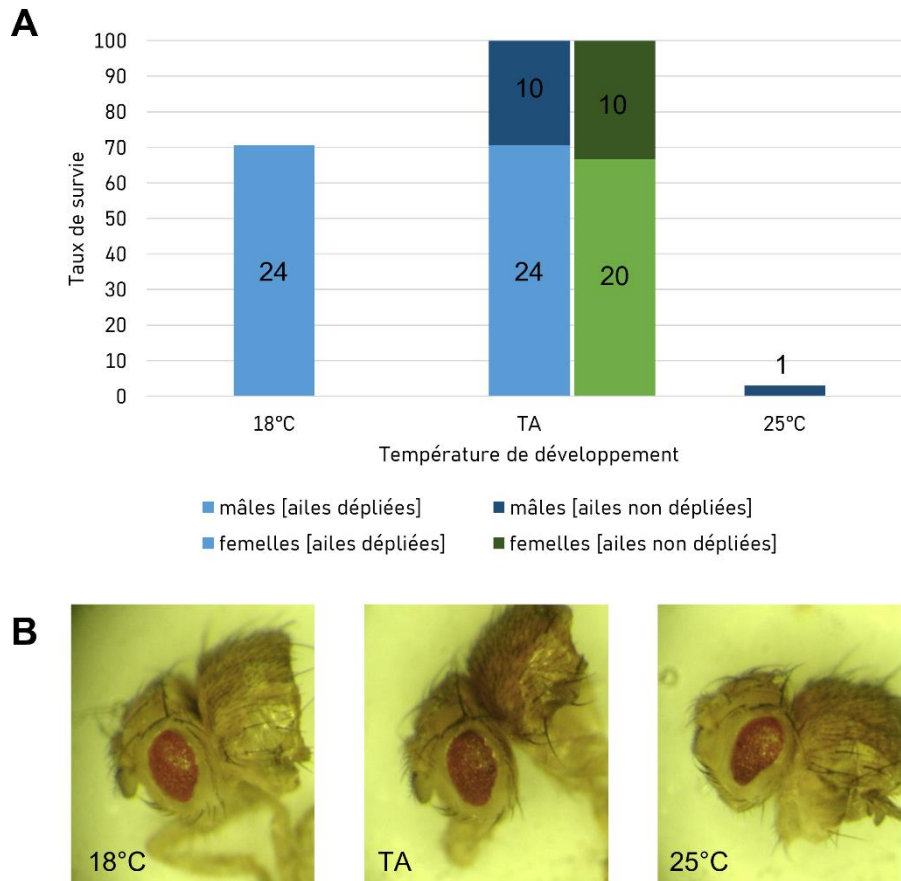
La recherche d'autres souches permettant l'expression dans des domaines équivalents à ceux que nous souhaitions étudier a mené à l'étude d'une construction particulière : GUS-*p53*. Cette construction place le gène *p53* sous contrôle de deux séquences régulatrices : le promoteur *GMR* et une séquence *UAS*. Ainsi, *p53* sera exprimé à la fois dans le domaine *GMR*, à savoir en postérieur du sillon morphogénétique du disque imaginal d'œil ; et dans le domaine d'expression d'un pilote permettant la synthèse de Gal4. Dans notre cas, en présence d'un pilote *elav-Gal4*, cette souche permettait la surexpression du pro-apoptotique *p53* à la fois dans les domaines *GMR* et *elav* (Figure 26).



**Figure 26 : Schéma de la construction GUS-p53.** Le système GUS-p53 permet l'expression de *p53* sous contrôle de séquences *UAS* et du promoteur *GMR*. Par utilisation du système UAS-Gal4, *p53* peut être induit à la fois dans le domaine *GMR* et le domaine du pilote utilisé pour activer les séquences *UAS*.

Nous avons donc testé si une construction *elav>GUS-p53* permettait l'obtention de phénotypes d'œil. Nous avons pu observer d'une part que la surexpression de *p53* dans ces deux domaines entraînait un phénotype d'œil réduit, caractérisé par un œil moins large et rugueux. D'autre part, la surexpression de *Rbf1* seul dans le domaine *elav* n'induit pas de phénotype d'œil, qui a une apparence sauvage. De manière intéressante, en plus de ce phénotype d'œil réduit, nous avons remarqué de nombreuses pupes non écloses et peu d'individus adultes du génotype d'intérêt lors de la surexpression de *p53*, comme cela avait été observé précédemment dans l'article Milet *et al.* Le pilote *elav-Gal4* étant pan-neuronal, la surexpression d'un pro-apoptotique dans ce domaine impacte tous les neurones, y compris les moteurs. Notre hypothèse était donc que les adultes formés après la métamorphose dans la pupa ne sont pas capables de s'en extirper. Cette hypothèse était appuyée par le fait que les quelques individus naissants ne semblaient pas capables de déplier leurs ailes, ce qui dépend de l'action de muscles thoraciques. Dans ce phénotype particulier, nous avons vu l'opportunité d'un nouveau test facilement utilisable pour un crible : le taux d'émergence des adultes.

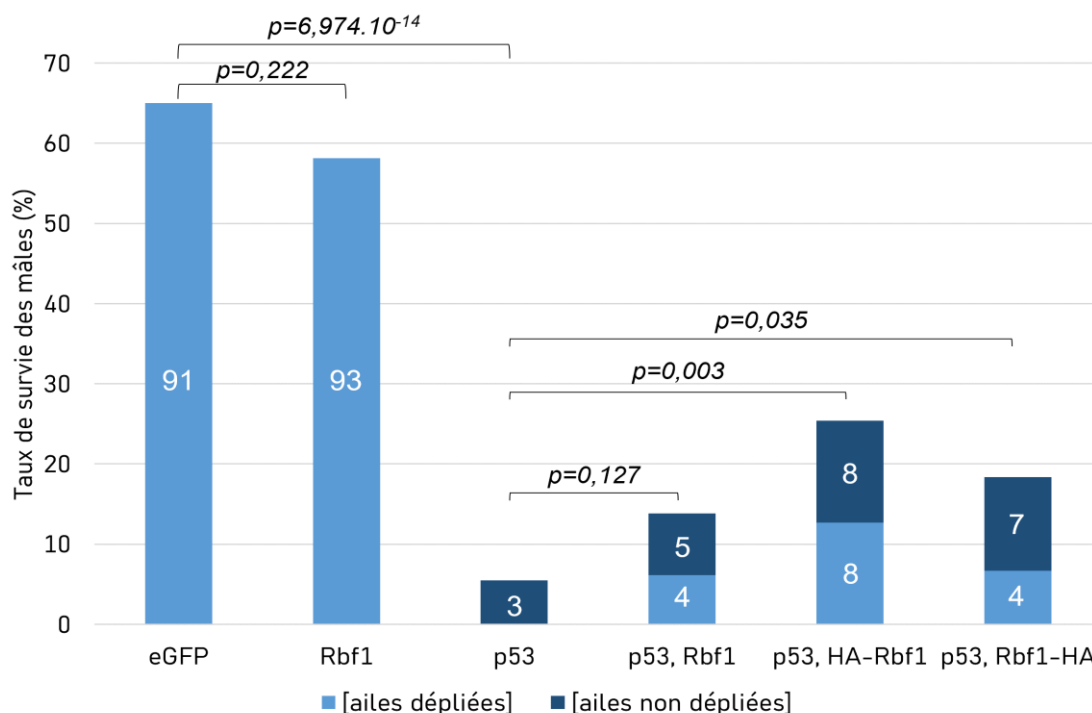
Afin de confirmer cette observation et de caractériser davantage ce défaut de naissances, nous avons croisé des femelles vierges *elav-Gal4* avec des mâles GUS-p53. Chez la drosophile, la température d'incubation va modifier la durée du cycle de développement. Dans le système UAS-Gal4, cette température a également un effet sur l'expression des gènes sous contrôle des séquences UAS, qui va augmenter avec la température. Ainsi, il y aura une expression de gènes plus élevée et des phénotypes associés plus forts à 25°C ou 29°C qu'à 18°C [193]. Dans l'optique de caractériser ce phénotype de létalité, les croisements ont été incubés à trois températures différentes : 18°C, température ambiante (~20°C) et 25°C.



**Figure 27 : L'expression de p53 dans les domaines *elav* et *GMR* simultanément entraîne une létalité pupale et des phénotypes d'œil chez les individus nés.** (A) Taux de survie rapportés à la survie des individus élevés à température ambiante (TA). Les chiffres noirs sur l'histogramme indiquent l'effectif de chaque catégorie. Les individus *elav>GUS-p53* présentent une létalité pupale dépendante de la température. C'est à température ambiante (TA) qu'on trouve le plus de survivants, mais environ un tiers d'entre eux présentent un phénotype d'ailes non dépliées. A 25°C, le seul individu naissant présente ce phénotype, contre aucun des 24 individus à 18°C. (B) Les yeux des adultes naissants sont réduits et rugueux, il est difficile de déterminer une différence de force de phénotype aux trois températures.

Nous avons observé un taux de survie par rapport aux individus élevés à température ambiante plus important à 18°C qu'à 25°C où un seul individu a émergé. De manière surprenante, la température ambiante est celle permettant le plus de naissances (64 individus) (Figure 27A). Pour comparaison, un croisement de femelles vierges *elav-Gal4* avec des mâles *UAS-eGFP* donne classiquement 120 à 150 individus à 25°C. La létalité est bien plus prononcée chez les femelles puisque nous n'avons pas pu en obtenir que lorsque le développement s'est fait à température ambiante. Parmi ces individus nés, nous avons également observé un phénotype d'aile non dépliée, probablement dû à l'apoptose induite dans les neurones. Ce phénotype a été observé pour 1/3 des individus à température ambiante et pour le seul adulte né à 25°C, mais pas chez ceux incubés à 18°C (Figure 27A). Ce phénotype d'aile non dépliée semble donc dépendre de la température de

développement. Quant aux phénotypes d'œil, il semblerait que l'œil soit moins rugueux et plus rond à des températures plus faibles (Figure 27B).



**Figure 28 :** La co-surexpression de *Rbf1* étiqueté supprime partiellement la létalité pupale des mâles *elav>GUS-p53*. Taux de survie par rapport à la survie des individus exprimant le balancier SM5, TM6B. Les chiffres blancs sur l'histogramme indiquent l'effectif de chaque catégorie. On observe un taux de survie très inférieur des mâles *elav>GUS-p53* (« p53 ») par rapport aux mâles *elav>eGFP* ou *elav>Rbf1*. Une suppression partielle significative de la létalité est observée lorsqu'on co-surexprime *GUS-p53* et la protéine *Rbf1* étiquetée sous contrôle du pilote *elav-GAL4*. HA-Rbf1 : HA-Rbf1 2MA $\alpha$ , Rbf1-HA : Rbf1-HA 1M. Test de Chi-2 d'homogénéité sur les effectifs totaux ([ailes dépliées] + [ailes non dépliées]),  $\alpha=0,05$ .

Pour la suite de ces expériences, nous avons voulu déterminer si la co-surexpression de *Rbf1* sous contrôle du pilote *elav-Gal4* permettait une suppression de la létalité pupale. Pour cela, nous avons placé les croisements réalisés à 25°C, température la plus délétère à l'éclosion des pupes. Compte tenu de l'absence de suppression observée dans les expériences avec le pilote *GMR-Gal4* (Figure 25), nous avons également co-surexprimé les deux *Rbf1* étiquetés. Ceux-ci présentaient des phénotypes d'ailes plus forts (Figure 24), ce qui pourrait témoigner d'un niveau d'expression plus élevé que la souche ne portant pas d'étiquette HA. Dans ces conditions, on observe une suppression significative de la létalité chez les mâles lorsque *p53* est co-surexprimé avec un transgène portant une forme étiquetée de *Rbf1*. En revanche, cette suppression de la létalité n'est pas significative lorsque le transgène *Rbf1* non étiqueté est utilisé. Cependant, on retrouve des mâles co-surexprimant *p53* et *rbf1* un phénotype d'aile dépliée, contrairement à la surexpression de *p53* seul pour laquelle la totalité des individus naissant présentent des ailes non dépliées. La co-surexpression de *rbf1* semble donc permettre une suppression partielle de la létalité et du phénotype d'aile non dépliée chez les mâles, mais le faible effectif obtenu

en condition de surexpression de *p53* (3 individus) ne nous permet pas d'évaluer avec certitude la répartition des phénotypes d'aile dans cette condition. Chez les femelles, aucune naissance n'a pu être observée pour les surexpressions de *p53*, avec ou sans co-surexpression de *rbf1* (Figure 28). Les phénotypes sont donc encore une fois plus forts chez les femelles et aucune suppression de la létalité n'a pu être observée.

Nous avons donc pu mettre au point un test de suppression phénotypique permettant d'observer facilement l'activité anti-apoptotique de Rbf1. Exprimer les transgènes *p53* et HA-Rbf1-2MA $\alpha$ , dont l'effet semble plus fort que Rbf1 et Rbf1-HA 1M (meilleure suppression des phénotypes de létalité et de non-déploiement des ailes), dans des contextes de déplétion de gènes cibles identifiés permettrait d'évaluer leur implication dans le rôle anti-apoptotique de Rbf1.

### III- Le ChIP-seq : une méthode d'identification des promoteurs liés par Rbf1

La présence d'une protéine au promoteur d'un gène ne renseigne pas l'activation ou la répression transcriptionnelle induite, ou même sur la présence d'une quelconque régulation de la transcription. A l'inverse, étudier le transcriptome seul ne permet pas de différencier les régulations directes et indirectes. Pour identifier les cibles directes de Rbf1, notre approche combine ces deux techniques complémentaires : une permettant d'identifier les promoteurs fixés par Rbf1, le ChIP-seq ; et une seconde permettant de déterminer le sens de régulation de la transcription de ces gènes, le RNA-seq. Dans ce chapitre, nous allons nous intéresser à la mise en place et à l'analyse de ce ChIP-seq.

Si le projet initial prévoyait la réalisation du ChIP-seq et du RNA-seq dans les disques d'aile et d'œil, la mise au point du protocole d'isolement en masse de disques imaginaux d'aile a été plus longue que prévue et bien que les protocoles historiques suggèrent que l'obtention des disques d'œil par la même approche soit assez aisée, il s'avère dans les faits que cela requiert des mises au point supplémentaires incompatibles avec les délais à respecter pour une thèse. Nous avons donc choisi de nous concentrer uniquement sur le disque imaginal d'aile pour les approches de séquençage. Nous aurions souhaité, une fois des cibles directes identifiées dans le disque d'aile, observer la régulation de ces cibles par Rbf1 dans le disque imaginal d'œil puis finalement, comme développé dans la section précédente, dans les neurones.

#### 1. Mise au point de l'immunoprécipitation de chromatine

La technique de ChIP-seq repose sur une première étape de ChIP, immunoprécipitation de chromatine. Dans cette approche, une première étape de crosslink permet de fixer les protéines à l'ADN par des liaisons covalentes. L'ADN est ensuite fragmenté par sonication et le facteur d'intérêt, ici Rbf1, est immunoprécipité avec le fragment d'ADN et éventuellement les protéines associées. La

dernière étape consiste à supprimer le pontage ADN-protéines, dégrader les protéines et finalement purifier les fragments d'ADN ainsi isolés pour les séquencer massivement.

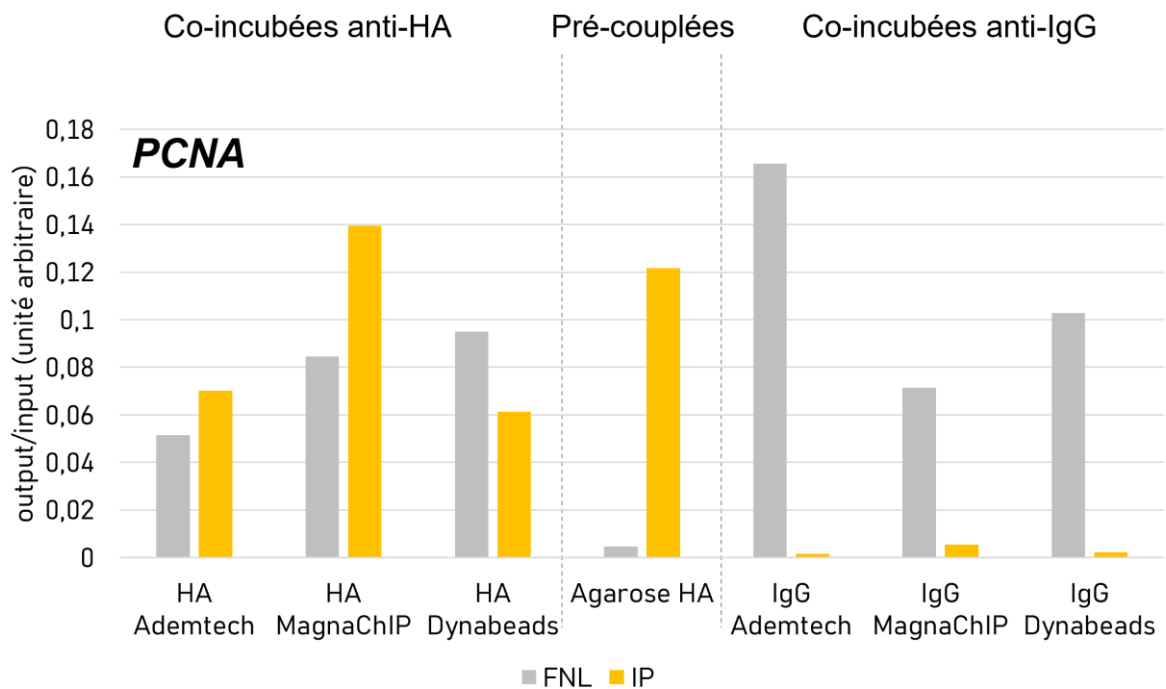
Parmi toutes ces étapes, il était nécessaire d'en adapter certaines à notre projet. Par chance, des expériences d'immunoprécipitation de chromatine avaient déjà été réalisées par d'anciens membres du laboratoire. Un protocole était donc pré-existant mais il a été nécessaire d'en vérifier les étapes et de les ajuster aux conditions de notre expérience. Certaines, en particulier les paramètres d'immunoprécipitation (étiquette, souche, billes, ...), ont nécessité d'importants ajustements, qui sont présentés ci-après.

#### a. Anticorps : viser la protéine ou une étiquette ?

Pour immunoprécipiter une protéine, deux possibilités sont envisageables : utiliser un anticorps visant directement la protéine d'intérêt, ou un anticorps dirigé contre une étiquette portée par la protéine d'intérêt. Au laboratoire, l'anticorps anti-Rbfl disponible n'est pas un anticorps commercial, sa quantité est donc limitée. Or la ChIP est une expérience qui utilise de grandes quantités d'anticorps. Il nous a donc semblé plus avisé d'utiliser un anticorps commercial, sachant que nous disposions par ailleurs au laboratoire de versions étiquetées de notre protéine. Parmi les souches disponibles pour ce faire, le choix de la souche à utiliser a été décrit dans la section II-1 de ce chapitre. Grâce à la surexpression sous contrôle du système UAS-Gal4 d'un transgène Rbfl-HA, nous nous assurons une immunoprécipitation uniquement dans les cellules d'intérêt, celles-ci étant les seules à posséder une forme étiquetée de la protéine Rbfl.

#### b. Quelles billes d'immunoprécipitation ?

Différentes expériences d'immunoprécipitation de chromatine ayant déjà été réalisées par d'anciens membres du laboratoire, nous possédions plusieurs types de billes et d'anticorps permettant l'immunoprécipitation de Rbfl. Principalement, il est possible de distinguer deux types de billes pour les immunoprécipitations : des billes permettant de lier l'anticorps indirectement, ou des billes portant directement l'anticorps lié covalamment. Nous avons à notre disposition trois types de billes pour les immunoprécipitations indirectes : Bio-Adembeads (Ademtech), MagnaChIP Protein G Magnetic Beads (Sigma) et Dynabeads Protein G (Thermofisher), toutes trois magnétiques et que je désignerai par la suite respectivement par les noms Adembeads, MagnaChIP et Dynabeads ; ainsi qu'un anticorps anti-HA ChIP grade déjà testé et validé au laboratoire. Nous possédions également une résine pré-couplée anti-HA du fabricant Sigma.



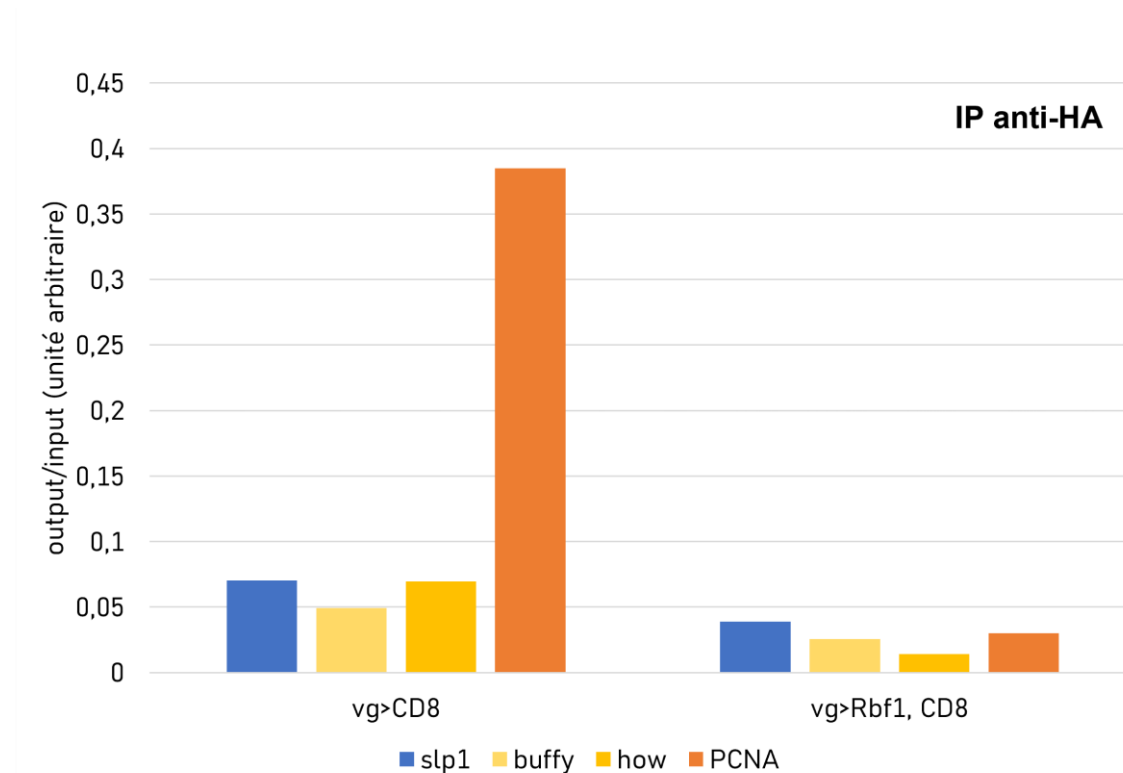
**Figure 29: ChIP-qPCR réalisées à partir d'expériences de ChIP faites avec différentes billes d'immunoprécipitation.** Les billes Ademtech, MagnaCHIP et Dynabeads ont été co-incubées avec un anticorps anti-HA (« HA ») ou avec un contrôle IgG (« IgG »). L'agarose HA est une résine pré-couplée avec des anticorps anti-HA. Les expériences de ChIP avec des quantités d'ADN initiales, identiques pour chaque condition, ont été dosées au Quantus à 0,388 ng/μL (soit 53 ng) et 0,310 ng/μL (soit 15,5 ng). Les qPCR ont été réalisées en duplicats ou triplicats techniques, la moyenne des deux expériences est considérée ici. Les données représentées correspondent au rapport entre la quantité d'ADN de *PCNA* obtenue dans les différentes conditions (output) et sa quantité initiale (input).

Afin de déterminer quelles billes seraient les plus efficaces pour nos expériences, nous souhaitons observer si celles-ci permettaient de récupérer des fragments d'ADN comprenant des promoteurs déjà caractérisés comme cibles de Rbf1. Nous avons alors réalisé des expériences de ChIP-qPCR avec des amorces permettant l'amplification de *PCNA*, une cible connue de Rbf1. Après ChIP avec des anticorps anti-HA ou anti-IgG co-incubés avec les différentes billes ou ChIP avec la résine pré-couplée anti-HA, des PCR quantitatives (qPCR) ont été faites sur les résultats d'immunoprécipitation (IP) ou sur le surnageant de ces IP, appelé « fraction non liée » (FNL). Nous pouvons observer que l'immunoprécipitation avec tous les types de billes a été efficace en comparaison avec une IP contrôle anti-IgG. Il semblerait que l'immunoprécipitation avec la résine pré-couplée HA ait été la plus probante puisqu'elle permet une déplétion plus importante en fragments contenant la séquence de *PCNA* de la fraction non liée et présente par ailleurs le meilleur enrichissement dans la fraction précipitée (IP) (Figure 29). Nous avons choisi d'utiliser ces billes pour la suite de nos expériences.

### c. Le début des ennuis : essais de ChIP contrôle VS Rbf1

Dans le but de confirmer si l'enrichissement que nous avons observé était bien spécifique d'une immunoprécipitation de Rbf1, ce qui semble cohérent en comparant les immunoprécipitations anti-HA

et anti-IgG (Figure 29), nous voulions réaliser des expériences de ChIP-qPCR en comparant les conditions contrôle envisagées (*vg>mCD8-GFP*) et les conditions test (*vg>Rbf1-HA 2MA $\alpha$ , mCD8-GFP*). Cette fois-ci, en plus de *PCNA*, nous avons choisi d'ajouter *buffy* et *how* car nous souhaitions également utiliser des cibles témoins de la régulation de l'apoptose par *Rbf1*.



**Figure 30 : ChIP-qPCR réalisée avec ou sans surexpression de *Rbf1*.** Les expériences de ChIP contrôle « *vg>CD8* » et en surexpression de *Rbf1* « *vg>Rbf1, CD8* » avec une quantité d'ADN initiale mesurée au Quantus pour l'échantillon « *vg>Rbf1, CD8* » est de 9,5 ng/ $\mu$ L, soit 950 ng. Les échantillons ont été déposés en duplicats, la moyenne des deux dépôts est considérée ici. *Buffy*, *how* et *PCNA* sont des gènes cibles connus de *Rbf1*, *slp1* n'est (a priori) pas une cible de *Rbf1*. Les données représentées correspondent au rapport entre la quantité d'ADN obtenue dans les différentes conditions (output) et sa quantité initiale (input).

Nous attendions un enrichissement des gènes cibles *PCNA*, *buffy* et *how* par rapport au contrôle *slp1* au sein des ChIP dans lesquelles la forme étiquetée de *Rbf1* était présente, et aucun enrichissement pour ces trois cibles dans les conditions contrôles *vg>CD8*. Or, nous n'avons observé aucun enrichissement particulier en conditions *Rbf1*. Nous pouvons également noter un enrichissement aberrant de *PCNA* dans les conditions *CD8* dont le rapport quantité après IP/quantité initiale (output/input) atteint presque 0,4, ce qui est beaucoup trop élevé pour une expérience de ChIP (Figure 30).

Plusieurs causes pouvaient être responsables de ces résultats non satisfaisants, parmi lesquelles :

- Les billes que nous avons sélectionnées ne sont finalement pas efficaces. La précipitation serait alors aspécifique et ce serait la raison de l'absence d'enrichissement des cibles connues. L'ADN récupéré ne serait alors pas spécifique de la fixation de *Rbf1*, mais simplement de l'ADN précipité par son poids ou par précipitation aspécifique des billes.

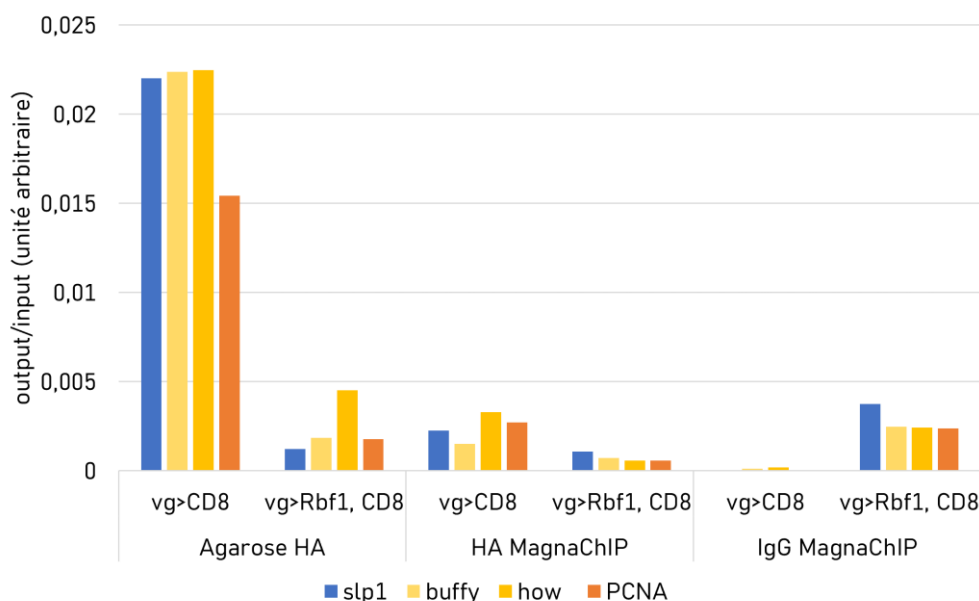


- Le contrôle CD8 contient une étiquette HA non signalée. Les billes d'immunoprécipitation pourraient alors précipiter CD8 plutôt que Rbf1. Tout comme dans l'hypothèse précédente, l'ADN récupéré ne serait pas spécifique de la fixation de Rbf1.

Bien qu'aucune de ces hypothèses n'explique l'enrichissement très important de *PCNA* sans surexpression de Rbf1, nous avons donc décidé de tester ces paramètres.

### Vérification de l'efficacité des billes d'immunoprécipitation

Dans un premier temps, nous avons choisi de vérifier l'efficacité des billes que nous avons sélectionnées. En effet, la situation sanitaire du début 2020 a fait que quelques mois se sont déroulés entre le choix des billes et ces essais de ChIP contrôle VS Rbf1. Durant ce délai, les billes auraient pu être contaminées ou accidentellement stockées de manière non optimale. De plus, la ré-examination des données de la **Figure 30** nous a interpellées sur les valeurs du ratio output/input. En effet, le rapport output/input de la fraction IP pour les billes pré-couplées HA est trop élevé pour des données de ChIP, dont le rendement est habituellement de moins de 10 %.



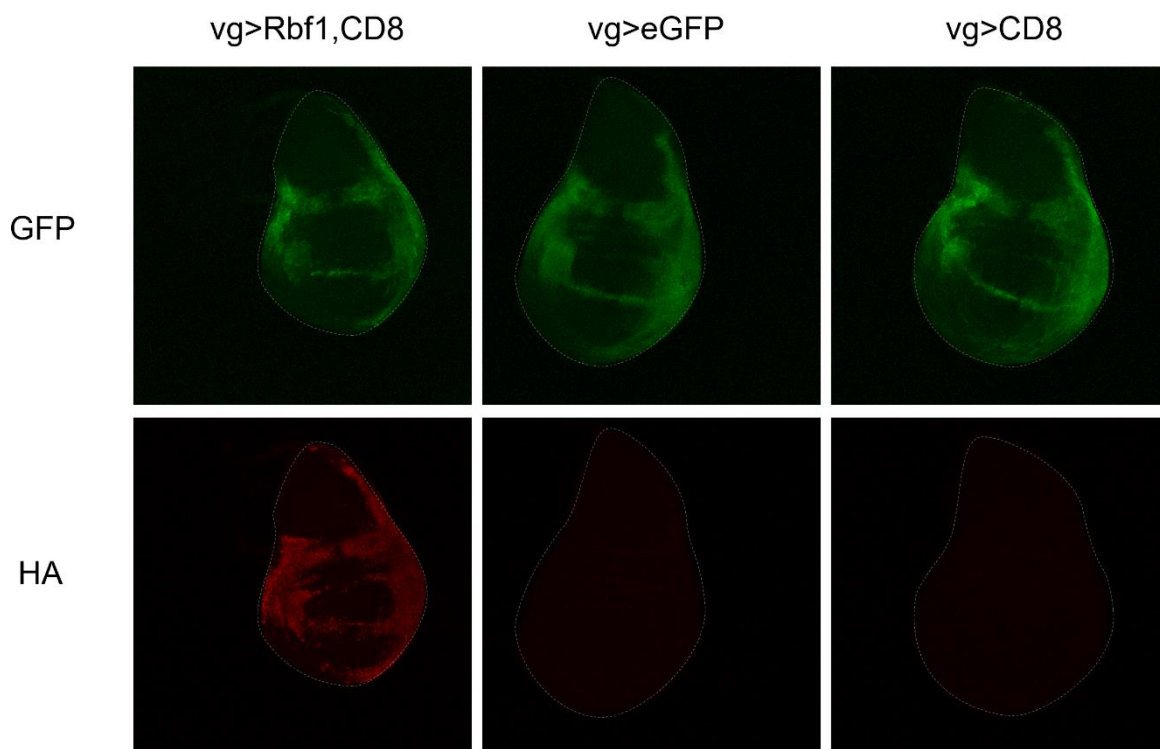
**Figure 31 : ChIP-qPCR réalisée en conditions contrôles ou en surexpression de Rbf1, avec différentes billes d'immunoprécipitation.** Les billes MagnaChIP ont été co-incubées avec un anticorps anti-HA (« HA MagnaChIP ») ou avec un contrôle IgG (« IgG MagnaChIP »). L'agarose HA est une résine pré-couplée avec des anticorps anti-HA. Les expériences de ChIP contrôle « vg>CD8 » et en surexpression de Rbf1 « vg>Rbf1, CD8 » ont été réalisées 100 disques imaginaires d'ailes. La quantité initiale n'a pas été dosée pour cette expérience. Les échantillons ont été déposés en duplicats, la moyenne des deux dépôts est considérée ici. Les données représentées correspondent au rapport entre la quantité d'ARN obtenue dans les différentes conditions (output) et sa quantité initiale (input).

Pour ce test, nous avons choisi d'utiliser des billes MagnaChIP : en effet, nous avons obtenu des résultats satisfaisants lors de tests précédents (**Figure 29**) et ces billes faisaient partie des plus

récemment utilisées au laboratoire pour des expériences de ChIP. Malheureusement, ni les billes à incuber avec les anticorps ni les billes pré-couplées n'ont permis d'obtenir un enrichissement en conditions Rbf1 par rapport aux conditions contrôles, ou même un enrichissement des cibles connues par rapport au contrôle *slp1* (Figure 31). Le choix des billes ne semblait donc pas être le responsable de l'absence d'enrichissement observée.

#### Vérification de l'absence d'une étiquette HA dans la souche contrôle

Enfin, nous avons voulu savoir si la souche CD8 utilisée comme contrôle ne possédait pas d'étiquette HA non signalée. La présence d'un HA dans cette souche aurait permis la précipitation dans notre contrôle également, qui aurait pu entraîner la précipitation de fragments d'ADN. Cette possibilité était toutefois la moins plausible de nos trois hypothèses, car CD8 est une protéine membranaire et non nucléaire qui n'est pas capable d'interagir avec l'ADN et, même si c'était le cas, il est peu probable que ces potentielles cibles ou interactions aléatoires correspondent aux cibles de Rbf1. Toutefois, il nous paraissait important d'écarter cette hypothèse.



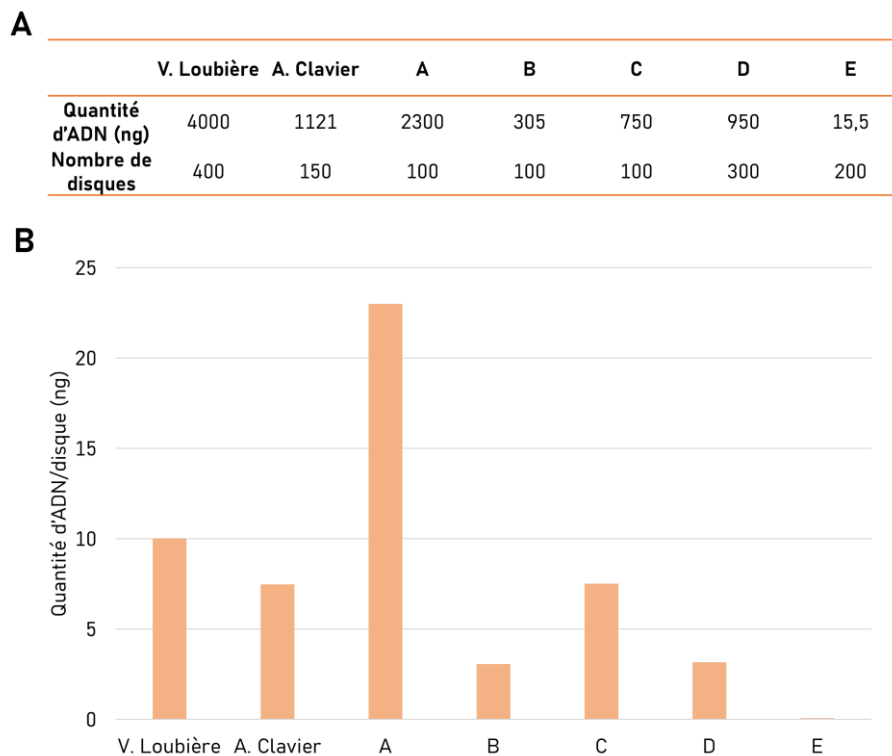
**Figure 32 : Evaluation de la présence d'une étiquette HA dans la souche contrôle.** Les disques imaginaux d'ailes des trois génotypes ont été disséqués et marqués par immunofluorescence avec des anticorps anti-GFP et anti-HA. Les disques présentés ici sont représentatifs des observations et le côté antérieur est orienté vers la gauche. Les images ont été prises à un grossissement x100, zoom 1,5.

Nous avons donc réalisé des immunomarquages anti-GFP et anti-HA dans la souche portant la protéine Rbf1 étiquetée utilisée dans nos expériences de ChIP comme contrôle positif, une souche *vg>eGFP* comme contrôle négatif et la souche *vg>mCD8-GFP* d'intérêt. Comme attendu, le

marquage GFP est présent dans toutes les souches. En revanche, contrairement à ce qui est observé en présence de Rbf1-HA, aucun marquage HA n'est visible en immunofluorescence lorsque seul CD8-GFP est exprimé, confirmant que ces disques imaginaires sont effectivement dépourvus d'une étiquette HA (Figure 32). Cette dernière hypothèse n'expliquait donc toujours pas les résultats obtenus en ChIP-qPCR.

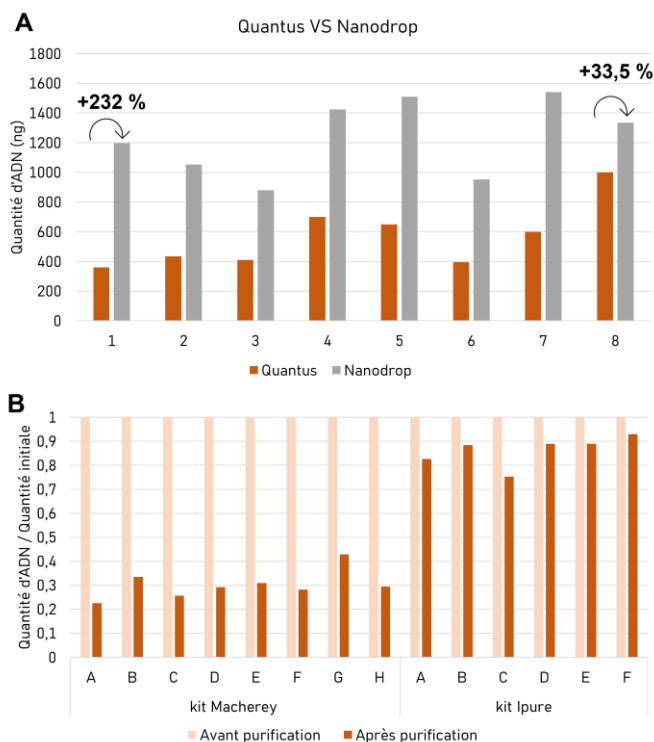
Un dernier espoir : une purification de l'ADN trop peu efficace

Les deux hypothèses que nous avons envisagées se sont donc révélées être fausses. Cependant, en reprenant les résultats des nombreuses ChIP-qPCR effectuées, nous avons pu observer de grandes disparités dans les quantités d'ADN initiales, et ce malgré un nombre de disques imaginaires d'ailes équivalent utilisé au sein de chaque expérience. De plus, la comparaison des quantités que nous obtenions avec celles obtenues par de précédents membres de l'équipe [159] et dans la littérature [187] a révélé un véritable problème dans la récupération de l'ADN ou dans sa purification (Figure 33).



**Figure 33 : Comparaison des quantités d'ADN initiales obtenues avec notre protocole ou dans la littérature.** (A) Tableau des quantités d'ADN obtenues et du nombre de disques imaginaires d'ailes utilisés dans la littérature (Loubière et al.), par d'anciens membres du laboratoire (Amandine Clavier) ou dans nos mains pour cinq expériences (A à E). (B) Comparaison de la quantité d'ADN en ng par disque imaginal d'aile associée à ces différents protocoles

Nous avons alors testé plusieurs paramètres afin d'évaluer la cause de cette perte de matériel. Parmi eux, nous avons évalué les effets de la méthode de dosage et du kit de purification des acides nucléiques, qui sont présentés (Figure 34). Le type de tubes utilisés, l'impact de la centrifugation ou encore l'effet de la présence de GFP sur le dosage ont également été testés et n'influent pas sur la quantité d'ADN récupéré. Ces données ne sont pas présentées dans ce manuscrit.



**Figure 34 : Comparaison des méthodes de dosage et de purification de l'ADN.** (A) Comparaison de la quantité d'ADN dosé selon la méthode de dosage (Nanodrop ou Quantus). L'ADN a été dosé par les deux méthodes à partir des huit mêmes échantillons, notés 1 à 8. (B) Comparaison de la perte de matériel durant la purification selon le kit utilisé : kit Macherey-Nagel « PCR Clean-Up » ou kit Ipure (Diagenode). Respectivement 8 et 6 échantillons ont été dosés au Quantus avant et après purification. Les résultats présentés sont les valeurs normalisées par rapport à la quantité avant purification.

Deux méthodes sont classiquement utilisées pour mesurer la quantité d'ADN dans un échantillon : la spectrophotométrie et la fluorimétrie. La spectrophotométrie mesure l'absorbance d'une goutte de 1 à 2  $\mu\text{L}$  d'échantillon et indique sa concentration en acides nucléiques d'intérêt (ARN ou ADN) ainsi que sa contamination en protéines et autres acides nucléiques. Au laboratoire, le spectrophotomètre disponible est un Nanodrop, qui permet des détections allant de 1,5 à 3700  $\text{ng}/\mu\text{L}$ . Il est habituellement utilisé par notre équipe pour quantifier les acides nucléiques purifiés. La fluorimétrie, quant à elle, se base sur des sondes spécifiques des différents acides nucléiques (ADN double brin, ARN, ...) couplées à des fluorochromes. Après réaction de ces sondes avec l'échantillon à quantifier, l'appareil mesure la fluorescence associée à l'échantillon pour mesurer une quantité. La plateforme de génomique de l'UVSQ possède un fluorimètre Quantus ainsi que la sonde permettant la quantification de l'ADN double brin entre 0,2 et 400  $\text{ng}/\mu\text{L}$ . Si le spectrophotomètre permet de quantifier de plus grandes quantités, il est aussi connu pour être moins sensible et moins précis. En effet, la mesure d'absorbance peut surestimer la quantité d'ADN présente dans le tube en considérant une partie de l'ARN comme de l'ADN. De plus, elle est peu efficace à de faibles concentrations.

## RESULTATS

Nous avons généré huit échantillons, différant par le type de tubes, le crosslink et les centrifugations utilisés, contenant chacun 50 disques imaginaires d'ailes disséqués. Ces échantillons ont subi les étapes du protocole de ChIP classique à l'exception de l'immunoprécipitation, à savoir la fixation (sauf pour deux des échantillons), la sonication et le « décroscrosslink ». Ils ont ensuite été purifiés et dosés au Nanodrop et au Quantus. Nous avons observé une forte surestimation de la quantité d'ADN du Nanodrop par rapport au Quantus, d'une moyenne de +131 %, indépendamment des paramètres de génération de l'échantillon. La plus forte surestimation observée est celle de l'échantillon 1, dans lequel la quantité d'ADN mesurée avec le Nanodrop est plus de trois fois supérieure à celle mesurée au Quantus (Figure 34A). La concentration mesurée au Nanodrop, de 23,9 ng/μL, entrait pourtant dans la gamme de détection de cet appareil. Le Nanodrop ne semble donc pas adapté pour quantifier l'ADN purifié après immunoprécipitation. En effet, si sans immunoprécipitation les valeurs obtenues entrent dans la gamme de détection du Nanodrop, la quantité obtenue après purification, diminuée au moins dix fois, n'est plus contenue dans cette gamme, ce qui risque d'augmenter encore la marge d'erreur. Certaines de nos données de référence sont donc potentiellement surestimées (protocole « A. Clavier », Figure 33).

Enfin, nous nous sommes intéressés à l'efficacité du kit de purification d'ADN que nous utilisons, le kit « PCR Clean-Up » (Macherey-Nagel). En comparant la quantité d'ADN dosée au Quantus avant et après purification, nous nous sommes aperçus d'une perte de matériel de l'ordre de 70 %. Nous avons alors testé un second kit, déjà utilisé au laboratoire. Celui-ci a permis une récupération de plus de 80 % du matériel initial (Figure 34B). Nous avons donc choisi d'utiliser ce dernier.

Malheureusement, l'amélioration de la purification n'a pas permis d'obtenir des ChIP de contrôle convenables. Sur les conseils des membres du comité de thèse, nous avons choisi de procéder directement à une expérience de ChIP-seq de contrôle. En effet, les amorces utilisées étaient les mêmes que ceux utilisés et publiés précédemment par le laboratoire, mais avec des fragments d'ADN plus longs : nous souhaitions des fragments de 200 à 300 paires de bases, alors que les fragments utilisés auparavant dans le laboratoire étaient soniqués à 400-500 paires de bases. La taille des fragments que nous obtenons étant du même ordre que la taille des amplicons issus de la PCR avec nos amorces (environ 150 paires de bases), il est donc possible que ces problèmes d'amplification des cibles viennent simplement du fait que la probabilité d'hybridation des deux amorces sur le même fragment de chromatine est faible. Cette expérience de ChIP-seq de contrôle avait pour but de déterminer plus précisément la localisation des pics associés à Rbf1 et ainsi de choisir de nouvelles amorces plus adaptées à la taille de nos fragments.

## 2. Séquençage des fragments obtenus

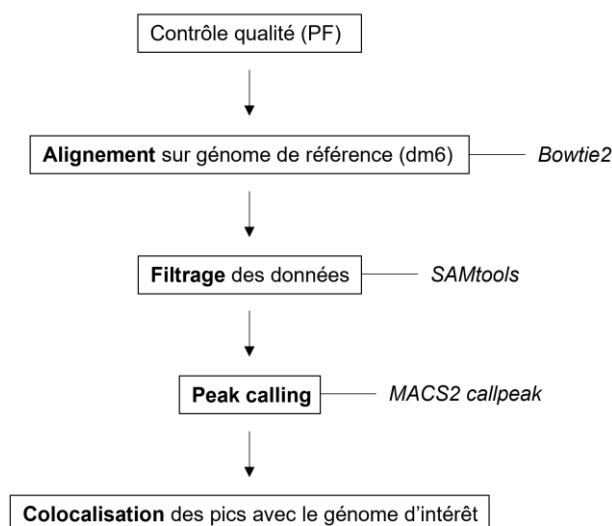
Pour la création des banques et le séquençage des échantillons, nous avons choisi de faire appel à la plateforme de séquençage de l'I2BC, recommandée par mes encadrantes et les membres du comité de suivi de thèse. Sur les conseils de ces derniers et des experts de la plateforme de séquençage, nous avons choisi de réaliser un séquençage Illumina paired-end d'une profondeur de 10 millions de reads. Pour cela, les échantillons devaient comporter un minimum de 10 ng d'ADN purifié.

Ces discussions ont eu lieu durant ma première année de thèse. Malheureusement, les mises au point et les nombreux problèmes encourus décrits dans la section précédente ont retardé mon envoi des échantillons de près d'un an. Entre temps, la crise de la COVID-19 nous a frappé et la plateforme de séquençage s'est trouvée surchargée. Si la prise en charge de notre expérience de RNA-seq ou d'une expérience complète de ChIP-seq n'était plus envisageable, ils ont cependant accepté de réaliser la banque et le séquençage d'un échantillon test qui nous permettrait de construire de nouvelles amorces.

Un échantillon de la fraction « IP » issue d'une ChIP en condition *vg>Rbf1*, *CD8* a donc été généré, d'une concentration de 2,76 ng/ $\mu$ L pour 15  $\mu$ L, soit 41,4 ng. Il a été déposé en mains propres à la plateforme de séquençage et les données ont été reçues quelques semaines plus tard.

## 3. Traitement des données de séquençage

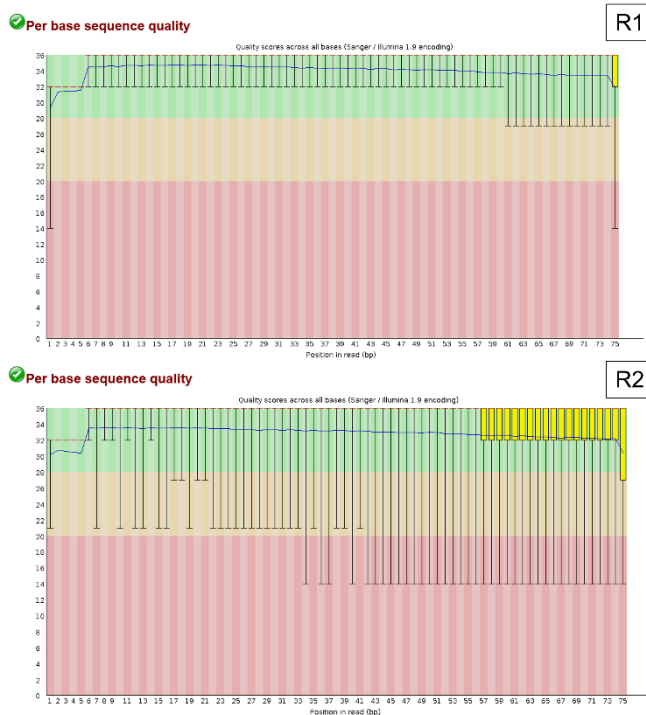
N'ayant pas de formation de bio-informatique ou de bio-informaticien au laboratoire, et dans un souci de temps et de facilité, nous avons choisi d'utiliser l'outil en ligne Galaxy du Galaxy Project [194]. Cette plateforme permet d'utiliser de nombreux outils et de les inclure dans des workflows (Figure 35), sans avoir à connaître la syntaxe de code nécessaire à leur utilisation. The Galaxy Project propose également des tutoriels expliquant les différentes étapes à suivre. J'ai d'ailleurs utilisé le tutoriel « Analysis of ChIP-seq data » [195] pour l'analyse de ces données.



**Figure 35 : Workflow utilisé pour l'analyse des données de ChIP-seq.** L'analyse a été réalisée grâce à la plateforme en ligne Galaxy, à l'exception du contrôle qualité réalisé par la plateforme de séquençage à la suite de la génération des données. Les outils utilisés pour la réalisation des différentes étapes est indiqué à droite en italique.

a. Contrôle qualité des données

La génération de données de contrôle qualité ainsi que le trimming des adaptateurs ont été réalisés par la plateforme. Le séquençage ayant été réalisé en paired-end, ces données sont sous la forme de deux fichiers html, un pour chaque sens de séquençage (appelés R1 et R2). Le programme fastqc, qui génère ces rapports, évalue les données selon une liste de 11 critères de qualité parmi lesquels la qualité de la séquence par base, le contenu en GC ou encore les niveaux de duplication de la séquence. Si certains d'entre eux posent problème, le rapport peut le signaler par deux types d'avertissement : un point d'exclamation dans un cercle orange, signifiant un point de vigilance, ou une croix dans un cercle rouge, signifiant un échec du test de qualité.



**Figure 36 : Diagramme de qualité de la séquence par base pour les reads issus du séquençage des brins sens (R1) et anti-sens (R2).** Le score de qualité Phred est indiqué pour chaque base de chaque read. Le score peut être bon (zone verte), moyen (zone orange) ou mauvais (zone rouge). La ligne bleue représente la qualité moyenne et les boîtes jaunes l'intervalle inter-quartile (25-75%). Ce diagramme a été généré par la plateforme de séquençage lors du contrôle qualité, en utilisant le programme fastqc.

Les rapports associés à notre échantillon indiquent une bonne qualité, comme le montrent les deux diagrammes de qualité de la séquence par base. Dans ces diagrammes, un score de qualité Phred est associé à chacune des bases de chaque séquence. Ces scores peuvent être bons (entre 28 et 36, zone verte), moyens (entre 28 et 20, zone orange) ou mauvais (entre 0 et 20, zone rouge). Pour R1, la grande majorité de ces scores sont dans la zone verte, indiquant une bonne qualité des séquences. Le programme fastqc n'a pas émis d'avertissement vis-à-vis de ce résultat. Pour R2, la qualité des reads semble moins bonne, mais le programme n'a pas non plus émis d'avertissement quant à ce résultat (Figure 36). Des recherches sur les forums de bioinformatique indiquent qu'il est habituel d'obtenir une qualité moindre pour le brin séquençé en deuxième.

### b. Alignement sur un génome de référence

Les reads ont ensuite été alignés sur le génome de référence, ici le dm6. Il correspond au 6<sup>e</sup> assemblage du génome de *Drosophila melanogaster*, réalisé en août 2014. C'est le dernier assemblage en date. Cette étape associe les séquences obtenues à une position correspondante sur le génome.

```

14520241 reads; of these:
  14520241 (100.00%) were paired; of these:
    1469740 (10.12%) aligned concordantly 0 times
    9561712 (65.85%) aligned concordantly exactly 1 time
    3488789 (24.03%) aligned concordantly >1 times
  ----
  1469740 pairs aligned concordantly 0 times; of these:
    180776 (12.30%) aligned discordantly 1 time
  ----
  1288964 pairs aligned 0 times concordantly or discordantly; of these:
    2577928 mates make up the pairs; of these:
      1854832 (71.95%) aligned 0 times
      448731 (17.41%) aligned exactly 1 time
      274365 (10.64%) aligned >1 times
93.61% overall alignment rate

```

**Figure 37 : Statistiques d'alignement.** Ces statistiques sont générées par le programme Bowtie2 à la suite de l'alignement des reads sur le génome dm6.

Pour l'alignement, le programme Bowtie2 a été utilisé et a pu générer un taux d'alignement global de 93,61 %. 100 % des reads ont pu être appairés (R1 et R2) et parmi eux, 65,85 % des reads sont alignés exactement une fois (Figure 37) soit 9561712 reads. Ces données sont satisfaisantes.

### c. Filtrage des données

Le filtrage des données permet de filtrer les reads dont la qualité ne serait pas suffisante. Nous avons choisi de retirer les reads dont le score de qualité MAPQ était inférieur à 20 (zone rouge Figure 36). Ce filtrage a été fait grâce à l'outil « Filter SAM or BAM, output SAM or BAM files on FLAG MAPQ RG LN or by region » qui fait partie de l'ensemble SAMtools et a permis de retirer environ deux millions de reads.

### d. Peak calling

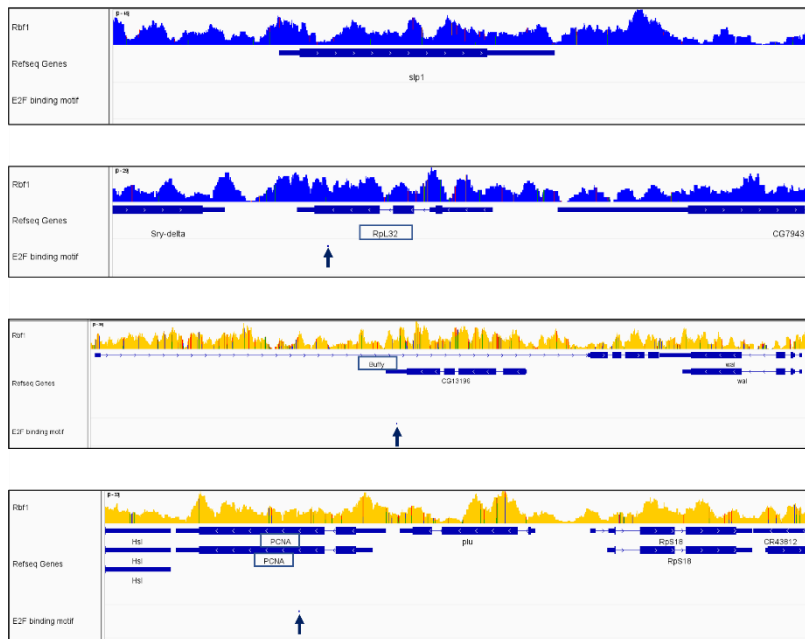
Le peak calling permet de lire la densité de reads générée par l'alignement aux différentes positions, et ainsi de la traduire sous forme de pics. Plus il y a de reads associés à une position, plus le pic sera important et plus la séquence d'ADN correspondante sera considérée comme une cible du facteur d'intérêt, ici Rbfl. L'outil MACS2 callpeak a été utilisé pour cette étape. Il détermine comme un pic une zone comprenant dix à trente fois plus de reads dans la condition d'intérêt que dans la condition contrôle sur une fenêtre équivalente à deux fois la taille des fragments (ici 2x75, soit 150 nucléotides).

Le peak calling est suivi d'une étape de colocalisation des pics. En effet, une fois que les coordonnées des pics sont déterminées, il faut ensuite les croiser avec les informations du génome de



référence pour obtenir des informations sur les gènes identifiés comme cible. Cette étape a été réalisée avec l'outil « Intersect the intervals of two datasets » pour croiser le génome dm6 chargé à partir du UCSC Genome Browser et le fichier obtenu à l'issue du peak calling. Les données sont alors prêtes à être visualisées dans des navigateurs de génome comme Integrated Genome Browser (IGB) et Integrative Genomics Viewer (IGV). C'est ce dernier que nous avons choisi d'utiliser.

#### 4. Visualisation des données dans IGV et analyse ChIP-seq



**Figure 38 : Visualisation du résultat de l'expérience de ChIP-seq au niveau de contrôles négatifs et positifs (gènes cibles de Rbf1).** La couverture de gènes contrôles (bleu, *slp1* et *RpL32/rp49*) et de gènes cibles de Rbf1 (orange, *buffy* et *PCNA*) est associée au schéma des gènes correspondants et à la présence du motif de liaison de E2F (TTGGCGC), signalé par une flèche bleue. Lorsque plusieurs gènes sont présents, le gène d'intérêt est encadré en bleu. Ces pistes ont été générées par IGV à partir des pics croisés avec le génome de référence dm6.

Les navigateurs de génome permettent de visualiser les données générées à l'issue de l'étape précédente. Rbf1 étant un co-facteur de transcription majoritairement retrouvé à proximité des TSS [112,113], nous nous attendions à trouver des pics au niveau des promoteurs des gènes cibles connus, comme *buffy* ou *PCNA*, et une absence de pics aux promoteurs des gènes contrôles comme *rp49* (aussi décrit sous le nom de *RpL32*) et *slp1*. Malheureusement, comme présenté sur la Figure 38, ni le promoteur de *buffy* ni celui de *PCNA* ne présentait de pic particulier. Le nombre de reads alignés sur ces gènes n'est pas différent de ceux alignés sur les gènes contrôle. L'observation globale de la répartition des reads sur le génome n'a révélé aucun pic ni même d'enrichissement particulier dans les régions contenant des gènes.

Cette expérience test ne nous aura pas permis d'identifier des régions d'intérêt pouvant être utilisées comme témoins positifs ou négatifs pour améliorer notre expérience. Cette expérience ne nous ayant pas plus renseignées sur la nature du problème technique nous empêchant de mener à bien nos ChIP, nous avons choisi d'en rester là. En effet, chercher à résoudre un problème dont nous n'arrivons pas à identifier la nature a de grandes chances d'être chronophage et vain, , nous avons donc choisi d'abandonner cette partie du projet pour nous concentrer sur la partie transcriptomique.

Nous avons ainsi fait évoluer le projet : une fois les cibles régulées par *Rbf1* identifiées, nous souhaitons en sélectionner quelques-uns impliqués dans les problématiques de la thèse (apoptose) et du laboratoire (mitochondrie et/ou réticulum endoplasmique). Nous aurions alors déterminé par ChIP-qPCR si ces cibles sont ou non des cibles directes de *Rbf1*.

## IV- Le RNA-seq : une méthode d'identification des régulations dépendantes de *Rbf1*

Comme évoqué précédemment, l'expérience de RNA-seq en condition contrôle et lorsque *Rbf1* est surexprimé devait nous permettre d'identifier des cibles régulées par *Rbf1* dans le disque imaginal d'aile. Pour limiter le bruit de fond généré par les cellules hors du domaine d'intérêt, nous ne souhaitons évaluer que le transcriptome des cellules surexprimant *Rbf1* et CD8 ou CD8 uniquement. Nous avons donc mis au point un tri cellulaire pour exclure les cellules hors du domaine *vestigial*. La crise sanitaire ne nous ayant pas permis de faire appel à la même plateforme de séquençage que pour le ChIP-seq, nous nous sommes tournées vers l'entreprise Fasteris, conseillée par l'un des membres du comité de suivi de thèse. Les discussions avec les experts de cette société nous ont menées à choisir un séquençage NextSeq500 Mid-output garantissant 15 à 25 millions de reads appairés (paired-end) d'une longueur de 2x75 paires de bases.

Entre les analyses du ChIP-seq et du RNA-seq, j'ai eu le temps de me former à la bio-informatique grâce à la formation NGS (Next Generation Sequencing) & Cancer proposée par le Cancéropôle Île-de-France et à l'aide du Dr. Frédéric Crémazy (UMR 1173, UVSQ). J'ai également pu accéder au serveur en ligne de la plateforme de génomique de l'UVSQ et ainsi réaliser mes analyses sous Linux. Un script complet d'une de ces analyses sous Linux est disponible en Annexe II. Une partie a également été réalisée sur RStudio.

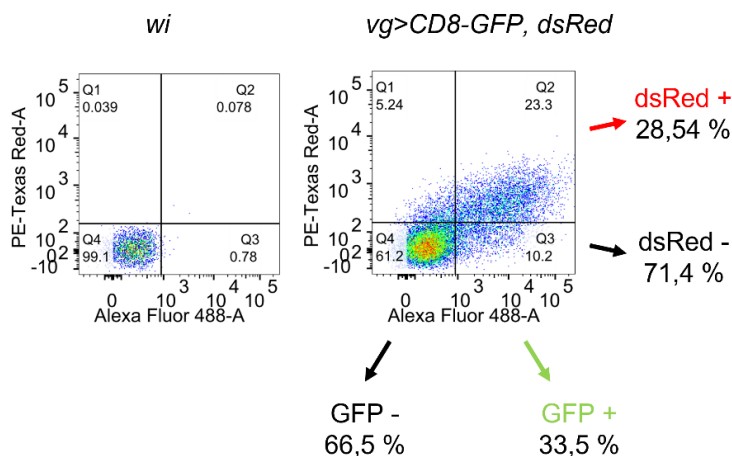
### 1. Tri magnétique ou tri par cytométrie en flux ?

. Le domaine *vestigial* représente 40 à 50 % des cellules du disque imaginal d'aile en condition sauvage [192]. Les cellules d'intérêt représentant donc moins de la moitié de la population récupérée à l'issue de l'isolement des disques, le signal spécifique pourrait se trouver écrasé par un bruit de fond important provenant des ARN issus des cellules hors du domaine *vestigial*. Afin de limiter ce bruit de fond, nous souhaitons récupérer uniquement les cellules du domaine *vestigial* surexprimant *Rbf1*. Pour cela, le projet initial prévoyait de mettre au point un tri magnétique permettant de précipiter les cellules d'intérêt. Nous nous sommes alors procuré des souches permettant l'expression d'une protéine chimérique CD8-GFP, CD8 permettant l'ancrage membranaire de la GFP sous contrôle de séquences UAS, appelée mCD8-GFP. Par recombinaison, nous avons créé une souche permettant la co-surexpression de *Rbf1* étiqueté et de ce mCD8-GFP. Des billes magnétiques couplées à des anticorps

anti-CD8 auraient reconnu les épitopes CD8 à la membrane des cellules d'intérêt, cet épitope étant co-surexprimé avec *Rbfl* dans le domaine *vestigial*. Nous souhaitions utiliser un protocole publié qui a permis d'isoler de la même façon des neurones de drosophile [196]. Cependant, cette technique aurait demandé une mise au point longue. Elle était également source de contraintes techniques car devait utiliser du matériel par ailleurs employé pour des cellules mammaliennes en culture. Les larves se développant dans un milieu contenant de la levure, utiliser cette technique signifiait risquer la contamination de ces cultures cellulaires.

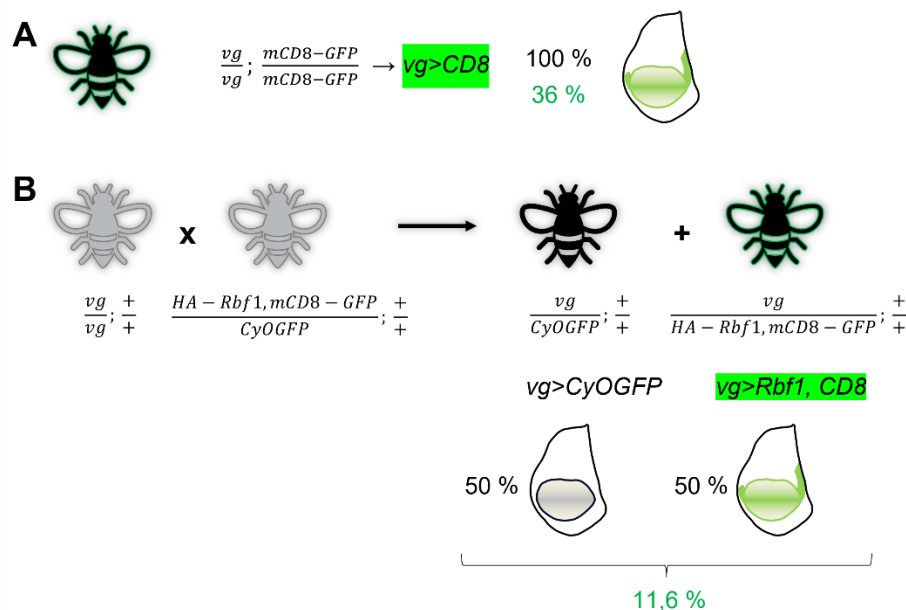
La plateforme d'imagerie et de cytométrie de l'UVSQ possède un cytomètre en flux capable de trier les cellules régulièrement utilisé pour le tri de cellules de mammifères. Bénéficiant de l'expérience de chercheurs de l'équipe voisine réalisant du tri cellulaire par cytométrie en flux en routine, nous nous sommes tournées vers cette technique. Le CD8 membranaire étant couplé à une GFP, nous souhaitions utiliser cette fluorescence pour récupérer uniquement les cellules du domaine *vestigial* (fluorescentes) et exclure les autres. Par ailleurs, des expériences de cytométrie en flux sur des cellules de disques imaginaux avaient précédemment été réalisées au laboratoire. Ces travaux pouvaient donc nous guider quant au paramétrage à utiliser pour réaliser un tri cellulaire, non plus sur des cellules mammaliennes, mais sur des cellules de disques imaginaux dissociés.

Nous souhaitions vérifier la compatibilité avec le tri de cellules de disque imaginal d'aile de drosophile, constitué de cellules de petite taille ; et la capacité de la GFP couplée au CD8 à se différencier de l'auto-fluorescence des cellules. Comme évoqué précédemment, le domaine *vestigial* représente environ 45 % des cellules du disque imaginal d'aile en condition sauvage, nous avons donc vérifié si cette proportion était retrouvée lors du tri cellulaire de disques d'aile. Afin de déterminer quel fluorochrome serait le plus adapté pour le tri de nos cellules, nous en avons testé deux différents : la GFP (couplée au CD8) et le dsRed, tous deux induits dans le domaine *vestigial*. Cette expérience a permis de déterminer un fenêtrage optimal pour la taille des cellules du disque imaginal d'aile, ainsi que de valider le tri des cellules par la GFP. En effet, c'est avec ce fluorochrome que la proportion de cellules fluorescentes était la plus proche des 40 % attendus (Figure 39).



**Figure 39 : Répartition des cellules de disques imaginaux d'aile exprimant mCD8-GFP et la dsRed dans le domaine *vestigial*.** Les cellules *wi* sont des cellules contrôles n'exprimant aucun des fluorochromes. Les cellules analysées ici ont précédemment été sélectionnées selon leur taille et leur granulométrie afin de retirer les débris et les amas cellulaires.

Pour évaluer le nombre de cellules que nous pouvions espérer récolter, nous avons ensuite voulu déterminer la proportion de cellules [GFP +] attendue dans des disques imaginaux surexprimant *Rbf1* et *CD8-GFP* sous contrôle du pilote *vg-Gal4*. Ces disques sont obtenus à l'issue d'un croisement entre des individus *vg-Gal4* et *UAS-Rbf1*, *UAS-CD8-GFP* balancés. Ainsi, seulement 50 % environ de la descendance de ce croisement possèdera le génotype d'intérêt *vg>Rbf1*, *CD8* ; l'autre moitié ayant hérité du chromosome balancier (Figure 40B). L'intérêt de ce tri étant de pouvoir récupérer uniquement les cellules d'intérêt, les larves issues de ce croisement ne sont pas triées manuellement pour séparer les larves possédant le génotype d'intérêt des larves possédant le balancier. Le balancier n'induisant pas de fluorescence dans les cellules, nous nous attendions donc à un maximum de 20 % de cellules GFP-positives. De plus, *Rbf1* est un inducteur d'apoptose dans le disque imaginal d'aile, cette proportion pouvait donc être bien inférieure selon la quantité d'apoptose induite. Les premiers tests nous ont effectivement permis d'observer qu'environ 8 % des cellules sont GFP-positives parmi les cellules issues de disques récupérés par la technique d'isolement en masse.



**Figure 40 : Pourcentage de cellules GFP-positives attendues pour chaque génotype.** (A) Les disques imaginaux *vg>CD8* sont issus d'une lignée établie et homozygote, 100 % des disques récupérés ont une fluorescence dans le domaine vestigial représenté en vert. La moyenne de dix expériences de tri indique une proportion attendue d'environ 36 % de cellules GFP-positives. (B) Les disques imaginaux *vg>Rbf1,CD8* sont issus de la descendance du croisement entre des femelles vierges homozygotes pour le pilote *vg-Gal4* et des mâles hétérozygotes recombinants *HA-Rbf1*, *mCD8-GFP*. Seule la moitié des descendants, et donc des disques imaginaux récupérés, ont une fluorescence dans le domaine *vestigial*. La moyenne de dix expériences de tri indique une proportion totale attendue d'environ 11,6 % de cellules GFP-positives.

Les nombreuses répétitions de ces expériences ont pu déterminer une moyenne de 36 % de cellules GFP-positives dans les disques imaginaux contrôles *vg>CD8* (Figure 40A) et de 11,6 % dans les disques surexprimant *Rbf1* (*vg>Rbf1*, *CD8*) (Figure 40B). Ces expériences nous ont permis de

## RESULTATS

déterminer qu'une expérience de tri permettait de récupérer plusieurs dizaines voire centaines de milliers de cellules par expérience. Ce nombre est bien sûr dépendant du nombre de larves disponibles pour la récupération des disques imaginaux. Extraction des ARN et séquençage

La plateforme de séquençage recommandait une quantité d'ARN totaux de 250 à 500 ng par échantillon. Nous souhaitions réaliser ces expériences en triplicats, cela représentait donc six échantillons à générer : trois échantillons de cellules surexprimant Rbf1 et CD8, et trois échantillons de cellules contrôles surexprimant CD8 seul. Lors de la mise au point de cette expérience, nous nous sommes aperçu qu'une extraction sur les cellules issues d'un seul tri ne suffisait pas à obtenir la quantité d'ARN requise et nous avons déterminé que  $2,5 \times 10^6$  cellules seraient nécessaires pour générer ces échantillons. Une étude de 2002 comparant les profils d'expression des disques imaginaux, qui détermine qu'un disque imaginal d'aile de larve de stade L3 contient 1 à 4 ng d'ARN [197]. En considérant qu'un disque imaginal d'aile est constitué d'environ 50 000 cellules, 250 ng d'ARN dans ces conditions représenterait  $3,1 \cdot 10^6$  à  $1,25 \cdot 10^7$  cellules. Notre estimation est donc concordante avec la littérature.

Cette quantité de cellules a représenté entre huit et onze tris cellulaires pour les cellules surexprimant Rbf1, CD8 et entre deux et quatre tris pour les cellules CD8. Ces différences sont dues au fait que le nombre de disques imaginaux, et donc de cellules, récupérés peut varier d'une expérience à l'autre selon le nombre de larves disponibles au début de l'expérience. Malheureusement, ces quantités n'ont pas suffi, puisque seul un des échantillons avait une quantité d'ARN total établie par dosage au Quantus supérieure à 250 ng. Les autres quantités mesurées étaient comprises entre 143 et 247 ng. Il est possible que la différence observée entre nos estimations et les quantités d'ARN extraites provienne du stockage des cellules après leur tri en cytométrie. En effet, celles-ci sont conservées dans du TRIzol à  $-20^{\circ}\text{C}$ . Il est possible que ces conditions dégradent certaines cellules, entraînant ainsi une perte de matériel.

De nouveaux tris et extractions ont été réalisés pour compléter les échantillons précédents. Ils ont ainsi pu être envoyés à la plateforme.

	Rbf1						CD8				
	A	B	C	D	E	F	A	B	C	D*	E
LGBC	308	305	297	375	637,5	240	380	382	379	450	513
PF	430,2	108,9	259,5	191	142,5	73,9	323,2	357,4	353,7	170	135,2

**Figure 41 : Des divergences de quantité entre les dosages du laboratoire et les dosages de la plateforme.** En tout, trois groupes d'échantillons ont été envoyés à la plateforme (PF), qui ont réalisés un dosage de la quantité d'ARN totaux et un contrôle qualité. Un de ces échantillon, signalé par un astérisque (\*), était dégradé. Nous avons choisi, sur les conseils de la plateforme, de séquencer les échantillons en vert. Dans les notations, ils sont alors devenus Rbf1\_1 (A), Rbf1\_2 (C), Rbf1\_3 (F), Rbf1\_4 (G) ; et CD8\_1 (A), CD8\_2 (C) et CD8\_3 (F)

Nous avons cependant rencontré deux problèmes avec les échantillons envoyés à la plateforme. Premièrement, les quantités d'ARN total dosées par la plateforme à l'arrivée des échantillons ne correspondaient pas à celles dosées par nos soins. Ainsi, la plupart d'entre elles se sont trouvées insuffisantes pour réaliser les banques et le séquençage (Figure 41). Nous avons donc dû réaliser trois envois d'échantillons. De plus, un des échantillons est arrivé dégradé, probablement à la suite d'un problème dans l'acheminement du colis. Deuxièmement, deux échantillons ont été accidentellement mélangés, compromettant ainsi leur indépendance biologique (Figure 41, échantillons Rbf1 A et C).

A la suite du troisième envoi (échantillons Rbf1 E/F et CD8 E), nos échantillons ne répondaient toujours pas aux critères demandés malgré une quantité d'ARN suffisante mesurée au Quantus par nos soins (Figure 41). Tous les échantillons ayant une qualité convenable (sauf CD8 D, signalé par un astérisque), la plateforme nous a alors proposé de séquencer sept des onze échantillons envoyés parmi ceux ayant les quantités les plus élevées, dont le nom est en vert dans les tableaux Figure 41, en nous assurant qu'une quantité légèrement inférieure à 200 ng ne devrait pas poser de problèmes pour la création des banques. Parmi eux, deux des échantillons Rbf1 ne sont pas totalement indépendants (A et C).

Malheureusement, au niveau de la plateforme, un problème technique a eu lieu lors de la génération des banques. Le matériel de départ étant déjà faible, le matériel restant pour une seconde préparation ne correspondait pas du tout aux quantités demandées. Malgré le risque de génération d'un nombre important de duplicats de PCR, la plateforme a décidé unilatéralement de préparer une nouvelle banque, qui, compte tenu de la faible quantité d'ARN, a nécessité 17 cycles de PCR au lieu des 14 habituels. Les duplicats de PCR sont des biais induits lors de l'amplification des ARN pour la création des banques, ils sont issus de la même molécule d'ADN complémentaire. Les reads issus des duplicats ne sont pas utilisables. Si le taux de duplicats est trop élevé, le nombre de reads restants est trop faible et la couverture et la profondeur qui en résultent ne sont alors plus adaptées à l'analyse. De plus, un taux trop élevé de duplicats de PCR risquerait de créer des pics artificiels et donc de fausser les résultats en faisant ressortir de fausses cibles et/ou en masquant de réelles amplifications biologiques.

La plateforme a alors proposé de réaliser un run test Mi-Seq de faible profondeur (1 million de reads) avant de lancer le séquençage prévu (15 à 25 millions de reads) afin de vérifier le taux de duplicats de PCR dans les banques créées. Pour cela, le programme fastp a été utilisé. Il permet de réaliser un contrôle qualité des échantillons grâce à des paramètres semblables à ceux de fastqc utilisé pour la ChIP-seq (score de qualité, taux de GC, taux de duplication, ...), ainsi que de filtrer les reads pour ne garder que ceux ayant une qualité suffisante et de retirer les adaptateurs nécessaires au séquençage. fastp permet de traiter ensemble les paires de reads (R1 et R2) des échantillons pour donner une qualité globale de la paire de données.

Rbf1_1	Rbf1_2	Rbf1_3	Rbf1_4	CD8_1	CD8_2	CD8_3
0,36 %	2,01 %	2,73 %	1,58 %	1,25 %	2,64 %	2,43 %

**Figure 42 : Taux de duplication des échantillons séquencés par le Mi-Seq.** Ces taux ont été déterminés par le programme fastp sur un séquençage test d'une profondeur de 1 million de reads.

Les taux de duplicats évalués par fastp étaient tous inférieurs à 5 % (Figure 42), nous avons donc considéré qu'ils étaient satisfaisants et que, de façon surprenante, les cycles de PCR supplémentaires effectués n'avaient pas eu d'impact sur la qualité des données. Nous avons donc réalisé le séquençage prévu.

## 2. Traitement et analyse des données

### a. Contrôle qualité

Après réception des données du séquençage prévu, un nouveau contrôle qualité a été réalisé avec fastp, selon la ligne de commande décrite ci-après (Figure 43). Ce programme permet également la détection et le retrait des adaptateurs de PCR, ainsi que l'élimination des reads de mauvaise qualité.

```
#Trimming/QC des reads

fastp \
  -h quality_report.html \
  -i échantillon_1_R1.fastq.gz \
  -I échantillon_1_R2.fastq.gz \
  -o échantillon_1_R1_trimmed.fastq.gz \
  -O échantillon_1_R2_trimmed.fastq.gz

#-h génération d'un rapport sous forme html
#-i fichier d'entrée (input), brin sens
#-I fichier d'entrée (input), brin anti-sens
#-o fichier de sortie (output), brin sens
#-O fichier de sortie (output), brin anti-sens
```

**Figure 43 : Exemple de commande utilisée pour le contrôle qualité et le retrait des adaptateurs.** Le programme fastp est utilisé.

Tous les indicateurs de qualité semblaient satisfaisants sauf un : le taux de duplication. Malheureusement, si les taux de duplication dans le séquençage préliminaire de vérification semblaient satisfaisants, ceux issus de ce séquençage sont, pour la grande majorité, bien trop élevés car supérieurs à 60 % (Figure 44). S'il ne semble pas y avoir de consensus sur le taux « normal » de duplicats de PCR pour une expérience de RNA-seq, nos résultats indiquent que plus de la moitié des reads obtenus sont des duplicats, ce qui nous semble anormal. En interrogeant les experts de la société ayant réalisé le séquençage à ce sujet, il semblerait que ce taux soit dû à l'augmentation de la profondeur de séquençage. Si la diversité de nos ARN est faible, ce qui semble être le cas, le premier séquençage en aurait séquencé la majorité. Ainsi, sur un séquençage plus profond, chaque nouveau read séquencé a une forte chance d'être un duplicat.

	Rbf1_1	Rbf1_2	Rbf1_3	Rbf1_4	CD8_1	CD8_2	CD8_3
Taux de duplication	25,40 %	66,96 %	74,34 %	60,99 %	52,81 %	70,63 %	67,40 %
Nombre de reads initiaux	29662563	21404033	23828154	25664084	20643937	22390296	27578867
Nombre de reads uniques	28310320	19851021	22574371	24651386	19323124	20723249	26406461

**Figure 44 : Taux de duplication et nombre de reads des échantillons séquencés par le Next-Seq.** Le taux de duplication a été déterminé par le programme fastp. Le nombre de reads initiaux et le nombre de reads uniques sont issus du rapport d'alignement par le programme STAR et correspondent respectivement au nombre de reads détecté par le programme et au nombre de reads alignés exactement une fois sur le génome (« number of input reads » et « uniquely mapped reads number »).

La question du retrait ou non des duplicats de PCR lors des expériences de RNA-seq ne semble pas faire consensus dans la communauté scientifique et certains déconseillent leur retrait qui pourrait induire un biais dans le traitement des données. En effet, il est possible que certains duplicats naturels, créés par des biais d'échantillonnage ou de fragmentation, soient reconnus comme duplicats de PCR alors qu'ils n'en sont pas [198]. Un moyen fiable d'identifier les duplicats de PCR est l'ajout de « codes-barres » appelés UMI (Unique Molecular Identifier) lors de la création des banques, avant toute étape de PCR. Cependant, cette étape n'a pas été incluse dans la préparation de nos données.

Nos taux de duplication étant très élevés, nous étions indécises sur l'impact de ces duplicats sur nos données. Compte tenu des cycles de PCR supplémentaires réalisés lors de la création des banques, il semble probable qu'au moins une partie de ces duplicats soient effectivement générés par la PCR. Cependant, leur proportion nous était inconnue, et, en les retirant, le nombre de reads restant serait très faible. Nous avons donc choisi de faire les analyses avec et sans retrait des duplicats. Pour cela, le programme Picard [199] a été utilisé (Figure 45). Le programme samtools flagstat a été utilisé pour évaluer le nombre de reads restants (Figure 45). Il est souvent utilisé pour regarder les statistiques d'alignement d'un fichier et parmi ces statistiques, il permet également de connaître le nombre de reads d'un fichier.



```
#Retirer les duplicats

java -jar picard.jar MarkDuplicates REMOVE_DUPLICATES=true \
  -I échantillon1_aligné.bam \
  -O échantillon1_aligné_sansduplicats.bam \
  -M résumé_échantillon1.txt

  #-I fichier à traiter
  #-O nom et format du fichier de sortie
  #-M résumé des opérations réalisées

#Voir le nombre de reads restants

samtools flagstat échantillon1_aligné_sansduplicats.bam
```

**Figure 45 : Exemple de commande utilisée pour le retrait des duplicats.** Le programme utilisé est picard. Pour voir le nombre de reads restants après retrait des duplicats, le programme utilisé est samtools flagstat.

Cependant, après retrait des duplicats, nous nous sommes trouvés face au problème anticipé plus haut : le taux de duplication étant très élevé pour certains échantillons, le nombre de reads restant après leur élimination devenait très faible. Pour un génome moyen, 15 à 20 millions de reads bruts sont recommandés, soit 30 à 40 millions lors d'un séquençage paired-end. Si on considère un taux d'alignement unique de 90 %, 27 à 36 millions de reads restent après l'alignement. Or, quasiment aucun de nos échantillons n'atteignait cette limite (Figure 46). Nous avons décidé de poursuivre tout de même l'analyse jusqu'au bout. En effet, nous espérions pouvoir identifier les gènes les plus différenciellement exprimés malgré le faible nombre de reads disponibles.

Echantillon	Taux de duplication	Reads initiaux (R1+R2)	Reads après filtre
Rbf1_1	25,4047 %	62 206 736	40 692 669
Rbf1_2	66,9574 %	46 815 800	7 091 880
Rbf1_3	74,3424 %	50 448 946	5 318 685
Rbf1_4	60,9938 %	54 277 158	12 719 676
CD8_1	52,8136 %	50 175 694	13 832 530
CD8_2	70,6257 %	50 220 326	7 110 840
CD8_3	67,4015 %	58 220 648	10 706 983

**Figure 46 : Nombre de reads avant et après filtrage des duplicats de PCR.** Le nombre de reads initiaux et le taux de duplication sont indiqués sur les rapports fastp. Le nombre de reads après filtre est compté par le programme samtools flagstat. Ce nombre est obtenu après retrait des duplicats sur les reads alignés une seule fois.

## b. Alignement sur le génome de référence

### Création du génome dm6

Si les génomes de référence étaient déjà préchargés sur la plateforme Galaxy, il a fallu le créer pour l'utiliser lors des analyses sous Linux. Pour cela, les données de la version dm6 du génome de *Drosophila melanogaster* sur la base de données Ensembl et l'assemblage a été réalisé grâce à la fonction « genomeGenerate » du programme STAR (Figure 47).

```
#Création du génome
STAR \
--runMode genomeGenerate \
--genomeSAindexNbases 12 \
--genomeDir chemin/génome \
--genomeFastaFiles dm6chr2L.fa dm6chr2R.fa dm6chr3L.fa dm6chr3R.fa dm6chr4.fa
dm6chrmito.fa dm6chrY.fa chrX.fa \
--sjdbGTFfile dm6.gtf \
--sjdbOverhang 74

#--genomeSAindexNbases permet de limiter la taille du fichier final et/ou
le temps de génération du fichier. Dépend de la taille du génome
#--genomeDir chemin vers le dossier contenant les fichiers du génome et
où sera généré l'assemblage
#--genomeFastaFiles fichiers du génome
#--sjdbGTFfile fichier des annotations du génome
#--sjdbOverhang taille des séquences génomiques à utiliser lors de
l'assemblage, correspond à la longueur des reads -1
```

**Figure 47 :** Exemple de commande utilisée pour la création du génome de référence dm6. Le mode genomeGenerate du programme STAR est utilisé.

### Alignement des reads sur le génome

Les reads ont ensuite été alignés sur le génome nouvellement créé. Cet alignement a été également généré avec le programme STAR [200] (Figure 48). Plus de 90 % des reads de chacun des échantillons ont pu être aligné de façon unique.

```
#Alignement avec STAR

STAR \
--genomeDir chemin/génome \
--readFilesIn échantillon_1_R1_trimmed.fastq.gz
échantillon_1_R2_trimmed.fastq.gz \
--readFilesCommand zcat \
--outSAMtype BAM SortedByCoordinate \
--outFileNamePrefix échantillon1_aligné \

#--genomeDir chemin vers le dossier contenant le génome de référence
#--readFilesIn fichiers R1 et R2 à aligner
#--readFilesCommand zcat décompression des fichiers gz
#--outSAMtype format du fichier de sortie (ici BAM) et type de tri des
données (ici par coordonnées)
#--outFileNamePrefix préfixe du fichier de sortie (output)

#Vérification du run : cat AVWR1_aligneLog.final.out
```

**Figure 48 :** Exemple de commande utilisée pour l'alignement sur le génome de référence. Le programme STAR est utilisé.

## Création d'un index du génome

La création d'un index du génome permet de retrouver plus facilement certaines séquences, à la manière d'un index dans un livre. Cela permet d'accélérer le temps de traitement des programmes. Pour cela, le programme samtools index de l'ensemble SAMtools a été utilisé ().

### c. Conversion des reads alignés en fichiers de comptage des reads

A la manière du peak calling, le comptage des reads permet de déterminer le nombre de reads alignés sur chaque gène. Cette étape est réalisée avec le programme htseq-count (Figure 49) et génère un fichier au format .counts .

```
#Conversion des BAM en raw counts (HTseq)
htseq-count -f bam -r pos -s no échantillon1_aligné.bam
Drosophila_melanogaster.BDGP6.32.107.gtf.gz > préfix_output.counts
  #-f : format de l'input
  #-r : paramètre de tri de l'input (ici, le fichier d'alignement est trié
par coordonnées donc par position (pos))
  #-s : pour "no", un read est considéré comme superposé avec un élément
génomique quelque soit le brin sur lequel il est aligné.
```

**Figure 49 : Exemple de commande utilisée pour le comptage des reads.** Le programme htseq-count est utilisé.

## 3. Analyse différentielle des données de séquençage

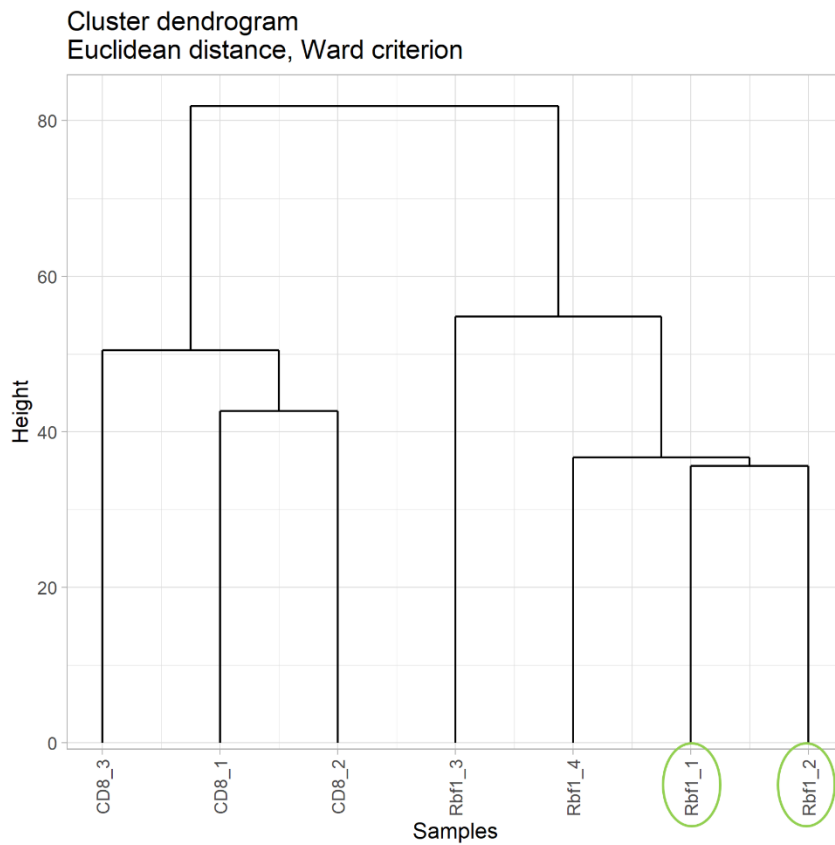
L'analyse des données a été réalisée grâce au package R SARTools [201], qui repose sur le programme DEseq2 et fonctionne sur R. Celui-ci réalise, à partir d'une table associant le nom des échantillons au fichier .counts et à leur condition (contrôle ou Rbf1). L'analyse différentielle des données et génère un fichier html contenant la description des données brutes, la variabilité au sein de l'expérience, la normalisation des données et l'analyse différentielle sous forme de différentes figures, dont certaines seront décrites ci-après.

### a. Description des données brutes, variabilité

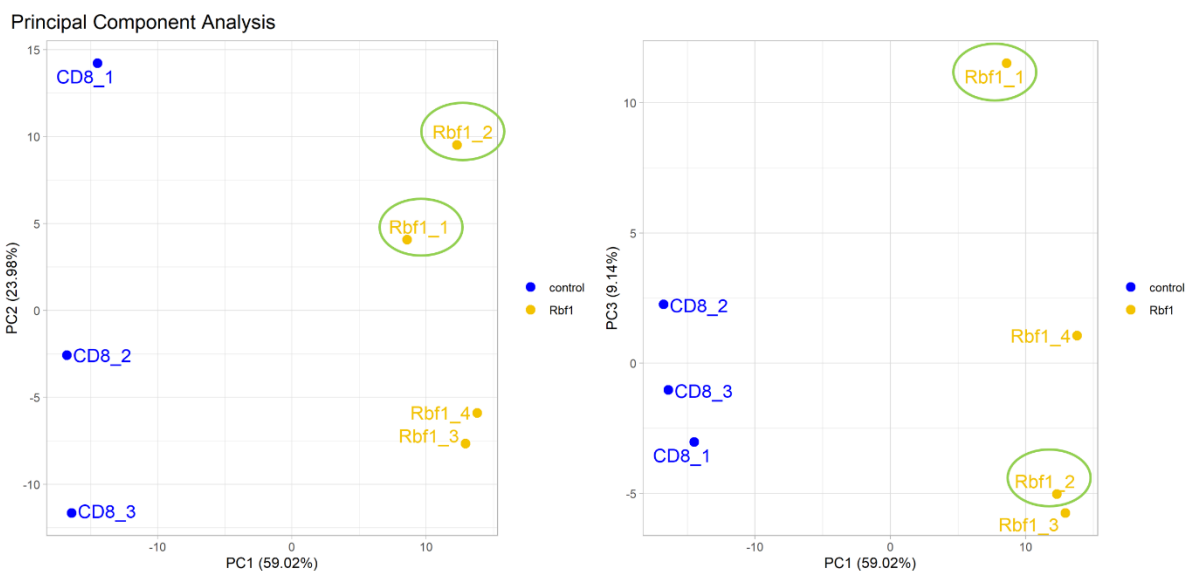
La description des données brutes inclut par exemple le nombre de reads associés à un gène pour chaque réplicat de chaque condition ou la description statistique des comptages bruts pour chaque réplicat (minimum, médiane, moyenne, ...). On y trouve également le total des comptages par réplicat et le pourcentage de gènes sur lesquels aucun read n'a pu être aligné.

La variabilité au sein d'une expérience de RNA-seq est normalement majoritairement issue des différences biologiques entre les échantillons. Pour vérifier cela, deux analyses sont réalisées.

**A**



**B**



**Figure 50 : Evaluation de la variabilité dans les échantillons séquencés.** Les échantillons entourés en vert sont ceux qui ne sont pas biologiquement indépendants. (A) Dendrogramme des clusters d'échantillons. On observe bien deux clusters différents pour les conditions contrôle (CD8) et les conditions de surexpression de Rbf1 (Rbf1). (B) Analyse en Composantes Principales. On observe également deux clusters pour les conditions Rbf1 (selon PC1 et PC3) et un pour CD8 (selon PC1). Les graphiques (A) et (B) sont générés par SARTools R, pour l'analyse dans laquelle les duplicats de PCR ont été retirés.

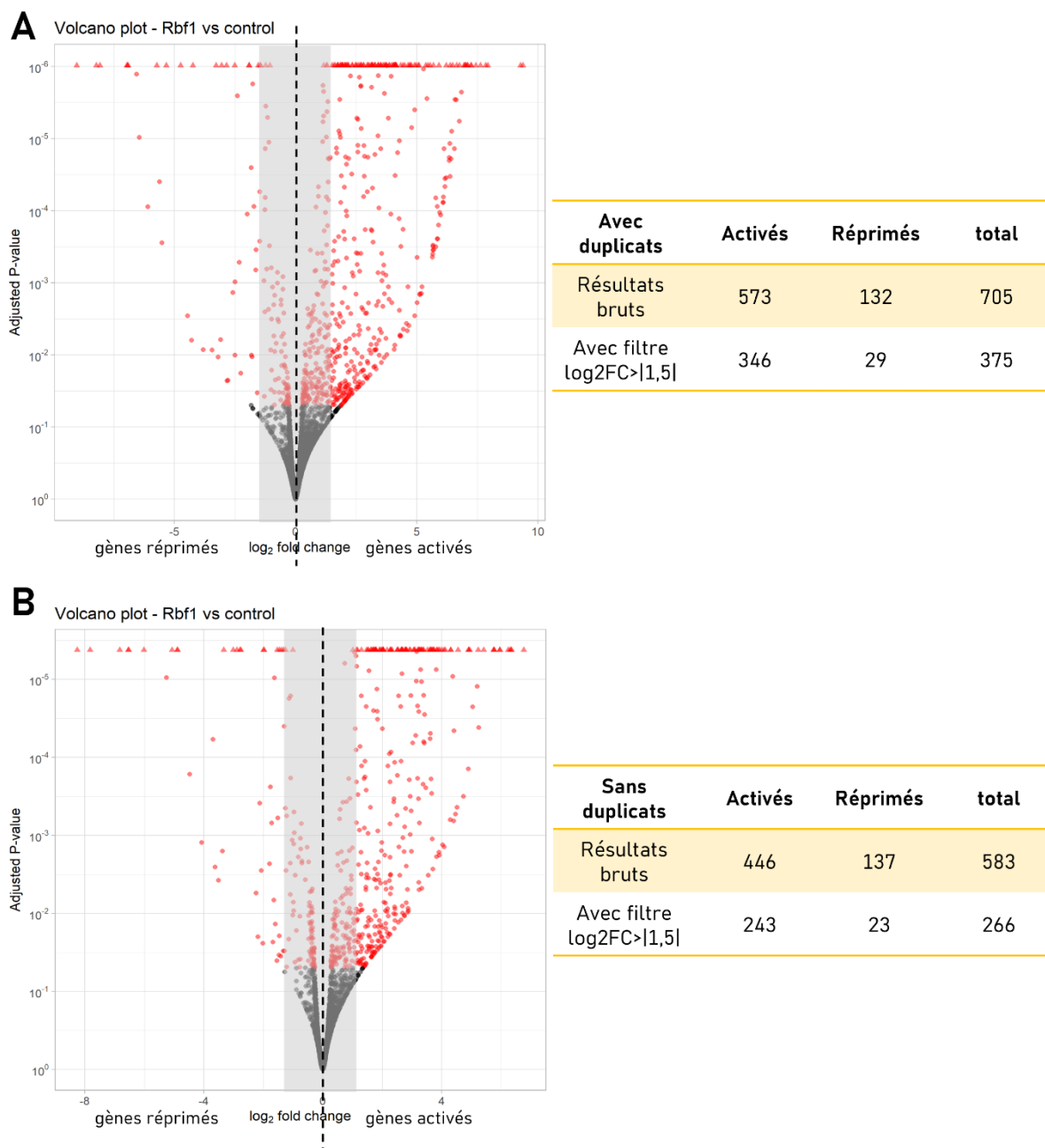
La première est une analyse de la variance qui permet d'obtenir un dendrogramme qui relie les échantillons à ceux ayant la variance la plus proche, créant ainsi des clusters d'échantillons. Dans notre cas, il y a bien formation d'un cluster d'échantillons CD8 et d'un cluster d'échantillons Rbf1 lorsque les duplicats sont retirés. Comme attendu, les deux échantillons qui n'étaient pas indépendants, Rbf1\_1 et Rbf1\_2, sont les plus proches (Figure 50A). La variabilité la plus importante semble donc bien être la variabilité biologique.

La seconde est une analyse en composantes principales. Ce type d'analyse permet de visualiser la répartition des données selon des variables, appelées « composantes principales ». Dans notre cas, on peut observer une bonne séparation des deux conditions selon la composante PC1, qui permet de représenter 59,02 % de la variabilité. (Figure 50B). Selon cette composante, la variabilité de notre échantillon permet de définir deux groupes correspondant aux deux conditions, la variabilité semble donc bien biologique. Selon les deux autres composantes, il n'est pas possible de définir des groupes, la variabilité décrite par ces composantes ne semble pas dépendre de caractéristiques de nos conditions (et en particulier les caractéristiques biologiques). S'il aurait été préférable de pouvoir observer la variabilité biologique selon plusieurs composantes, PC1 permet de représenter presque 60 % de la variabilité qui permet de différencier nos échantillons, ce qui est satisfaisant.

### b. Analyse différentielle

L'analyse différentielle considère un gène comme activé si le  $\log_2\text{FoldChange}$  ( $\log_2\text{FC}$ ), qui représente l'expression différentielle entre le contrôle et la condition test, est supérieur à 0 et la p-value inférieure à 0,05. A l'inverse, un gène considéré réprimé est associé à  $\log_2\text{FC}<0$  et  $p\text{-value}<0,05$ . De façon plus concrète, un  $\log_2\text{FC}=3$  signifie que l'expression du gène associé est augmentée  $2^3$  fois, soit huit fois (800 %) dans les conditions d'intérêt par rapport aux conditions contrôle. De la même manière, un  $\log_2\text{FC}=-3$  signifie que l'expression du gène associé est diminuée  $2^{-3}$  fois, soit 0,125 fois (12,5 %) dans les conditions d'intérêt par rapport aux conditions contrôles.

Le programme SARTools génère entre autres trois fichiers contenant respectivement l'ensemble des gènes quel que soit le nombre de reads associés, l'ensemble des gènes considérés comme activés (ayant un  $\log_2\text{FoldChange}>0$  et une  $p\text{-value}<0,05$ ) et l'ensemble des gènes considérés comme réprimés (ayant un  $\log_2\text{FoldChange}<0$  et une  $p\text{-value}<0,05$ ). Ces tables contiennent, pour les gènes listés, de nombreuses informations notamment le nombre de reads associés à chacun des réplicats, le  $\log_2\text{FC}$ , la p-value ajustée. Le croisement des tables des gènes activés et réprimés avec la base de données bioMart sur R a permis de récupérer des données supplémentaires sur ces gènes, comme le nom du gène ou le chromosome sur lequel il est situé.



**Figure 51 : Analyse différentielle.** En (A) sont représentés les résultats de l'analyse sans retrait des duplicats et en (B) les résultats avec retrait des duplicats. Sur les volcano plots, chaque point est la représentation de l'expression différentielle (p-value ajustée et Log2FC) pour chaque gène. Les points rouges représentent les gènes dont l'expression différentielle est considérée significative (p-value < 0,05). Les points représentés sous forme de triangle ont une p-value associée inférieure à  $10^{-3}$  (0,001). Les gènes dont le Log2FC est supérieur à 0 sont considérés comme activés, ceux dont le Log2FC est inférieur à 0 sont considérés comme réprimés par le programme SARTools qui a généré ces volcano plots. Nous considérons que les gènes activés sont ceux pour lesquels  $\log_2FC \geq 1,5$  (et réprimés  $\leq -1,5$ ). Les points se situant dans le rectangle grisé sont exclus de ces conditions. Le nombre de gènes avant et après ce filtre sont indiqués dans les tableaux correspondants.

L'analyse différentielle a, dans notre cas, identifié 573 gènes activés et 132 gènes réprimés par Rbf1 lorsque les duplicats de PCR n'ont pas été retirés ; et 446 activés et 137 réprimés lorsqu'ils l'ont été. Nous avons limité notre liste aux gènes présentant une valeur absolue de  $\log_2FC > 1,5$ . L'analyse conservant les duplicats de PCR identifie 346 gènes activés (Annexe III) et 29 réprimés (Annexe IV) ;

celle réalisée en éliminant les duplicats de PCR identifie 242 gènes activés (Annexe V) et 23 réprimés (Annexe VI) ; soit respectivement 375 et 265 gènes régulés par *Rbf1* (Figure 51). 238 gènes (217 activés et 21 réprimés) sont identifiés dans les deux analyses. Ces résultats ne correspondent pas à ce que nous attendions. En effet, des études ont pu évaluer la fixation de *Rbf1* à environ 2000 gènes chez la drosophile [112,113]. S'il est possible que cette fixation n'entraîne pas toujours de régulation des gènes comme c'est le cas pour le complexe dREAM [202], nous identifions ici à peine 300 gènes, ce qui paraît très peu. De plus, *Rbf1* est majoritairement décrit comme un répresseur transcriptionnel, mais nous identifions surtout des gènes activés. Cependant, ces résultats ne sont pas surprenants compte tenu du faible nombre de reads que nous pouvions utiliser et de la faible couverture qui en résulte.

Si l'on s'intéresse à l'identité des gènes dérégulés, on observe bien la surexpression de *Rbf1* (Figure 52A), en revanche, aucune des cibles connues de *Rbf1* que nous utilisons habituellement comme contrôles n'a pu être trouvée parmi les gènes identifiés. La visualisation des données dans IGV (Figure 52), a semé le doute quant à la validité des gènes cibles identifiés dans ces analyses. En effet, nous avons observé :

- Pour 90 gènes (~37%) identifiés comme activés, la couverture correspond à la structure du gène et l'activation est visible (Figure 52B).
- Pour 21 des gènes identifiés comme activés en réponse à la surexpression de *Rbf1*, la couverture ne correspond pas à la structure du gène dans la condition « activation » et pour ainsi dire aucun reads n'est présent dans la condition contrôle : ces gènes ne sont vraisemblablement pas activés mais il s'agit simplement d'un biais induit par le trop faible nombre de reads (Figure 52C).
- Pour 122 gènes identifiés comme activés par *Rbf1*, là encore il s'agit a priori d'un biais lié au faible nombre de reads, la couverture s'avérant même insuffisante pour être comparée à la structure du gène (Figure 52D).

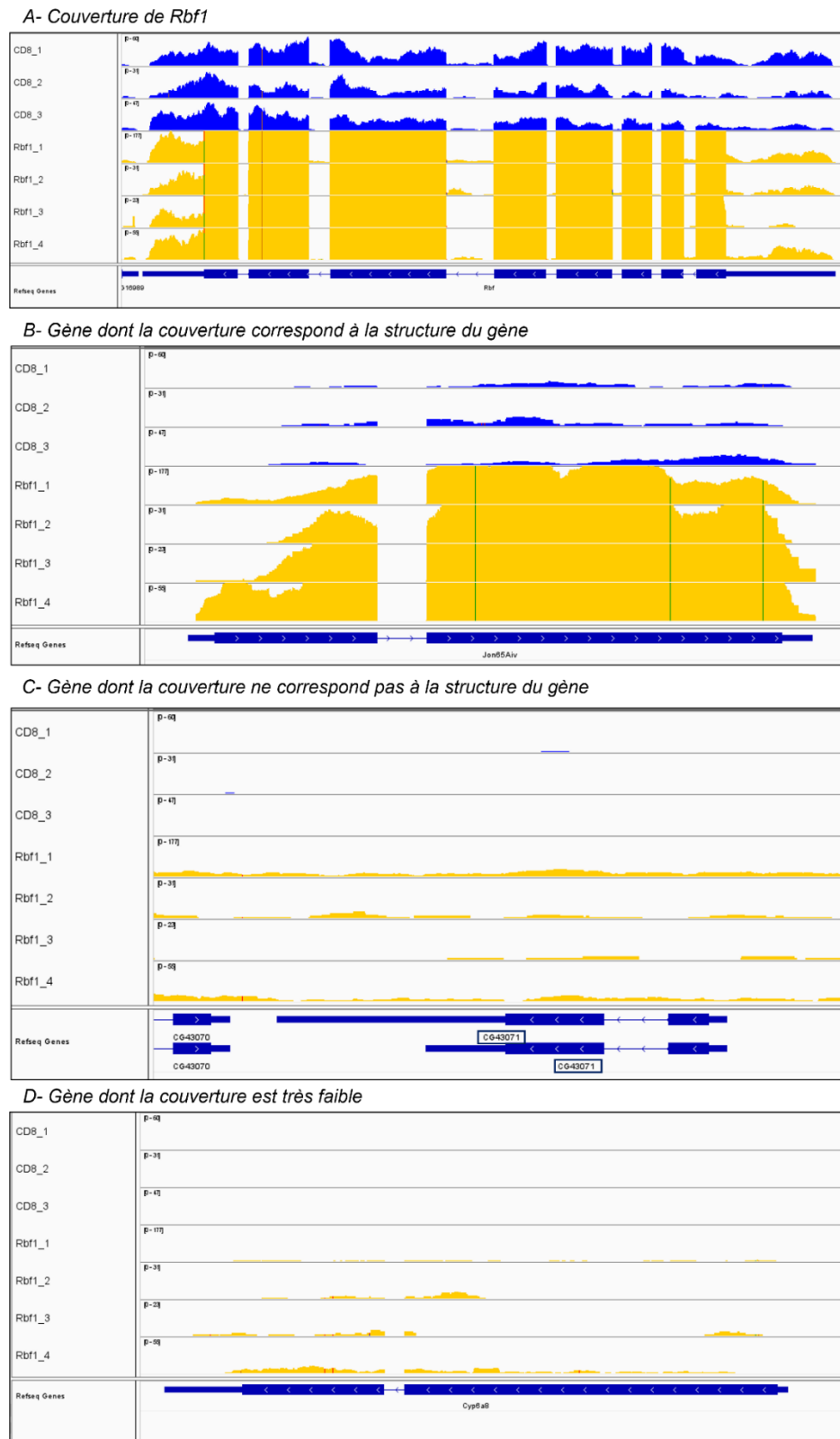
Les gènes identifiés comme réprimés peuvent également être répartis dans ces trois groupes. Concernant les cibles connues de *Rbf1* n'apparaissant pas comme différentiellement exprimées, *buffy* et *how* présentent une couverture très insuffisante. *PCNA* a quant à lui une couverture correspondant à la structure du gène mais sans différence entre les conditions contrôles et la surexpression de *Rbf1*. Les locus correspondant à dix gènes activés et quatre gènes réprimés n'ont pas été trouvés sur IGV et n'ont donc pas pu être analysés.

La faible proportion des gènes dont la couverture correspond à la structure des gènes et l'incohérence de nos données par rapport aux données connues nous pousse à arrêter ici l'analyse du RNA-seq. En effet, même si *Rbf1* est susceptible d'être un activateur transcriptionnel, il est quand même majoritairement considéré comme un répresseur et nous nous attendions à identifier une majorité de gènes réprimés par *Rbf1* et non l'inverse. De plus, aucune des cibles connues de *Rbf1*

n'est identifiée dans notre analyse. Il apparaît que le taux de duplication induise un nombre de reads obtenus qui s'avère insuffisant pour réaliser une analyse robuste. Dans la mesure où le seul échantillon ayant une couverture convenable est un échantillon en contexte de surexpression de Rbf1, cela est sans doute le biais majoritaire de notre expérience et est sans doute la raison pour laquelle nous obtenons autant de gènes activés en réponse à Rbf1.

Les problèmes techniques rencontrés par la plateforme de séquençage ont fait que nos données ne sont pas assez fournies pour proposer des résultats fiables. Un de nos échantillons, Rbf1\_1, possède quant à lui une couverture suffisante. S'il n'est pas suffisant pour porter à lui seul des données exploitables, nous aurions souhaité les croiser avec d'autres données de transcriptomique générées dans le disque imaginal d'aile en condition de surexpression de Rbf1. Nous aurions peut-être pu identifier quelques cibles communes potentiellement intéressantes. Malheureusement, nous n'avons pas pu récupérer ces données extérieures. Ce problème de couverture nous contraint donc à ne pas poursuivre cette analyse du RNA-seq. Elle aura toutefois été formatrice dans le traitement d'approches bioinformatiques.





**Figure 52 : La couverture des reads sur les gènes ne correspond pas toujours à leur organisation.** L'étendue de la visualisation a été normalisée proportionnellement au nombre de reads uniques alignés et dont les duplicats ont été retirés pour chaque réplicat. Des couvertures représentatives de chaque catégorie de gène (B, C, D) ont été sélectionnées parmi les gènes considérés comme activés avec ou sans retrait des duplicats. Dans les cas où plusieurs gènes sont représentés, le nom du gène d'intérêt est encadré en bleu. Ces figures sont issues de captures d'écran d'IGV. Rbf (=Rbf1) :  $p = 2,38 \cdot 10^{-73}$  ;  $\text{Log}_2\text{FC} = 3,885$  / Jon65Aiv :  $p = 8,93 \cdot 10^{-13}$  ;  $\text{Log}_2\text{FC} = 3,651$  / CG43071 :  $p = 0,0130$  ;  $\text{Log}_2\text{FC} = 2,537$  / Cyp6a8 :  $p = 5,3371 \cdot 10^{-4}$  ;  $\text{Log}_2\text{FC} = 4,464$

# DISCUSSION

Les régulations transcriptionnelles médiées par Rbfl sont multiples et jouent un rôle essentiel dans de nombreux processus cellulaires, le plus étudié étant la régulation de la transition G1/S dans le cycle cellulaire. Le laboratoire s'intéresse à l'apoptose et a produit plusieurs publications qui ont permis d'élucider en partie le rôle de Rbfl dans ce processus. Celui-ci est particulièrement complexe car selon l'état prolifératif des tissus dans lequel il est exprimé, *rbfl* peut être pro- ou anti-apoptotique. En effet, dans les cellules en prolifération comme celles du disque imaginal d'aile, la surexpression de *rbfl* entraîne de l'apoptose. Au contraire, dans les cellules différenciées ou non prolifératives comme les neurones ou le disque imaginal d'œil, la surexpression de *rbfl* n'est non seulement pas pro-apoptotique, mais supprime partiellement l'apoptose induite par des facteurs pro-apoptotique tels que p53 ou le membre de la famille Bcl-2 *Debcl*.

Si certaines des régulations transcriptionnelles médiées par Rbfl dans l'induction de l'apoptose ont été élucidées par le laboratoire, l'ensemble des gènes régulés par Rbfl n'est pas connu, en particulier les cibles impliquées dans son rôle anti-apoptotique. De plus, les facteurs de transcription de la famille E2F sont les partenaires privilégiés de Rbfl, mais ce ne sont pas les seuls. En effet, l'action de Rbfl peut être liée à son inclusion dans le complexe de régulation transcriptionnelle dREAM. Ce complexe et sa conservation chez d'autres organismes modèle dont les mammifères font l'objet d'une revue qui sera prochainement soumise à un journal.

Ce projet de thèse proposait une approche globale qui devait permettre d'identifier l'ensemble des cibles directes de Rbfl dans le disque imaginal d'aile, dans lequel Rbfl est pro-apoptotique, par une approche combinée de ChIP-seq et RNA-seq. Rbfl, tout comme le complexe dREAM, est majoritairement caractérisé comme un répresseur transcriptionnel, mais peut également activer l'expression des gènes. Si une expérience de ChIP-seq seule nous aurait permis de déterminer les promoteurs sur lesquels se fixe Rbfl, elle n'aurait pas suffi à caractériser la régulation des gènes associés. C'est pourquoi nous souhaitons combiner cette approche à une expérience de RNA-seq, qui permet de renseigner sur l'activation ou la répression de l'expression des gènes.

Bien que ces deux approches n'aient pas produit les résultats escomptés, la mise au point de ces expériences dans le disque imaginal d'aile m'a donné l'occasion de développer une technique d'isolement en masse de ces organes. Cet outil permet à un expérimentateur seul de récupérer de grandes quantités de disques imaginaires sans avoir recours à la dissection. Ce protocole a été récemment publié dans le journal *Biology (MDPI)*. Par ailleurs, bien que les résultats des approches ChIP-seq et RNA seq n'aient pas pu être valorisés, j'ai pu me former aux approches bioinformatiques pour analyser ces données, d'abord en utilisant la plateforme Galaxy puis en apprenant à réaliser ces analyses en lignes de commandes. Compte tenu de l'utilisation de plus en plus importantes de ces approches, cette compétence me paraît importante et valorisable dans la suite de mon parcours.

Enfin, les outils que j'ai mis en place (protocole d'isolement de disques imaginaux, crible de suppression de la létalité pupale) faciliteront de nouvelles analyses génomiques ultérieures au-delà du travail réalisé ici.

## I- Un protocole pour s'affranchir de la dissection

Les approches telles que le ChIP-seq nécessitent de grandes quantités de matériel biologique. Pour s'affranchir de la dissection et des problèmes de variabilité qu'elle pourrait poser, nous avons mis au point un protocole d'isolement en masse de disques imaginaux d'aile en adaptant des protocoles des années 1960-1970. Cette méthode permet d'obtenir des fractions enrichies en disques imaginaux d'ailes, avec un rendement de 11,5 %. Si ce rendement paraît faible par rapport au rendement de la dissection qu'on peut considérer autour de 85 %, ce protocole présente plusieurs avantages.

Premièrement, il est réalisable par un seul expérimentateur. Lorsque de grandes quantités de disques imaginaux, comme c'est le cas pour nos expériences, il est courant que plusieurs chercheurs participent à la dissection. Chaque participant ayant sa manière de disséquer, cela induit de la variabilité dans le traitement des disques. Grâce à ce protocole semi-automatisé, un seul expérimentateur est nécessaire et les résultats ne sont plus impactés par cette variabilité.

Deuxièmement, la durée de ce protocole est peu impactée par la quantité de larves utilisée comme matériel de départ. Au contraire, le temps de dissection est plutôt proportionnel à la quantité de larves à disséquer. Or, plus la dissection est longue, plus la qualité des disques se trouve impactée. Ici, la durée du protocole ne devrait pas excéder deux à trois heures peu importe le nombre de larves utilisées.

Enfin, lors de la dissection, les différents génotypes sont en général traités les uns à la suite des autres, le traitement n'est donc pas homogène. Dans cette technique, différents génotypes peuvent être traités en parallèle.

Bien que le rendement puisse sembler peu avantageux, ce protocole permet donc d'isoler de nombreux disques imaginaux en quelques heures, en s'affranchissant des principaux défauts de la dissection manuelle.

Les principaux contaminants des fractions récupérées sont les glandes salivaires. Celles-ci étant constituées de cellules bien plus grosses que celles des disques imaginaux d'aile, il est facile de récupérer uniquement les cellules d'intérêt en filtrant après dissociation.

Nous avons également évalué la qualité de ces disques en contrôlant le niveau d'apoptose induit dans ces cellules et en utilisant un test de viabilité cellulaire, qui ont tous deux permis de déterminer que cette technique permet d'obtenir des disques de qualité équivalente à ceux obtenus par dissection. Nous étions donc confiantes quant à la possibilité d'utiliser ces disques pour nos approches de ChIP-seq et RNA-seq. Elle a également été utilisée par une autre équipe pour la récupération de disques

imaginaires d'aile pour la réalisation d'un peptidome dans le cadre de l'étude de la spondylarthrite ankylosante. Ces résultats ne sont pas encore publiés.

## II- ChIP-seq : des pistes d'amélioration

### 1. Un problème récurrent de quantité d'ADN

La première des étapes a été de sélectionner le mode d'immunoprécipitation de la protéine d'intérêt, Rbfl. Peu d'anticorps anti-Rbfl sont disponibles. Nous avons donc choisi de réaliser nos expériences de ChIP-seq en recherchant les sites de fixation d'une forme étiquetée de Rbfl. Au laboratoire, des souches permettant l'expression de formes de Rbfl étiquetées par 3 étiquettes HA à l'extrémité amino- ou carboxy-terminale étaient disponibles au laboratoire. Pour l'ensemble de ces souches, l'expression de *rbfl-HA* est sous contrôle de séquences UAS et le transgène qui porte la construction a été inséré dans le génome de façon aléatoire, par transposition d'un élément P. Parmi les lignées disponibles, nous en avons sélectionné une dont le niveau d'expression est semblable à celui de la protéine Rbfl non étiquetée, sur la base de sa capacité à induire des phénotypes d'encoches sur les ailes proches de ceux induits par Rbfl non HA.

Afin d'isoler les protéines Rbfl-HA, nous avons également testé différents types de billes d'immunoprécipitation pour choisir la résine agarose pré-couplée HA, qui semblait donner les meilleurs résultats (Figure 29).

Une fois tous les paramètres de la ChIP ajustés à nos conditions, nous avons souhaité vérifier si nous pouvions amplifier des cibles connues de Rbfl en ChIP-qPCR. En étudiant l'absence d'enrichissement de ces cibles, nous nous sommes aperçues que les quantités d'ADN de départ que nous utilisions étaient bien trop faibles par rapport aux données de la littérature. En effet, Loubière *et al.* [187] proposent 5 à 10 µg d'ADN comme matériel de départ pour une ChIP. Nos préparations de chromatine contenaient jusqu'à 300 fois moins d'ADN (15,5 ng) que la fourchette basse proposée par Loubière *et al.* Il est donc possible que la quantité de nos inputs soient trop faibles et que ce que nous détectons lors des ChIP-qPCR de contrôle ne soit que du bruit de fond, nos quantités étant trop faibles pour passer le seuil de détection.

L'absence d'amplification des cibles que nous observions pouvait être dues à des amorces de PCR non adaptées. En effet, les amorces utilisées s'étaient révélées efficaces sur des fragments d'ADN de 400 à 500 paires de bases lors d'une précédente étude réalisée au laboratoire. Dans les expériences que j'ai réalisées nous avons soniqué la chromatine de façon à obtenir des fragments de 200 à 300 paires de bases. Compte tenu de la taille des amplicons, proche de la taille de ces fragments, la probabilité d'hybridation des deux amorces est faible et pourrait expliquer nos difficultés d'amplification. Plus de deux ans de thèse s'étant déjà écoulés et compte tenu de cette hypothèse, en concertation avec mon

comité de suivi de thèse, nous avons choisi de réaliser une expérience test avec séquençage. Celle-ci avait pour objectif d'identifier des sites de liaison majeurs de Rbfl dans notre tissu d'intérêt et aurait dû nous permettre de redéfinir des amplicons pour contrôler l'efficacité de nos CHIP. La crise de la COVID-19 ayant frappé entre l'élaboration du projet avec la plateforme de séquençage et la génération des échantillons, nous avons dû restreindre le nombre d'échantillons à envoyer à un unique échantillon test. Nous avons donc généré un échantillon contenant 69 ng d'ADN en fin d'IP. Celui-ci a été dilué pour correspondre aux critères de la plateforme et atteindre une quantité d'ADN de 41,4 ng, bien supérieure aux attentes minimales de la plateforme (10 ng). Si nous avions pu séquencer d'autres échantillons, il aurait été également intéressant d'étudier un échantillon « input » (avant immunoprécipitation).

Malheureusement, à la suite des analyses réalisées sur Galaxy, la distribution des reads ne permettait pas de distinguer leur accumulation au niveau des promoteurs, appelée « pics », caractéristiques des données de CHIP-seq. En effet, Rbfl est un cofacteur de transcription et des expériences de CHIP-on-chip ou CHIP-seq réalisées sur adultes ou embryons de drosophile décrites dans la littérature permettent d'anticiper l'existence de pics dans les 2 kb autour du site d'initiation de la transcription [112,113]. Nous nous attendions donc à trouver des pics en amont des gènes cibles, au niveau des promoteurs, ou à proximité de sites consensus de liaison de E2F, le principal partenaire des protéines de la famille RB. Or, comme cela est montré dans la [Figure 38](#), la distribution du signal ne semble pas dessiner de pics, ni même montrer un enrichissement au niveau de séquences promotrices/régulatrices. Le signal semble distribué de façon relativement homogène/aléatoire tout le long du génome et semble plus compatible avec du bruit de fond qu'avec un signal spécifique.

La raison de cet échec n'est pas clairement déterminée. Nous pouvons supposer une nouvelle fois un problème de quantité d'ADN. La quantité d'ADN utilisée comme quantité initiale pour un échantillon d'immunoprécipitation a été dosée à 6 µg. A l'issue de l'IP, la fraction immunoprécipitée contenait 69 ng, soit 1,15 % de l'input. Lors du tri cellulaire utilisé pour récupérer les cellules du domaine *vestigial*, nous avons pu déterminer que ce domaine comprenait seulement 10 % environ des cellules du disque imaginal en contexte de surexpression de Rbfl. Par conséquent, seules 10 % des cellules possèdent la protéine Rbfl étiquetée HA, et donc 10 % de l'ADN total extrait est susceptible d'être fixé par HA-Rbfl. Sur les 6 µg d'ADN dosés dans l'input, seulement 10 %, soit 600 ng, représentent donc l'ADN « d'intérêt » et l'input réel. La quantité d'ADN de la fraction immunoprécipitée est donc de 69 ng pour une quantité initiale d'intérêt de 600 ng, donc un rendement de 11,5 %. Cette valeur semble trop élevée par rapport aux valeurs attendues pour une immunoprécipitation de chromatine. De plus, ces 6 µg sont associés à une préparation issue de 150 disques imaginaires d'aile. Cela représente 40 ng d'ADN par disque d'aile, soit quatre fois plus que la valeur indiquée dans Loubière *et al.* [187] que nous utilisons comme référence.

Plusieurs hypothèses peuvent alors être émises concernant ces valeurs très élevées et l'échec de cette expérience de ChIP-seq :

- *Le crosslink a été trop long.* Une fixation trop longue induirait la précipitation par entraînement d'une grande partie du matériel.
- *Les lavages sont insuffisants.* Notre immunoprécipitation se faisant avec une résine et non des billes magnétiques, les lavages sont possiblement moins efficaces. Il se pourrait alors qu'il reste des fragments d'ADN liés de façon aspécifique.
- *L'immunoprécipitation de l'étiquette HA ne s'est pas faite correctement.* Cette hypothèse semble peu probable. En effet ; des tests préalables d'immunoprécipitation suivie de Western Blot avaient été réalisés et ont montré un enrichissement des étiquettes HA dans les immunoprécipitations anti-HA par rapport aux immunoprécipitations anti-IgG, ce qui est encourageant quant aux capacités d'immunoprécipitation de l'étiquette HA.
- *Le dosage de l'ADN surestime la quantité réelle.* Dans le cadre de cette hypothèse, il y aurait moins d'ADN dans les fractions de départ et les fractions immunoprécipitées que ce que nous avons estimé. Ainsi, le matériel de départ ne serait pas en quantité suffisante pour discriminer les résultats spécifiques de la surexpression de *Rbf1* et le bruit de fond. Si ce cas de figure s'est présenté pour les dosages ARN, ici le réactif utilisé n'est pas le même et la plateforme de séquençage ne nous a pas alerté d'une quelconque incohérence entre les quantités annoncées et celles mesurées ; ou d'une quantité inférieure au seuil demandé. Cette hypothèse paraît donc également peu probable.
- *L'étiquette HA perturbe les interactions protéine-protéine.* *Rbf1* est un co-facteur de transcription et ne se lie donc pas directement à l'ADN. L'interaction se fait par l'intermédiaire d'autres protéines, comme des facteurs de transcription. Pour faciliter les expériences de ChIP, nous utilisons une protéine *Rbf1* étiquetée HA en N-terminal. Il se pourrait que la présence de cette étiquette modifie les interactions de *Rbf1* avec les facteurs de transcription partenaires. Toutefois, nous avons veillé à utiliser une souche dont les phénotypes d'ailes en conditions de surexpression de HA-*Rbf1* se rapprochent le plus possible de ceux induits par la surexpression de *Rbf1* non étiquetée. Cette induction de l'apoptose suggère que les interactions *Rbf1/dE2F2* sont maintenues, car tous deux sont nécessaires à la régulation de *buffy* et *how* et présents à leurs promoteurs [159].
- *La fixation de HA-*Rbf1* à la chromatine n'est pas suffisante.* Il se pourrait que ces interactions protéine-protéine, pour une raison inconnue, ne résistent pas au traitement subi par les échantillons. Nous avons testé l'impact de la fixation et nous avons pu voir que sans fixation, aucun enrichissement n'a pu avoir lieu en qPCR. Il y a donc bien eu fixation, mais nous n'avons aucune indication sur la qualité de celle-ci.

A ce jour, plusieurs de ces hypothèses, voire leur combinaison, semblent probables pour expliquer les difficultés que nous avons rencontrées. Premièrement, il est possible que la mauvaise fixation de HA-Rbfl sur la chromatine ait entraîné la précipitation aspécifique de fragments d'ADN ou que l'étiquette ait été masquée par un crosslink trop long. Deuxièmement, la quantité d'ADN trop faible utilisée au départ est certainement un frein au bon déroulement de l'expérience. Enfin, il se pourrait que les lavages, trop peu nombreux ou pas assez forts, n'aient pas permis de débarrasser les billes de la fixation aspécifique. Il serait alors intéressant de tester différents temps de crosslink et de réaliser des expériences de Western Blot sur les fractions input et IP afin de vérifier la présence de Rbfl dans ces fractions, en plus des qPCR sur les cibles connues. Ces expériences devraient bien sûr être réalisées avec une quantité d'ADN de départ plus élevée.

## 2. La ChIP pour étudier le complexe dREAM

En plus de l'identification des cibles directes de Rbfl grâce à la combinaison avec le RNA-seq, cette expérience de ChIP-seq aurait pu nous informer sur la présence d'autres sous-unités du complexe dREAM aux promoteurs cibles. En effet, les facteurs de transcription ciblent des sites de fixation consensus spécifiques. Par exemple, le motif consensus de fixation de E2F est TTGGCGC chez la drosophile et TTTSGCGC chez l'humain. Grâce à des outils de bio-informatique, il aurait été intéressant de déterminer si des motifs redondants étaient associés à la fixation de Rbfl. Par combinaison avec les données du RNA-seq, une éventuelle corrélation entre les sites de fixation présents et le sens de régulation de la transcription du gène associé (activation ou répression) aurait pu être étudiée. La présence effective de ces facteurs de transcription aurait pu être vérifiée par ChIP avec immunoprécipitation des facteurs d'intérêt et qPCR des gènes cibles.

## III- RNA-seq : de la purification d'ARN à la couverture des gènes

### 1. Des difficultés d'obtention des ARN

Pour réaliser le RNA-seq, le cahier des charges de l'entreprise à laquelle nous avons confié le séquençage massif préconisait une quantité d'ARN totaux comprise entre 250 et 500 ng par échantillon. A l'issue des expériences de tri cellulaire, les ARN totaux ont été extraits à l'aide de TRIzol. Ils ont ensuite été quantifiés grâce à une méthode fluorimétrique (Quantus, Promega) plus fiable qu'une technique de spectrométrie comme le Nanodrop (Figure 34A). Cependant, les dosages étaient rarement en accord avec les quantités évaluées par la société qui a pris en charge les échantillons pour le séquençage massif. Il est arrivé que pour un des échantillons, la société détecte une concentration supérieure à celle évaluée au laboratoire. Cela s'explique sans doute par un culot



d'ARN encore non totalement solubilisé lorsque j'ai réalisé la mesure. En revanche, pour la majorité des échantillons envoyés, les quantités mesurées à l'arrivée chez le prestataire en Suisse se sont avérées bien inférieures, de quelques pourcents à plus de 50 % (Figure 41).

Deux possibilités sont à envisager pour expliquer ce phénomène. La première concerne une déficience de la sonde utilisée pour doser les ARN au Quantus. Celle-ci pourrait surestimer la quantité d'ARN en fixant plusieurs sondes sur le même ARN. Il n'est pas envisageable que la surestimation vienne d'une erreur de dosage qui considérerait l'ADN comme de l'ARN car d'une part, l'intérêt de la mesure fluorimétrique (Quantus) par rapport à la mesure de l'absorbance (Nanodrop) est l'utilisation de sondes spécifiques des différents acides nucléiques ; et d'autre part, les échantillons ont été traités à la DNase avant le dosage. La seconde concerne la dilution des échantillons. Avant l'envoi à la plateforme de séquençage, les échantillons sont dilués dans de l'eau RNase/DNase-free pour homogénéiser les quantités entre réplicats et/ou adapter les volumes conformément au cahier des charges établi par la plateforme. Si un nouvel aliquot a été utilisé pour chaque lot de purifications, il est possible que le stock soit contaminé et qu'il entraîne une dégradation partielle des échantillons. Enfin, une dégradation durant le transport est envisageable, même si celui-ci a été fait dans les meilleures conditions possibles (glace carbonique, camion réfrigéré, envoi express).

Trois lots d'échantillons ont été envoyés à la société de séquençage, chacun nécessitant de longs mois de préparation des échantillons. Bien que nos échantillons ne satisfassent pas toutes les exigences de la plateforme, ils étaient de bonne qualité et les quantités détectées pas si faibles pour interdire totalement leur usage pour la génération de banques. La récupération du matériel biologique prenant par ailleurs trop de temps pour chercher à obtenir l'échantillon « parfait », nous avons décidé, sur les conseils de la plateforme de séquençage, de réaliser les banques et le séquençage sur quatre des échantillons en contexte de surexpression de *rbf1* (sur six envoyés) et trois échantillons contrôles CD8 (sur cinq). Parmi ces sept échantillons, deux des *Rbf1* ne sont pas biologiquement indépendants et trois contiennent moins de 250 ng d'ARN, quantité recommandée par la plateforme (Figure 41, échantillons en vert).

### 2. Des taux de duplication trop élevés

Par malchance, un « problème technique important » lors de la préparation des banques a nécessité de répéter cette étape et, compte tenu de la faible quantité de matériel restant à l'issue de la première préparation, de réaliser un nombre élevé de cycles de PCR (17), augmentant fortement le risque d'obtenir des duplicats de PCR. En effet, si le nombre de cycles de PCR augmente mais que tout le matériel est déjà amplifié (en particulier lorsque la quantité est faible), toute amplification supplémentaire sera en réalité un duplicat. L'entreprise nous a alors proposé de réaliser un séquençage test pour nous assurer que le taux de duplication dans ces banques était acceptable. A l'issue de ce séquençage, les taux de duplications observés, compris entre 0,36 et 2,73 %, étaient tout à fait

convenables (Figure 42) et nous avons donc procédé au séquençage des banques. Malheureusement, lors du contrôle qualité des données du séquençage plus profond, les taux de duplication se sont finalement avérés bien trop élevés (Figure 44). Ce phénomène est dû à l'augmentation de la profondeur de séquençage. En effet, la quantité de matériel de départ étant faible, sa diversité l'est aussi et le séquençage test (1 million de reads) était apparemment suffisant pour rendre compte de l'essentiel de la diversité de nos ARN, ce qui est suprenant pour un génome tel que celui de la drosophile qui contient plus de 13 000 gènes. Dans ce cas, chaque nouveau read séquencé au-delà du million a donc de fortes chances d'être un duplicat de PCR et non un read unique et un séquençage de 15 millions de reads implique donc le séquençage essentiellement des duplicats de PCR.

Pour réaliser l'analyse, la question du retrait ou non des duplicats de PCR s'est alors posée. Il ne semble pas y avoir de consensus clair sur la question, mais il est parfois recommandé de ne pas les retirer pour ne pas induire de biais dans le traitement des données [198]. Le programme *fastp* considère deux reads comme étant des duplicats si leur séquence est exactement la même. Le programme *picard*, qui identifie et retire les duplicats, considère deux paires de reads (dans le cas d'un séquençage paired-end) comme étant des duplicats s'ils sont alignés au même endroit dans le génome et que les cinq premiers nucléotides de chaque membre de la paire se correspondent respectivement avec un taux de mésappariement de moins de 0,03 %. Cette comparaison ne prend donc pas en compte les 70 paires de bases suivant ces cinq premiers nucléotides, un polymorphisme présent dans cette séquence ne serait donc pas détecté et les reads considérés comme des duplicats alors qu'ils ne le sont pas. Cependant, dans notre cas, ces taux de duplicats sont très élevés et pourraient, lors de l'analyse de l'expression différentielle, induire de faux positifs si de nombreux duplicats correspondent aux mêmes gènes : cela induirait la confusion entre un enrichissement dû à un nombre de duplicats élevés et un enrichissement attribuable à un effet biologique.

Si, à première vue, l'analyse la plus pertinente serait celle après soustraction des duplicats, leur retrait pose le problème suivant : si 70 % des reads sont des duplicats à retirer, alors seulement 30 % de reads uniques restent pour l'analyse, ce qui n'est en général pas suffisant pour avoir une profondeur adaptée à des analyses d'expression différentielle. Dans notre cas, cela représente entre 14 et 5 millions de reads restants (Figure 46). Or, les plateformes de séquençage que nous avons consultées au sujet du RNA-seq nous avaient conseillé 15 à 30 millions de reads pour un génome moyen comme celui de la drosophile. En considérant un taux de duplication normal faible, cela représente au moins 15 à 20 millions de reads uniques. Une profondeur de séquençage trop faible réduit la puissance statistique de l'expérience, et peut conduire à une augmentation de faux positifs ou négatifs. De plus, une analyse dans ces conditions, plutôt que d'être représentative du transcriptome, ne correspond plus qu'à une fraction des transcrits les plus abondants. Au contraire, si les 70 % de duplicats ne sont pas retirés, il est possible que les duplicats, s'ils s'accumulent dans une condition donnée, simulent des régulations qui n'existent pas et conduisent à l'identification de faux positifs. Afin de prendre la

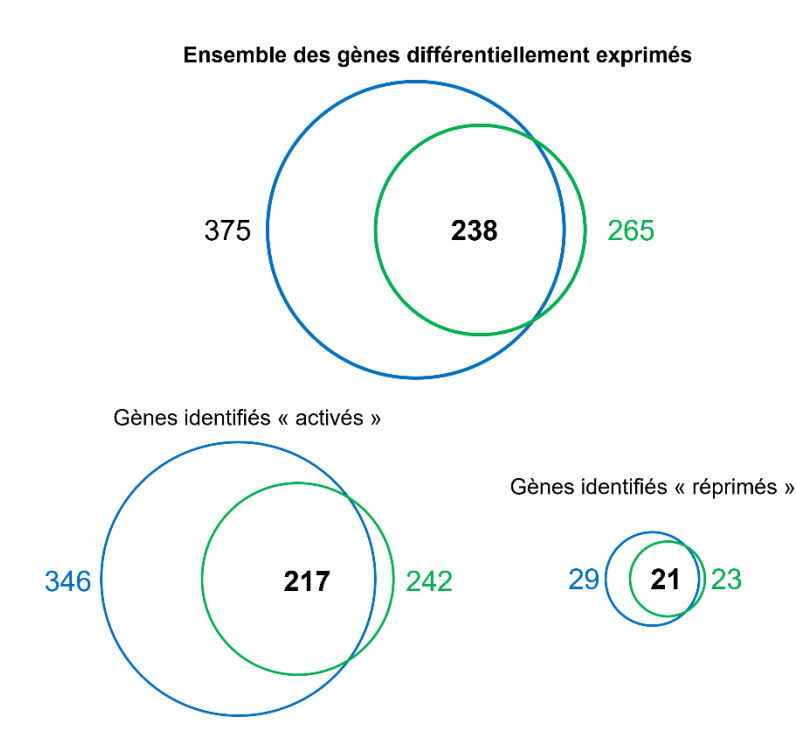
mesure de chacun de ces biais, nous avons donc choisi de réaliser deux analyses en parallèle : une où les duplicats ne seraient pas retirés, et une autre où ils le seraient après l'alignement. Le retrait des duplicats aurait également pu être réalisé avant l'alignement, mais dans notre cas, en l'absence d'ajout de « code-barres » d'identification des duplicats (UMI) lors de la création des banques, il était plus adapté de le réaliser après l'alignement.

### 3. Analyse bio-informatique des données de RNA-seq

L'analyse des données de ChIP-seq à l'aide de la plateforme Galaxy m'a permis d'appréhender facilement l'analyse de données NGS. En effet, elle permet l'utilisation des outils de bioinformatique de façon simple, en se concentrant sur le paramétrage des outils mais sans nécessité d'en connaître la syntaxe puisque celle-ci est automatiquement générée selon les paramètres indiqués. Cela m'a permis de me familiariser avec les outils disponibles, leur fonction et options et les différentes étapes de l'analyse de ce type de données. Cependant, un des inconvénients principaux de cette plateforme est le temps de traitement des différentes demandes (« queries ») soumises. Les serveurs étant en ligne et partagés par des chercheurs à l'échelle internationale, la charge du serveur influe beaucoup sur la durée du traitement. Si cet inconvénient n'était pas limitant lors de l'analyse du ChIP-seq qui ne comportait qu'un échantillon, il aurait pu être problématique pour le traitement du RNA-seq pour lequel sept échantillons étaient à traiter. J'ai également eu la chance d'interagir avec le Dr. Frédéric Crémazy (UMR 1173, UVSQ) et la plateforme de génomique de l'UVSQ, qui m'ont donné accès à leur serveur et des conseils pour une analyse bio-informatique plus directe.

L'accumulation de suffisamment de cellules pour générer les échantillons d'ARN a nécessité de longs mois, et cette durée a été prolongée par les multiples envois à la plateforme. J'ai pu profiter de ce temps pour approfondir mes connaissances en bio-informatique, notamment grâce à la formation « NGS & Cancer » proposée par le Cancéropôle Île-de-France qui m'a permis d'avoir de solides connaissances théoriques sur ce type d'analyses. Ainsi, j'ai réalisé l'analyse du RNA-seq sous Linux et R.

A l'issue de cette analyse, nous avons pu identifier respectivement 375 et 265 gènes dont la quantité de transcrits différait entre les conditions avec et sans conservation des duplicats (Figure 53). Parmi ces gènes, 238, soit la majorité, sont identifiés par les deux analyses. L'inspection visuelle sur IGV des loci identifiés exclusivement dans l'une ou l'autre de ces analyses nous montre une couverture extrêmement faible (Figure 52C) correspondant à du bruit de fond aléatoire. En effet, lorsqu'aucun read n'est disponible pour l'une des conditions, la présence de seulement quelques reads dans l'autre apparaît comme une expression différentielle. Pour l'essentiel, les gènes identifiés le sont dans les deux analyses et l'inclusion ou non des duplicats de PCR ne semble pas avoir une influence. Nous avons donc choisi de nous concentrer sur les 265 gènes identifiés avec élimination des duplicats de PCR.



**Figure 53 : Nombre de gènes identifiés comme différentiellement exprimés avec ou sans duplicats.** Le nombre de gènes avec duplicats de PCR (sans retrait des duplicats) est indiqué en bleu, celui sans duplicats de PCR (avec retrait des duplicats) en vert.

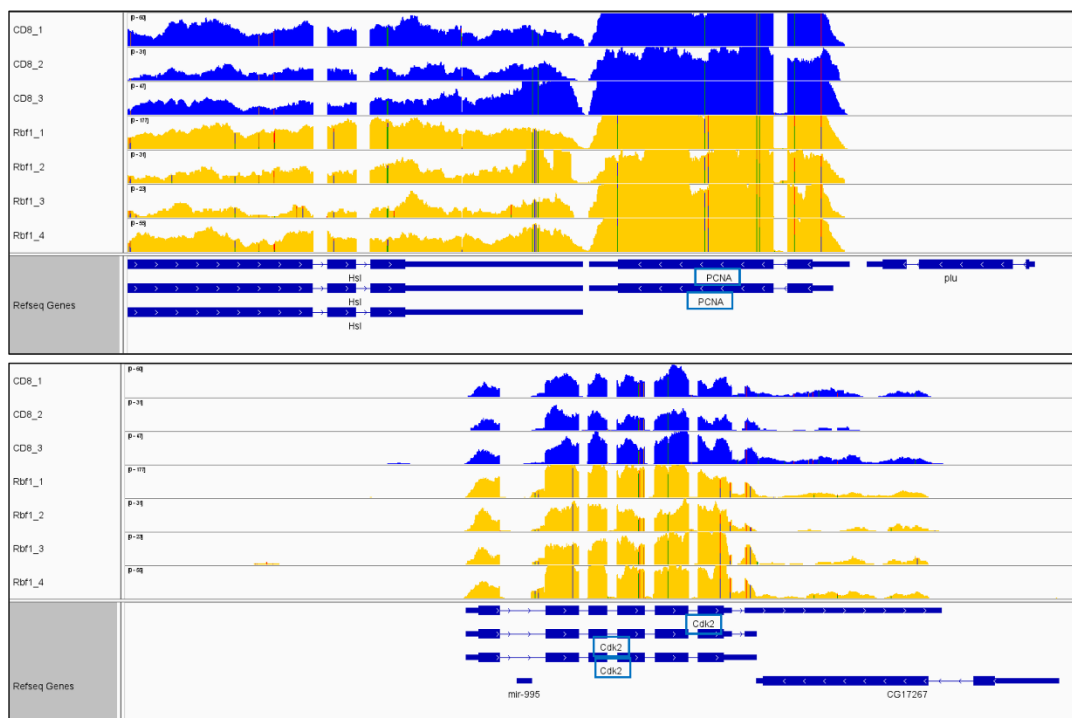
Grâce à la visualisation sur IGV, nous nous attendions à trouver une couverture correspondant aux ARNm, donc un signal détectable au niveau des exons mais pas des introns. Malheureusement, à peine plus d'un tiers des gènes identifiés présente une couverture cohérente avec l'organisation du gène. En effet, près de la moitié des gènes identifiés par l'analyse différentielle présentent une couverture très faible. Il nous semble plus raisonnable d'exclure ces gènes, car il est impossible de savoir si l'expression différentielle mise en avant est due à une réalité biologique ou simplement à un hasard de présence de ces ARN faiblement détectés.

Compte tenu de la faible proportion pour lesquels les résultats semblent cohérents, nous avons choisi d'arrêter là l'analyse de ces données. De plus, nous ne retrouvons aucun des gènes cibles connus de Rbf1, notamment *buffy*, *how* ou *PCNA*. La couverture des deux premiers est très faible, et la couverture du dernier ne présente pas de différence entre contrôle et Rbf1 (Figure 54). Nous n'avons donc aucune garantie quant à la fiabilité de nos données.

### Cibles de Rbf1 impliquées dans l'apoptose



### Cibles de Rbf1 impliquées dans le cycle cellulaire



**Figure 54 : Couverture des reads sur des gènes cibles connus de Rbf1.** L'étendue de la visualisation a été normalisée proportionnellement au nombre de reads uniques alignés et dont les duplicats ont été retirés pour chaque réplicat. Buffy et how sont des cibles impliquées dans l'apoptose, PCNA et Cdk2 dans le cycle cellulaire. Dans les cas où plusieurs gènes sont représentés, le nom du gène d'intérêt est encadré en bleu. Ces figures sont issues de captures d'écran d'IGV.

## IV- Perspectives

### 1. Identifier des gènes cibles

L'objectif du projet était d'identifier les cibles directes de Rbf1 par une approche combinée CHIP-seq et RNA-seq. Le CHIP-seq devait permettre l'identification des promoteurs liés par Rbf1. Cette expérience n'ayant pas été menée à son terme, nous avons prévu de croiser nos données de RNA-seq avec celles d'Acharya *et al.* [113], qui ont réalisé une expérience de CHIP-exo de Rbf1 dans des embryons de drosophile. Ces données n'ont pas été acquises dans le tissu que nous souhaitons étudier, mais elles auraient éventuellement permis de sélectionner des candidats d'intérêt dans la liste des gènes identifiés comme différentiellement exprimés par l'analyse du RNA-seq.

Ce type d'approche globale permet d'identifier plusieurs dizaines voire centaines de gènes. Pour valider voire caractériser ces cibles, il est nécessaire de réduire leur nombre. Nous avons choisi de prioriser les cibles intervenant dans les problématiques du laboratoire, à savoir l'apoptose et la mitochondrie. Pour les sélectionner, nous aurions souhaité réaliser une analyse de Gene Ontology (GO). Le projet Gene Ontology, a pour but d'associer chaque gène ou protéine à un ensemble de termes relatifs à sa ou ses fonctions [203]. Nous aurions pu utiliser ces termes pour limiter le nombre de gènes à analyser :

- Une approche sans *a priori*, en analysant quels termes sont les plus enrichis et ainsi déterminer les fonctions principales régulées par Rbf1 dans notre modèle
- Une approche orientée, en recherchant des termes associés à l'apoptose et à la mitochondrie pour déterminer quels gènes identifiés répondent à ces fonctions.

Si ce type d'analyse aurait permis de ne s'intéresser qu'à des cibles de Rbf1 impliquées dans des processus très enrichis ou dans des fonctions correspondant aux problématiques du laboratoire, elle aurait également pu induire un biais. En effet, les termes GO sont parfois annotés sur la base de séquences prédites comme pouvant être impliquées dans les processus indiqués. Ainsi, certaines cibles sélectionnées pourraient ne pas avoir de rôle dans les processus d'intérêt, ou au contraire, des cibles qui auraient pu être intéressantes ne seraient pas sélectionnées car pas encore annotées.

### 2. Valider ces cibles

Une fois les gènes d'intérêt sélectionnés parmi les centaines identifiés par l'approche combinée CHIP-seq/RNA-seq, il est nécessaire de valider le fait qu'ils sont i) bien des cibles, directes ou non, de Rbf1 et ii) impliqués dans le processus apoptotique régulé par Rbf1.

a. Vérifier si les gènes identifiés sont de réelles cibles de *Rbf1*

La validation des cibles directes aurait été faite par ChIP-qPCR. Comme les ChIP de contrôle que nous avons essayé de réaliser lors de la mise au point de l'expérience de ChIP-seq, un moyen de validation des cibles directes serait de réaliser une ChIP avec immunoprécipitation de *Rbf1* (grâce à l'étiquette HA) et de vérifier par qPCR si les cibles en question font bien partie des fragments d'ADN précipités. Le design des amorces de qPCR aurait de plus été facilité par les résultats de ChIP-seq : en connaissant la position des pics sur les promoteurs, il est possible de créer des amorces visant précisément le site de fixation de *Rbf1*. Sans ChIP-seq, nous aurions pu vérifier l'augmentation ou la diminution du niveau des transcrits des gènes cibles par qPCR en conditions de surexpression de *Rbf1* ou non.

Cette étape de validation des gènes cibles aurait également été l'occasion de s'intéresser de nouveau à l'aspect tissu-spécifique de ce projet de thèse. En effet, nous avons dû délaissier cette partie car l'ajustement de la méthode d'isolement en masse pour l'isolement de disques imaginaires d'yeux aurait été trop longue. Cependant, une fois les gènes cibles de *Rbf1* dans le disque imaginaire d'aile identifiés, il aurait été intéressant de rechercher s'ils étaient également cible dans le disque imaginaire d'œil ou les neurones par ChIP-qPCR. Le cas échéant, le sens de régulation (activation ou répression transcriptionnelle) aurait également pu être évalué par RT-qPCR. Une des limites de cette approche est que nous aurions uniquement identifié les cibles de *Rbf1* dans le disque imaginaire d'aile et observé leurs régulations dans les neurones. Certaines cibles dérégulées uniquement dans le disque d'œil n'auraient alors pas pu être identifiées car elles n'auraient pas été sélectionnées comme cibles d'intérêt dans le disque d'aile.

Les disques imaginaires d'œil sont plus délicats à disséquer que les disques imaginaires d'aile. De plus, leur taille moins importante nécessite d'en récupérer une plus grande quantité pour obtenir le matériel de départ nécessaire aux approches ChIP-seq et RNA-seq. Nous avons donc envisagé d'utiliser le cerveau larvaire comme modèle pour tester l'activité anti-apoptotique de *Rbf1*. Des expériences préliminaires de détection des caspases clivées en conditions de surexpression de *Rbf1*, *p53* et de co-surexpression avaient été réalisées, mais nous n'avons pas observé de marquage plus important en conditions « *p53* ». En absence de gènes candidats à valider, nous n'avons pas poursuivi la mise au point de ce test.

b. Évaluer l'implication des cibles dans le processus apoptotique

Après sélection des cibles d'intérêt, nous aurions souhaité valider leur implication dans l'apoptose. Pour cela, nous aurions utilisé deux outils, permettant d'évaluer l'implication de la cible dans les activités pro- et anti-apoptotiques de *Rbf1*. Ces outils permettent d'observer l'effet de mutants ou d'ARN interférents des cibles sur l'apoptose régulée par *Rbf1*. Grâce à leurs phénotypes facilement

observables, il est aisé de cribler un grand nombre d'individus pour obtenir des résultats avec une bonne puissance statistique. Si les résultats obtenus se sont révélés finalement inexploitable, nous avons toutefois réalisé la mise au point de ces outils, qui ne sera pas vaine puisqu'ils pourront être utilisés lors de la réalisation d'autres projets du laboratoire.

Au laboratoire, l'évaluation de l'effet pro-apoptotique de Rbf1 dans l'aile est possible grâce au phénotype d'encoche à la marge de l'aile. Le nombre d'encoche présentes est corrélé positivement à la quantité de cellules en apoptose dans le disque imaginal (Figure 24A). Nous utilisons cette échelle pour déterminer l'implication d'acteurs potentiels de la voie d'apoptose grâce à des ARN interférents ou des mutants. Si l'absence de la protéine entraîne l'apparition d'un phénotype de force différente (plus ou moins d'encoche), alors elle est le plus généralement impliquée, directement ou indirectement, dans le processus apoptotique médié par Rbf1.

Contrairement aux phénotypes d'aile, l'observation et l'analyse des phénotypes d'œil nécessitent de prendre en compte plusieurs paramètres comme la taille, l'organisation des ommatidies ou la rugosité. Ces phénotypes sont parfois difficiles à observer ou à quantifier et, la variété des types cellulaires présents dans le domaine GMR font que la force et la nature des phénotypes observés n'est pas corrélée au niveau d'apoptose. Nous n'avons donc pas utilisé cet outil pour tester l'effet des gènes candidats sur de l'activité anti-apoptotique de Rbf1. Nous avons cependant mis au point un crible se basant sur un autre phénotype associé à l'induction d'apoptose dans les neurones : la létalité pupale. En effet, nous avons observé que, lors de l'induction de *p53* à la fois dans l'ensemble des neurones et en postérieur du sillon morphogénétique grâce aux pilotes *elav* et *GMR*, peu d'individus naissent et les quelques rescapés présentent des phénotypes d'œil réduit et d'ailes non dépliées (Figure 27A). Nous supposons que cette létalité est due à l'atteinte des motoneurones dans les adultes. En effet, la dissection de pupes non écloses nous a permis d'observer une métamorphose complète sans problème apparent de développement. L'induction d'apoptose dans les neurones, et en particulier les neurones moteurs, provoquerait une perte de mobilité des adultes métamorphosés qui ne seraient alors pas en mesure de s'extirper des pupes. Cette hypothèse expliquerait également le fait que les ailes des quelques rescapés ne soient pas déployées, ce déploiement étant dépendant de muscles et donc de leur innervation.

En co-surexprimant les gènes HA-*rbf1* (permettant la production de la protéine étiquetée) et *p53* utilisé pour les expériences de CHIP-seq et de RNA-seq, nous avons pu supprimer partiellement la létalité pupale chez les mâles. En effet, nous observons un taux de survie significativement plus élevé chez les individus co-surexprimant HA-*rbf1* et *p53* que chez ceux surexprimant *p53* seul, dont une partie présente des phénotypes moins forts avec des ailes dépliées (Figure 28). De façon surprenante, l'effet que nous observons avec la protéine Rbf1 étiquetée en N-terminal est plus fort que celui observé avec la protéine étiquetée en C-terminal, à l'inverse des observations faites sur les phénotypes



d'aile (Figure 24). Cet effet est probablement dû à un niveau d'expression des transgènes différent. En effet, ces souches ont été générées par insertion aléatoire, le niveau d'expression de chacun des transgènes peut donc être différent, et également être modulé selon le tissu d'intérêt. Cela expliquerait les différences que nous observons entre la modulation des phénotypes pro- et anti-apoptotique par rapport à la protéine non-étiquetée. Nous avons ainsi pu mettre à nouveau en évidence le rôle anti-apoptotique de Rbf1 dans les cellules différenciées, ici les neurones. Ce phénotype de suppression partielle de la létalité peut être facilement criblé pour caractériser le rôle anti-apoptotique de Rbf1.

L'évaluation du rôle pro-apoptotique de Rbf1 par le phénotype d'ailes encochées permet de déterminer à la fois une atténuation et une aggravation en observant le décalage de la répartition des phénotypes respectivement vers les phénotypes faibles ou forts (Figure 24B). Cette nuance est possible grâce à l'existence d'un lien de causalité entre le nombre d'encoches observées et la quantité d'apoptose induite. Nous ne savons pas si une telle relation existe entre la force de l'activité anti-apoptotique et le nombre de naissances observées : la quantité d'individus survivants est-elle proportionnelle à la quantité d'apoptose restante après la suppression par Rbf1 ? Pour le moment, nous avons un système ON/OFF qui permet d'évaluer s'il y a eu survie ou non, et donc une activité anti-apoptotique, mais il n'est pas exclu que ce système puisse permettre de quantifier cette activité. La combinaison des taux de survie et des phénotypes d'ailes (dépliées ou non) observés sur les rescapés serait à explorer pour éventuellement pouvoir déterminer des échelles de gravité du phénotype. De plus, le taux de létalité chez les individus surexprimant *p53* dans l'ensemble des neurones étant déjà très important, cet outil ne permet pas, dans ces conditions, de détecter une aggravation du phénotype. Nous pouvons cependant supposer qu'une augmentation de l'apoptose induirait une aggravation des phénotypes d'yeux. Un des moyens de limiter cette létalité serait d'abaisser la température d'incubation. Nous avons observé qu'à température ambiante, la létalité est limitée et un nombre plus élevé d'individus naît, dont une partie avec des ailes non dépliées (Figure 27A). Dans ce cas, il serait possible d'observer une augmentation de la létalité. Cependant, cette baisse de température, en baissant l'efficacité du système UAS-Gal4, est susceptible de diminuer l'expression de *Rbf1* à un niveau insuffisant pour exercer son activité anti-apoptotique. En effet, à cette température, le niveau d'expression du transgène semble très limité puisque nous n'observons que peu ou pas d'encoches dans les ailes. Nous ne savons donc pas si son effet anti-apoptotique pourrait être évalué dans ces conditions. Dans notre cas, le but était d'évaluer si un mutant de la cible étudiée module les capacités de suppression partielle de la létalité de Rbf1, mais pas de quantifier cette modulation. Nous n'avons donc pour l'instant pas besoin d'échelle de gravité des phénotypes.

Le développement d'un tel outil aurait pour intérêt de nous permettre de tester rapidement la capacité d'un gène à moduler l'activité anti-apoptotique de Rbf1, de la même façon que nous sommes capables de tester la modulation de l'activité pro-apoptotique dans l'aile. Peu de données étant disponibles sur les mécanismes de régulation de l'apoptose par Rbf1, et en particulier son effet anti-

apoptotique, des outils de ce type peuvent s'avérer précieux car ils peuvent permettre la découverte de nouvelles voies de régulation grâce à la réalisation de cribles génétiques.

### 3. Etudier le rôle du complexe dREAM

Un autre objectif de ce projet était d'étudier le rôle du complexe dREAM dans les régulations médiées par Rbfl dans l'apoptose. Le complexe dREAM est un complexe de régulation transcriptionnelle dont l'activité est difficile à prédire. En effet, celui-ci peut être activateur ou répresseur de la transcription des gènes et sa composition est variable. Pour un même gène, l'activité du complexe dREAM peut être différente selon le tissu étudié : le gène *how* est activé par dREAM dans le disque imaginal d'aile [159], alors qu'il est réprimé dans les cellules embryonnaires [202]. Des expériences de CHIP-on-chip suggèrent qu'un complexe contenant Rbfl est répresseur, alors qu'un complexe contenant le facteur de transcription Myb est plutôt activateur. Cette séparation n'est cependant pas très claire car il existe des complexes contenant à la fois Rbfl et Myb, en général répresseurs [202]. De plus, Rbfl peut également activer l'expression de gènes, comme le laboratoire l'a montré avec *how*. Cette régulation est impliquée dans l'apoptose médiée par Rbfl, elle-même dépendante du complexe dREAM [159]. A ce jour, nous ne savons pas dans quelle mesure les régulations transcriptionnelles médiées par Rbfl se font par l'intermédiaire du complexe dREAM ou par des complexes Rbfl/facteur de transcription indépendants. De plus, nous avons pu mettre en avant, dans une revue de la littérature, l'importance et la pertinence de l'utilisation des organismes modèles dans l'étude des régulations médiées par les complexes DREAM. Le rôle de ces complexes dans l'apoptose est très peu étudié et de nombreuses pistes restent à explorer dans cette direction. L'étude des complexes de drosophile serait un point de départ de choix pour déterminer l'étendue des activités du complexe dREAM dans l'apoptose.

Pour confirmer l'implication du complexe dREAM dans ces régulations, l'utilisation de mutants ou d'ARN interférents contre Mip120 et/ou Mip130, les membres du cœur du complexe dREAM, en conditions de surexpression de Rbfl, permettrait de savoir si le complexe est impliqué dans la régulation des gènes cibles identifiés. Par RT-qPCR en conditions mutantes, nous aurions pu observer si les régulations sont conservées ou non. La perte de ces régulations en conditions mutantes indiquerait que ces sous-unités sont nécessaires.

Dans le cas où l'expression de certains de ces gènes serait dépendante de la présence du complexe dREAM, des CHIP-qPCR auraient également pu être réalisées pour préciser sa composition. Il serait notamment intéressant de se pencher plus particulièrement sur la protéine Myb, qui semble être le déterminant majeur du changement d'activité du complexe de la répression vers l'activation. Nous possédons au laboratoire deux souches permettant une surexpression de Myb étiqueté GFP (#BL32044) ou possédant une suite d'étiquettes GFP, FLAG, Tev et Biotine (#BL58482) sous contrôle de promoteurs endogènes. Après vérification que ces étiquettes ne perturbent pas

l'assemblage du complexe par co-immunoprécipitation des membres cœur du complexe Mip120 et Mip130 *via* ces Myb étiquetés, nous aurions pu tester la présence de Myb aux promoteurs régulés par Rbf1. Nous pourrions alors observer si la présence de Myb est corrélée à un sens de régulation particulier. L'étude de Georgette *et al.* semble indiquer que les complexes contenant Rbf1 et Myb sont en général répresseur et que Rbf1 est absent aux promoteurs activés. Nos résultats tendent à montrer que Rbf1 pourrait également être activateur, il serait intéressant d'évaluer si cette régulation positive fait intervenir Myb au sein du complexe dREAM.

Ces différentes pistes pourraient permettre d'identifier des patrons de régulation associés à la composition du complexe, et de préciser le rôle antagoniste de Myb et de Rbf1 au sein du complexe dREAM.

## V- Conclusion

Si ce projet de thèse n'a pas produit les résultats escomptés, il a tout de même permis de produire les outils nécessaires à la réalisation de projets de génomique dans un laboratoire dont les thématiques sont plus axées sur la biologie cellulaire et la génétique.

Le démarrage d'un projet de recherche est souvent laborieux, comme cela a été le cas ici. Si la publication de données scientifiques est en général mise en avant, toute la mise en place nécessaire est souvent oubliée. Cependant, c'est elle qui permet d'avoir des bases solides et optimisées pour le bon déroulement du projet. Ici, j'ai pu mettre au point un protocole de récupération des disques imaginaires d'ailes permettant de s'affranchir de la dissection et de ses contraintes. Je me suis également formée au ChIP-seq et au RNA-seq, et j'ai travaillé à toutes les étapes qui mènent à la production des acides nucléiques nécessaires au séquençage. J'ai également pu discuter avec les différentes plateformes de cytométrie, de séquençage et de génomique qui m'ont énormément appris sur ces différentes techniques. Je suis également capable à présent de réaliser moi-même l'analyse bio-informatique de ces données de séquençage. Cette compétence est précieuse car ces approches sont de plus en plus développées dans les laboratoires. Si ces analyses pourraient être réalisées par des personnes extérieures (plateforme de séquençage, bio-informaticien), ma capacité à pouvoir analyser mes propres données me permettra à la fois d'être autonome dans leur gestion et de comprendre le fonctionnement et les étapes de ces analyses. Mon travail a également permis le développement d'un outil pour évaluer facilement l'effet anti-apoptotique de Rbf1 qui dans le futur, pourra être utilisé par le laboratoire pour caractériser cette activité encore méconnue de Rbf1.

Ces différentes approches et mises au point m'ont amenée à me former à un large éventail de techniques de génétique de la drosophile, de biologie cellulaire (imagerie, cytométrie en flux), de biologie moléculaire (ChIP, qPCR, NGS) et bio-informatique. Si les résultats n'ont pas toujours été à la hauteur de nos espérances, ces expériences ont été très formatrices et m'ont permis d'évaluer et de

contourner les principales difficultés de ce projet. J'ai ainsi développé des outils et protocoles pour le laboratoire et la communauté scientifique et acquis des compétences précieuses que je pourrai utiliser pour mener à bien mes projets de recherche à venir.

# ANNEXES

**Annexe I : Article « Apoptosis Quantification in Tissue:  
Development of a Semi-Automatic Protocol and Assessment of  
Critical Steps of Image Processing », de Noiron *et al*, 2021**

## Article

# Apoptosis Quantification in Tissue: Development of a Semi-Automatic Protocol and Assessment of Critical Steps of Image Processing

Juliette de Noiron <sup>1,2</sup>, Marion Hoareau <sup>1</sup>, Jessie Colin <sup>1,2,\*</sup> and Isabelle GuénaI <sup>1,\*</sup> 

<sup>1</sup> Laboratoire de Génétique et de Biologie Cellulaire (LGBC), UVSQ, Université Paris-Saclay, 78000 Versailles, France; juliette.bertheault-de-noiron@ens.uvsq.fr (J.d.N.); marion.hoareau@uvsq.fr (M.H.)

<sup>2</sup> Ecole Pratique des Hautes Etudes, PSL Research University, 75014 Paris, France

\* Correspondence: jessie.colin@ephe.psl.eu (J.C.); isabelle.guenal@uvsq.fr (I.G.)

† Present address: Unité Biologie des ARN des Pathogènes Fongiques, Institut Pasteur, F-75015 Paris, France.

**Abstract:** Apoptosis is associated with numerous phenotypical characteristics, and is thus studied with many tools. In this study, we compared two broadly used apoptotic assays: TUNEL and staining with an antibody targeting the activated form of an effector caspase. To compare them, we developed a protocol based on commonly used tools such as image filtering, z-projection, and thresholding. Even though it is commonly used in image-processing protocols, thresholding remains a recurring problem. Here, we analyzed the impact of processing parameters and readout choice on the accuracy of apoptotic signal quantification. Our results show that TUNEL is quite robust, even if image processing parameters may not always allow to detect subtle differences of the apoptotic rate. On the contrary, images from anti-cleaved caspase staining are more sensitive to handle and necessitate being processed more carefully. We then developed an open-source Fiji macro automatizing most steps of the image processing and quantification protocol. It is noteworthy that the field of application of this macro is wider than apoptosis and it can be used to treat and quantify other kind of images.

**Keywords:** apoptosis; TUNEL; caspase; image processing; thresholding; signal quantification; *Drosophila*



**Citation:** de Noiron, J.; Hoareau, M.; Colin, J.; GuénaI, I. Apoptosis Quantification in Tissue:

Development of a Semi-Automatic Protocol and Assessment of Critical Steps of Image Processing.

*Biomolecules* **2021**, *11*, 1523. <https://doi.org/10.3390/biom11101523>

Academic Editor: Stéphen T. Manon

Received: 22 July 2021

Accepted: 12 October 2021

Published: 15 October 2021

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Apoptosis is a programmed cell death characterized by caspase activation, subsequent dismantling of cell components, including DNA fragmentation, and final phagocytosis of so called “apoptotic bodies” by surrounding cells or macrophages [1]. Importantly, apoptosis is not only critical for correct development of metazoan organisms, but also for their survival. Indeed, apoptosis failure is observed in many diseases including cancers. Therefore, it is widely studied, and new actors are regularly identified. Apoptosis detection can be performed by multiple methods based on various features of apoptotic steps or regulators. Imaging of apoptosis in whole tissues can rely on a more limited number of methods. The first developed and best known among them is TUNEL (Terminal deoxynucleotidyl transferase dUTP Nick End Labeling) which relies on labeling of DNA 3' ends whose numbers increase during the DNA fragmentation step of apoptosis. However, TUNEL is costly, time consuming, and also detects necrotic cells [2]. Alternatively, the use of antibodies raised against cleaved—and thus activated—executioner caspases has proved to be more specific and convenient since immunodetection protocols are less time consuming as they include fewer steps than TUNEL. In mammals, the cleaved form of executioner caspase 3 is often targeted [2]. In *Drosophila melanogaster*, the antibody used was raised against the executioner caspase Dcp-1 cleaved at Asp 216. This antibody was recently shown to actually detect the cleaved forms of both Dcp-1 and DrICE executioner caspases [3].

In this study, we wanted to compare TUNEL and cleaved caspase stainings in the most objective way. To this end, we co-stained apoptotic wing imaginal discs (the larval tissue giving the adult wing) with TUNEL and anti-cleaved Dcp-1 antibody and addressed their sensitivity and requirements in terms of image processing.

As a single image can give a wide variety of information, the first step of image analysis consists of choosing a readout (i.e., which data is worth collecting). For example, protein quantity can be assessed by measuring staining intensity, whereas tumor or bacterial colony growth can be followed by measuring the stained area. In the case of apoptosis, the most commonly found readouts are the “number of apoptotic cells” or an “apoptotic index” that has various definitions depending on the lab or experimenters [4–7].

When the readout is the “number of apoptotic cells”, many studies use a manual counting, implying that an experimenter defines interesting spots and count them. Manual counting is reliable because the expert eyes of experimenters are able to distinguish the signal of interest from background noise better than any machine. This counting can be computer-assisted by software that record every experimenter’s click but even with this assistance, this approach remains time consuming and might involve estimation bias that can potentially raise ethical questions. This is why it is better to rely on automatized—or semi automatized—computer-based methods, whenever it is possible, regardless of the readout. However, machines do not have eyes trained to recognize specific signal from background. This discrimination is allowed by the image processing steps carried out prior to quantification in order to decrease background noise and amplify the signal of interest. This proper discrimination of foreground signal from background noise is called segmentation and defines the boundaries of the objects of interest. Thus, segmentation quality directly affects quantification accuracy.

Software such as Imaris (Bitplane, Belfast, UK) or Matlab (MathWorks, Natick, MA, USA) display default functions for signal quantification [8,9] which usually comprise image processing to yield a rapid result. However, these programs are not open source and the methodology used to obtain the values is often hard to access, meaning that users have only a limited control on their implementation. Moreover, they usually are computationally demanding and thus require powerful devices to run. For all those reasons, many researchers prefer to work with open-source software with which one can develop its own protocol for image processing and quantification. Several open-source image analysis software such as QuPath (University of Edinburgh, Edinburgh, UK) or Icy (Institut Pasteur, Paris, France) have been developed but ImageJ (NIH, Bethesda, MD, USA), and its implemented version Fiji (Fiji Is Just ImageJ, NIH, Bethesda, MD, USA), remain the most widely used as they are highly versatile tools [10–12].

Once the readout has been chosen (i.e., intensity, number of objects, area), the image has to be modified to get rid of background noise and artefacts in order to improve segmentation. Images typically display three major kinds of defects: 1. general background noise; 2. isolated background pixels with aberrant high intensities; 3. groups of background pixels with aberrant high intensities. Many functions are available to improve image quality on ImageJ/Fiji software, but they often involve the experimenter’s appraisal. Unfortunately, the more the experimenter is involved, the more it is difficult to ensure that all the images have undergone the same process. However, most of the time, it is impossible to totally exclude experimenter’s involvement.

We previously showed in the *Drosophila* model that overexpressing *rbf1*, the homolog of the human tumor suppressor *RB1*, induces apoptosis in proliferating cells [13]. This apoptosis requires the pro-apoptotic Bcl-2 family member *Debcl*, and involves caspases activation [14]. It can be visualized using TUNEL on *rbf1* overexpressing wing imaginal discs. In this study, we co-stained wing imaginal discs overexpressing *rbf1* alone (*vg > rbf1*) or in the context of a *debcl* partial inactivation (*vg > rbf1, debcl<sup>E26</sup>*) with both TUNEL and anti-cleaved Dcp-1. We used the images we obtained to compare several methods of image processing and estimate their impact on the quantification of apoptosis for both assays. We then developed a semi-automatic protocol available as a free access Fiji macro



called CASQITO (Computer Assisted Signal Quantification Including Threshold Options). This protocol enables to process images from both labelings and quantify the number of apoptotic cells (count readout) or the stained area (area readout). It is worth noting that our analysis and protocol can be relevant to quantify other types of stainings outside the field of apoptosis.

## 2. Materials and Methods

### 2.1. Fly Stocks

Flies were raised at 25 °C on a standard medium. The *UAS-rbfl* and *vg-gal4* strains were generous gifts from Joel Silber (Institut Jacques Monod, Université de Paris, France). The *debcl<sup>E26/E26</sup>* was obtained from the Bloomington Drosophila Stock Center (Bloomington, IN, USA) (BL 27342) and we used a *w<sup>1118</sup>* fly stock as the reference strain.

### 2.2. Immunostaining and Images Acquisition

Third-instar larvae were dissected in 1X PBS pH 7.6 in order to remove all internal tissues except wing imaginal discs, then carcasses were fixed with 3.7% formaldehyde in 1X PBS for 20 min at room temperature and washed three times for 10 min in PBST (1X PBS, 0.3% Triton X-100). Discs still attached to cuticles were then saturated for 1 h in PBST-BSA (1X PBS, 0.3% Triton X-100, 2% BSA) and dissected again to isolate wing imaginal discs which were then incubated overnight with 1:100 dilution of anti-cleaved Dcp-1 (Asp216, Cell Signaling Technology, Danvers, MA, USA) at 4 °C. The following day, after three washes in PBST, wing discs were incubated for two hours with anti-rabbit secondary antibody (1:400, Alexa-Fluor-612-conjugated goat anti-rabbit-IgG (H + L) antibody, Molecular Probes, Thermo Fisher Scientific, Waltham, MA, USA) in PBST. Following three washes in PBST, TUNEL staining was performed according to manufacturer's instructions (ApopTag Red In situ apoptosis detection kit, merck-millipore, Temecula, CA, USA). Finally, wing discs were mounted in ProLong Diamond (Invitrogen, Waltham, MA, USA) and images were acquired using a Leica SPE inverted confocal microscope (Leica, Wetzlar, Germany) at 568 nm for TUNEL and 612 nm for anti-cleaved Dcp-1 stainings (28 discs for the *vg > rbfl* genotype and 39 discs for the *vg > rbfl, debcl<sup>E26</sup>*). Image analysis was performed exclusively with the Fiji software and the exact same zone was selected for both assays

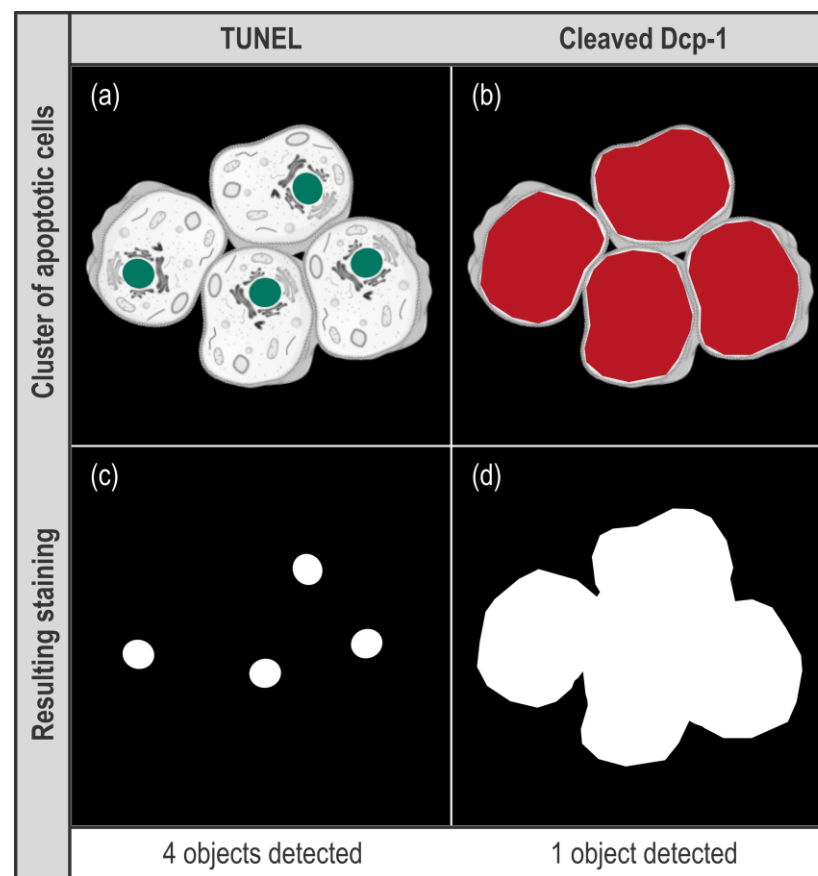
## 3. Results

### 3.1. Choosing a Readout According to the Biological Question

The readout is the data used to translate the intensity of the biological effect into numbers. Therefore, the chosen readout should be consistent with the biological question and the tool used to study it. For instance, intensity can be measured to assess the amount of a stained component. In the study of apoptosis rate, regardless of the assay, a cell is apoptotic or not. Thus, quantifying the staining intensity, even if it can somehow make sense, does not seem to be the best option for accurate quantification of apoptosis. Conversely, as long as apoptotic cells can be separated from each other (low apoptosis rate, scattered pattern, or intracellular discrete staining), counting the number of objects equates counting the number of apoptotic cells, which constitutes a valid readout. If this readout cannot be used, another valuable readout is the stained area. These count or area measurements can be used as a readout per se or can be used as primary data for calculation of an apoptotic index or score. For instance, such apoptotic index can correspond to a percentage of apoptotic cells obtained by dividing the number of apoptotic cells (count readout) by the total number of cells in the tissue section. As well, another apoptotic index can be obtained from the stained area by dividing it by the cells number (obtained by plasma membrane or nuclear co-staining), by the area of interest (surface of a cellular clone or of the tissue section). These apoptotic indexes do not indicate the actual number of apoptotic cells, which can be estimated using the count readout, but this number of apoptotic cells is rarely necessary, and those readouts provide estimates that are sufficient to compare the apoptosis rate between different samples. Moreover, in a heterogenous

tissue, indicating the percentage of apoptotic cells among a specific cell type usually makes more sense.

Here, we used TUNEL and anti-cleaved Dcp-1 to detect apoptosis. These highlight different features of apoptosis as TUNEL labels fragmented DNA in the nuclei while anti-cleaved Dcp-1 staining is cytosolic. As *rbf1* overexpression is a potent apoptosis inducer in the wing imaginal discs, the probability to have clusters of adjacent apoptotic cells is rather high. This can possibly become problematic for accurate quantification. Indeed, when adjacent cells are apoptotic, TUNEL labeling is expected to remain punctiform as nuclei remain spaced by cytoplasm (Figure 1a,c). On the contrary, with anti-cleaved Dcp-1 staining, it is expected that such adjacent apoptotic cells become indistinguishable from each other and thus appear as a single object (Figure 1b,d).



**Figure 1.** Effect of different types of stainings on the “Count” readout. (a,b) scheme showing a virtual cluster of four apoptotic cells. In (a), green spots represent nuclei stained by TUNEL while in (b), red patches represent cytosols stained by anti-cleaved Dcp-1. (c,d) schemes present the result of image processing for these signals.

Therefore, when the readout is the count of apoptotic cells, these clusters of labeled cells are not expected to alter the quantification for TUNEL while they may cause an underestimation of the number of apoptotic cells with anti-cleaved Dcp-1 staining. The extent of this underestimation is difficult to anticipate as it depends on many parameters. Still, this underestimation surely increases with the apoptotic rate—as the probability to have clusters of apoptotic cells increases—which could lead to an artificial flattening of the difference of apoptosis rate that may exist between two conditions. As for the area readout, the size of the wing imaginal disc cells (and their nucleus) being homogenous, the stained area indirectly reflects the number of apoptotic cells without being impacted by their relative localization. In the end, counting cells seems, at least at first sight, a more precise, more direct, readout of apoptosis than area. However, this readout might

be altered by apoptotic cell clusters. As it is not possible to anticipate how these clusters will affect the quantification in our experimental set up, we chose to use both count and area readouts. With ImageJ/Fiji, these two readouts can be obtained using the “Analyze Particles” function, which only works on binary 2D images. This means that our image processing protocol should include both z-projection and binarization using a threshold, these two treatments being compatible with our set up. Indeed, as our tissue of interest is a monolayer, z-projection should not affect quantification. Moreover, whether a cell is apoptotic or not, our readouts do not depend on signal intensity and binarization by itself should not affect the quantification.

### 3.2. Designing an Image Processing Protocol

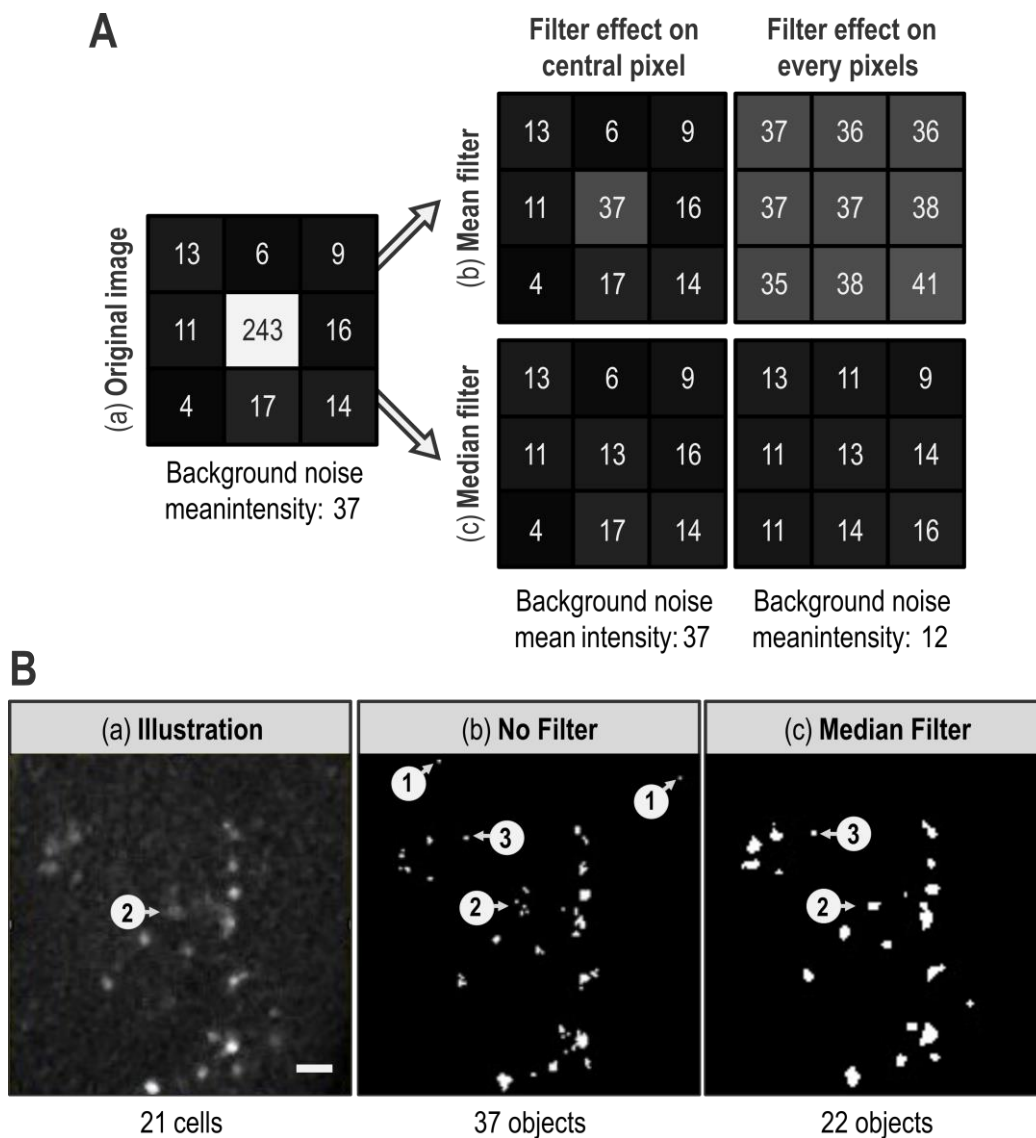
Steps of image processing directly depend on the chosen readout. In order to get both the number of objects and the stained area, our image processing protocol is based on three major steps: 1. Background noise reduction, 2. Compression of our 3D images into 2D by a z-axis projection, 3. Thresholding. Those steps allow appropriate segmentation required for relevant quantification by the “Analyze Particles” function. Importantly, on ImageJ/Fiji, there are many ways to minimize background signal, several ways of compressing a 3D image in 2D and multiple ways to determine a threshold, resulting in countless combinations of possible image processing. In this study, we investigated the weight of these parameters on signal segmentation to end up with an optimized and unbiased protocol for apoptosis quantification.

#### 3.2.1. Median Filter and Size Limitation Efficiently Reduce Artefacts

When quantification is automatized, definition of the signal of interest by segmentation is even more critical. Indeed, wrong segmentation can lead to quantification of unreal objects and thus give useless results. Thus, the signal of interest boundaries has to be better defined while background noise has to be decreased. General background noise can be minimized in many ways depending on the kind of images, the desired readout, and the defaults faced. In our case, mandatory use of a threshold would blacken every low intensity pixel constitutive of the general background noise. However, if binarization of the image efficiently removes diffuse low intensity background noise, it is not sufficient to erase artifactual pixels with aberrant high intensity, i.e., whose intensity is higher than the threshold value. Fortunately, isolated aberrant high pixels can be dealt with filters. Filters are matrix operations that re-calculate a pixel intensity value based on itself and its neighbors. The two mainly used filters are “Mean Filter” and “Median Filter” (Figure 2A).

A “Mean Filter” with a radius of 1 gives a pixel an intensity value that corresponds to the mean of its value and those of its direct neighbors (Figure 2A(b)). Hence, the value of an isolated pixel with aberrant high intensity is attenuated by the intensity value of its neighbors. However, as extreme values impact mean calculation, every neighboring pixel is affected by the isolated aberrant high pixel and their intensity is artificially increased (Figure 2A(b)). By contrast, a “Median Filter” with a radius of 1 gives to a pixel an intensity value corresponding to the median of its value and those of its direct neighbors (Figure 2A(c)), which is expected to be much closer to local intensity value. Besides, as extreme values effect on median calculation is low, high intensity isolated pixel impact on its neighbors is negligible. Overall, the “Mean Filter” tends to spread an aberrant high value whereas a “Median Filter” tends to confine it. This effect is illustrated in Figure 2B, where the “1” arrows of the “No filter” panel show typical isolated aberrant high pixels that are efficiently erased by a “Median Filter” with a radius of 1 (Figure 2B compare (b) and (c)). Aside from this benefit, the area pointed by the “2” arrow exemplifies the median filter ability to preserve edges of an object. Indeed, on the illustration, human eyes easily detect that the “2” arrow targets a marked cell (Figure 2B(a)). However, this object is heterogeneous: in a restricted space, it contains few pixels of high intensity and many pixels of low intensity (i.e., below the chosen threshold). Without any filter, only high intensity pixels are kept after the thresholding step thereby fragmenting this object in

several small groups of pixels (Figure 2B(b)). Thus, with no further treatment, multiple objects will be counted in this area, which does not reflect reality. However, as these high intensity pixels are close to each other, the “Median Filter” with a radius of 1 homogenizes intensity values within this object. This allows its reconstruction and gives a segmentation consistent with reality (Figure 2B(c)). Thus, when the “Median Filter” with a radius of 1 is applied, the “number of objects” decreases only to be closer to what a human eye would count.



**Figure 2.** Filters effect on background noise and segmentation. (A): The grid in (a) presents intensity values of a  $3 \times 3$  pixels image. The grid in (b) depicts how a “Mean Filter” with a radius of 1 affects the central pixel of the original image (left) and the whole image (bottom right). In order to make intensity differences more visible, each boxes background color corresponds to the double of each pixel intensity value in greyscale. (B): (a) image of a portion of a TUNEL-labeled wing imaginal disc after a Max Intensity z-projection. (b) same as (a) after binarization using a manually determined threshold. (c) same as (b) but a “Median Filter” with a radius of 1 was applied to the image before Max Intensity z-projection. White bar corresponds to  $10 \mu\text{m}$ . Arrows with circled numbers 1, 2, and 3 target areas of interest. The number of visible objects in each panel was manually counted.

Although the “Median Filter” with a radius of 1 efficiently reduces the number of artifactual objects by erasing isolated high pixels, the issue of groups of pixels with an intensity higher than the threshold value remains. A way to eliminate most of those artifacts

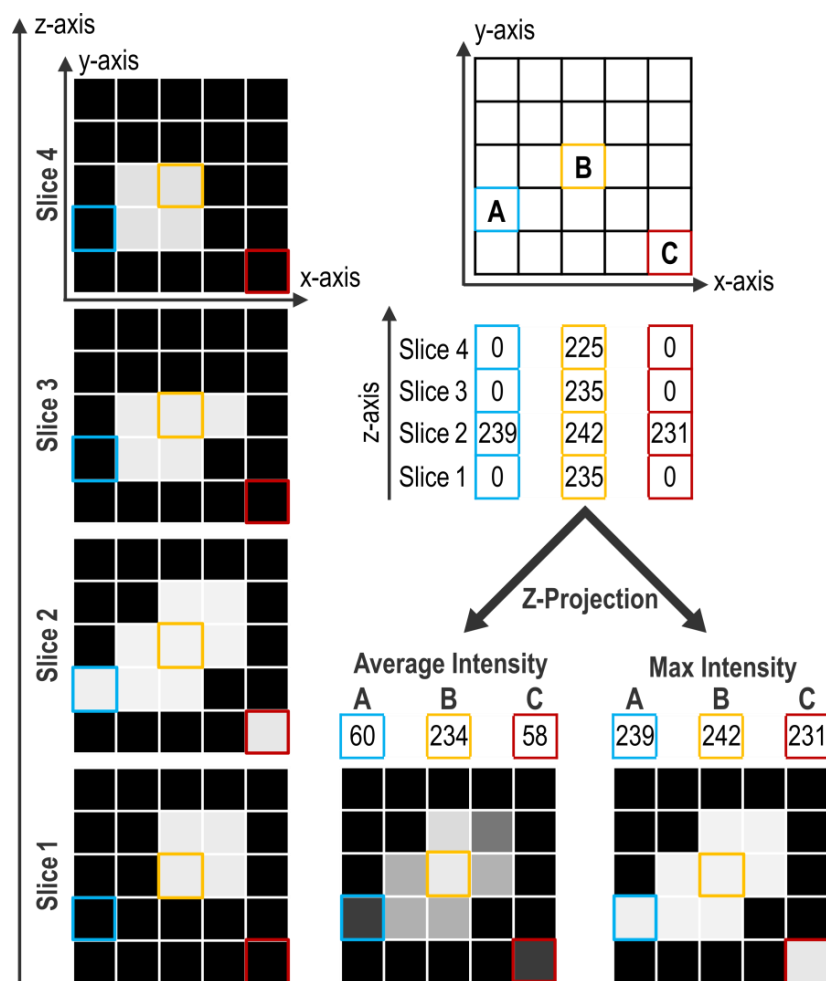
is to limit our analysis to objects with a size consistent with the smallest biological object of interest. In our case, this smallest biological object is a TUNEL-labeled nucleus, we assessed their size on a few random images and thus set a size limit at 2  $\mu\text{m}$ . Importantly, this “>2  $\mu\text{m}$  size limitation” fits our data but should not be taken as a default value and must be adapted for other kinds of signals or cell types. Moreover, as the “Analyze Particles” function records the size of every object, this filtering can be carried out after quantification. In Figure 2B, the “>2  $\mu\text{m}$  limitation” eliminates artifactual objects pointed out by the “3” arrow as well as individual pixels such as those pointed by the “1” arrows. It thus appears very powerful in order to “clean” the image. However, as efficient as the size limitation may be, it cannot replace “Median filter”. Indeed, as already explained, in the absence of a “Median Filter”, the cell indicated by the “2” arrow in Figure 2B becomes fragmented in several small groups of pixels, each one being smaller than 2  $\mu\text{m}$  (Figure 2B(b)). Thus, without the “Median Filter”, these pixels are eliminated by the “>2  $\mu\text{m}$  limitation” and the actually labeled cell indicated by the “2” arrow is not included in the quantification of the apoptotic signal. Here, reconstruction of the object by the “Median Filter” prevents its elimination by the “>2  $\mu\text{m}$  limitation” (Figure 2B(c)). In the end, combination of a “Median Filter” with a radius of 1 and “>2  $\mu\text{m}$  limitation” allows a better segmentation and a more accurate quantification.

### 3.2.2. Max Intensity Z-Projection Improves Contrast

Confocal microscopy gives the possibility to capture objects in 3D. However, image processing often requires transforming volumes into 2D images by compressing the z-axis. In our case, the “Analyze Particles” function used to quantify the signal of interest requires 2D images. Flattening a 3D volume can seem counterproductive, as separate objects on the same z-axis will be reduced to a single one on the final 2D image. In our case, this is unlikely to happen since imaginal disc cells are organized in a monolayer with only limited folding and this can also be true for tissue sections as long as they are thin enough. In the case of a multilayer tissue, z-projection can still be used if the apoptosis rate is low enough to prevent the occurrence of having multiple apoptotic cells in the same z-axis.

Projection consists of compressing the signal contained in every pixel of a z-axis into a single one. With ImageJ/Fiji, projection can be performed in several ways but two of them are mainly used. The first method, Average Intensity (AI), calculates the average intensity of all the pixels of a z-axis. The second, Max Intensity (MI), only retains the maximal intensity value along the z-axis. Figure 3 presents examples of these projection methods on a virtual object without any other treatment (i.e., median filter).

The z-axis presented in “A” shows only one illuminated pixel on slice #2 and this pixel is included in the object. The “B” z-axis shows numerous pixels highly illuminated, comprised in the object. The “C” z-axis presents an artifactually illuminated pixel on slice #2 that is not comprised in the object. Comparison of the projection methods shows that AI projection decreases the importance of the artifactual pixel of the “C” z-axis, while the MI projection increases its weight. However, in the processing protocol, the preceding use of a “Median Filter” with a radius of 1 eliminates most of those artefacts that are thus not present anymore at the projection step. Conversely, the AI projection of the “A” z-axis leads to loss of signal even if it is part of the object. Furthermore, with the AI projection, the contrast between object and background is weak, therefore the range for the appropriate threshold value is limited (Figure 3). After a MI projection, contrast is enhanced and thus threshold determination is easier for the experimenter, which helps limiting the experimenter bias. This is particularly important for signals with low contrast such as TUNEL. In our case, these two projection methods do not end up in drastically different results but, overall, MI projection presents more benefits than AI projection.



**Figure 3.** Projection methods. The left panel presents different planes obtained by imaging a virtual object surrounded by a perfect background noise of 0. On the right, the upper panel presents the coordinates of three pixels: A, B, and C, and their respective intensity values along z-axis. Bottom panels show the resulting projection obtained either by an Average Intensity or a Max Intensity projection with respective intensity values obtained for A, B, and C pixels. For all representations, boxes background color corresponds to their pixel intensity in grey scale.

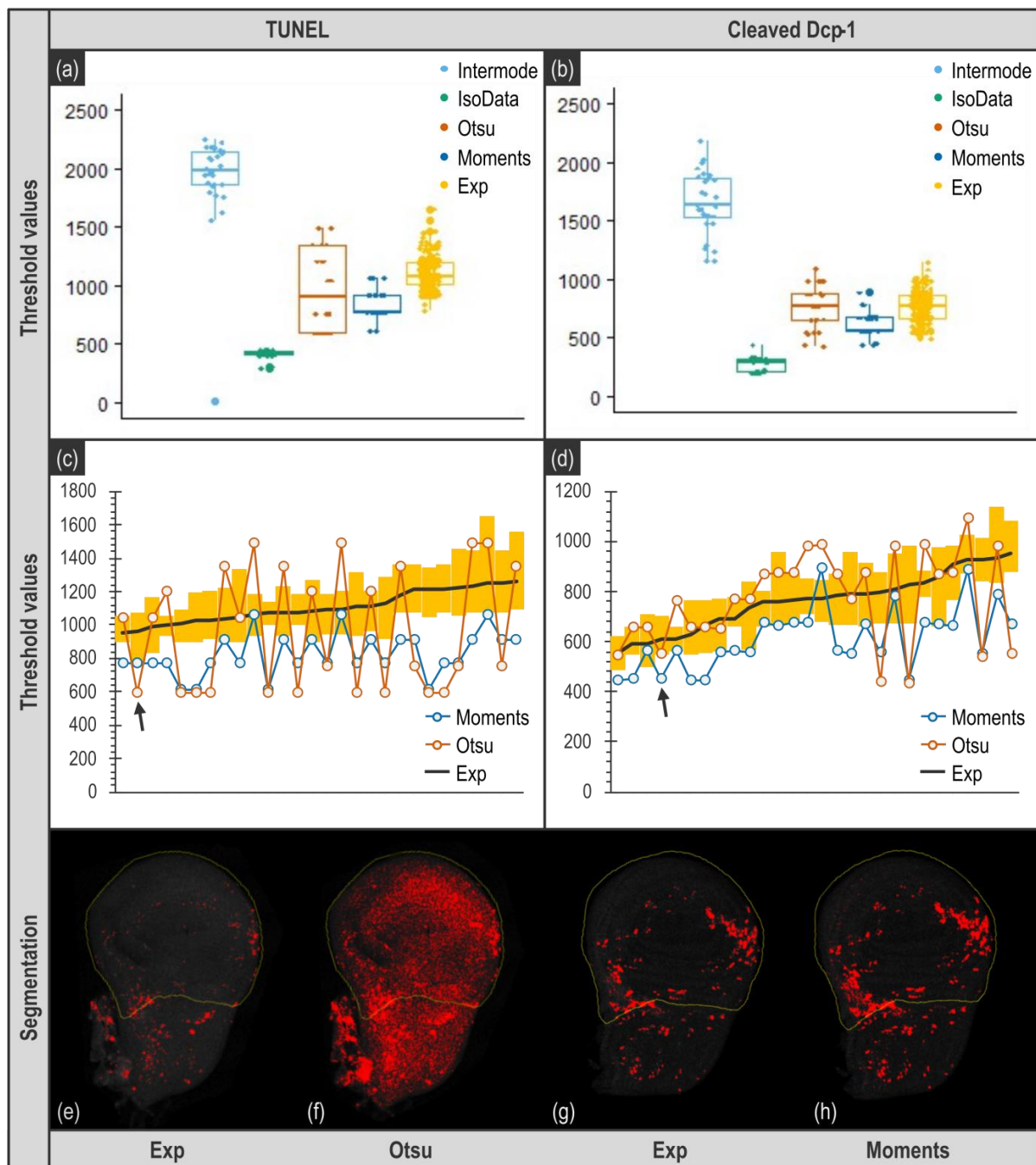
### 3.2.3. Use of Custom Manual Thresholds Gives the Best Segmentation for Relevant Quantification

The “Analyze Particles” function used for quantification requires the image to be binary. The transition from a greyscale image to a black and white image involves the setting of a threshold that defines an intensity value above which a pixel is turned to white and under which a pixel is turned to black. Ideally, this value should enable to get an image where white only corresponds to the signal of interest. Thresholding is the last step of segmentation and finally defines the objects of interest, which is critical for accurate quantification. Therefore, among the steps of image processing, it is the one that has the most dramatic effect on quantification accuracy, thus we dedicated particular attention to the threshold determination method.

Threshold can be automatically set by algorithms that analyze specific features of the image intensity histogram to determine a threshold value using either simple indicators such as the mean, maximal, or minimal intensity values, or more complex formulas. Hence, algorithms appear as an unbiased method to obtain a specific threshold value per image. We thus wondered whether any of the 16 thresholding algorithms currently available in the menu Image > Adjust > Threshold of ImageJ/Fiji could be used to determine a threshold that properly segments apoptotic signal in our images. Using some randomly

chosen images, after application of a “Median filter” with a radius of 1 and a Max Intensity projection, we visually checked if these algorithms could provide a threshold value allowing a relevant segmentation (i.e., consistent with apoptotic signal). Most algorithms did not pass the visual inspection step as they yielded unrealistic segmentation either by ignoring a great portion of the signal or by including artifactual signal. However, two of them, Otsu [15] and Moments [16], seemed capable of discriminating actual apoptotic staining from background. We then performed a more detailed analysis of the threshold values obtained with these algorithms by comparing them to the ones obtained by experimenters. To this end, for the 28 images of the *vg > rbf1* genotype, three experimenters determined the threshold to use for each staining (TUNEL or anti-cleaved Dcp-1) by eye and in triplicate (see Supplementary Figure S1). The criteria for manual threshold determination were to set a threshold value that allows the positive signal to be consistent with the original image and located in the vestigial domain. Thresholds for these images were also determined using the 16 algorithms. As expected, algorithms inducing obvious unrealistic segmentation of the apoptotic signal yielded threshold values very far from the range of the ones determined by experimenters (Figure 4a,b; compare IsoData and Intermode with Exp. Data not shown). On the contrary, Moments, Otsu, and experimenter threshold values are in the same range (Figure 4a,b).

From this global analysis, it could seem that algorithms can be as efficient as human eye for threshold determination (compare for instance Otsu and Exp in Figure 4b). However, visually, Otsu capability to determine a relevant threshold seemed irregular. We thus further deepened our data analysis and compared the threshold values obtained not globally but for each image. As shown in Figure 4c,d, the values obtained with Otsu, if they tend to be roughly the same on average than the ones obtained by experimenters, are actually most of the time out of the range of experimenters values. This is particularly striking for images obtained from TUNEL (Figure 4c) as Otsu’s determined values are far higher or lower than the ones obtained by any experimenter. This would not be an issue as long as the threshold values obtained still allow a realistic segmentation of apoptotic signal and subsequent relevant quantification. However, such deviation of the threshold value results in an inappropriate segmentation (compare Figure 4e,f), that necessarily ends in a biased or most likely totally wrong quantification. When it comes to Moments, it provides threshold values that are usually lower than experimenters’ ones (Figure 4a,b), which means that using this algorithm tends to include some background noise to the quantification. The question resides then in determining whether this amount of background noise is important enough to alter quantification. In the case of TUNEL staining (Figure 4c), values are quite low, thus it certainly affects quantification rather significantly. By contrast, when images come from anti-cleaved Dcp-1 staining, the threshold values given by Moments are much closer to the ones obtained by experimenters, they actually seem very similar to the lowest values determined by experimenters (Figure 4d). Therefore, one could assume that the variability of threshold values between Moments and an experimenter is comparable to the one that exists between experimenters. We tested this by comparing the relationship between the two most distant experimenters’ batch of measurements to the one between Moments and its closest dataset. As shown in Supplementary Figure S2B, if two experimenters do not determine exactly the same value for the threshold, their estimates remain consistent with each other ( $p = 10^{-5}$  and  $R^2 = 0.53$  for the most distant measurements), the difference can be more or less described as a given experimenter tending to set thresholds always lower than the other. This is a systematic error that should affect quantification only moderately. On the contrary, threshold values determined by Moments are not consistent with the values of experimenters ( $p = 0.8$  for the closest in Supplementary Figure S2B). This indicates that Moments can set a low threshold value when all experimenters would have chosen a higher one, but it is not always the case, and most of all, the extent of this underestimation (i.e., the range of the difference between experimenters and Moments’ threshold values) is variable. This is more problematic as the extent of background incorporation in the quantification will then vary and might alter quantification relevance.



**Figure 4.** Comparison of thresholding methods. Upper panel presents overall distribution of threshold values per image obtained with various thresholding methods for images of *vg > rbf1* genotype. (a,b) Exp distribution (in yellow) corresponds to the whole of the threshold values determined by three experimenters in triplicate (raw data are presented in Supplementary Figure S1). Intermode (in light blue) and IsoData (in green) are examples of algorithms yielding inadequate values, very far from experimenters' distribution. Otsu (in brown) and Moments (in blue) are algorithms that seemed usable. (c,d) show a pairwise organization of the threshold values for each image obtained by Otsu (in brown), Moments (in blue) and experimenters (Exp, the black line corresponds to the mean of the nine values determined the experimenters and the yellow bars correspond to the range of the threshold values determined by the experimenters for each image). Images were ranked according to the mean of the experimenter values (black line). Bottom panel illustrates the result of the binarization using the threshold values determined by experimenters (e,g), Otsu (f) and Moments (h). Black arrows of the middle panels target the image used for illustrations presented on the bottom panel. Importantly, these images were chosen as they are representative of the deviation between the algorithm and the experimenter average value (chosen images have a deviation equal to the median of the deviations).



Contrary to algorithms, manual determination of the threshold values appears quite robust. Indeed, comparison of manually determined threshold values between experimenters for individual images shows a low variability and a good reproducibility both between several determinations of a given experimenter and between experimenters (Supplementary Figure S1). As all images from an experiment are acquired identically, with the same microscope settings and come from samples treated with the same solutions, at the same time, theoretically, the appropriate threshold value is expected to be the same for every image. Moreover, using the same unique threshold value for all images can be considered as more objective and unbiased.

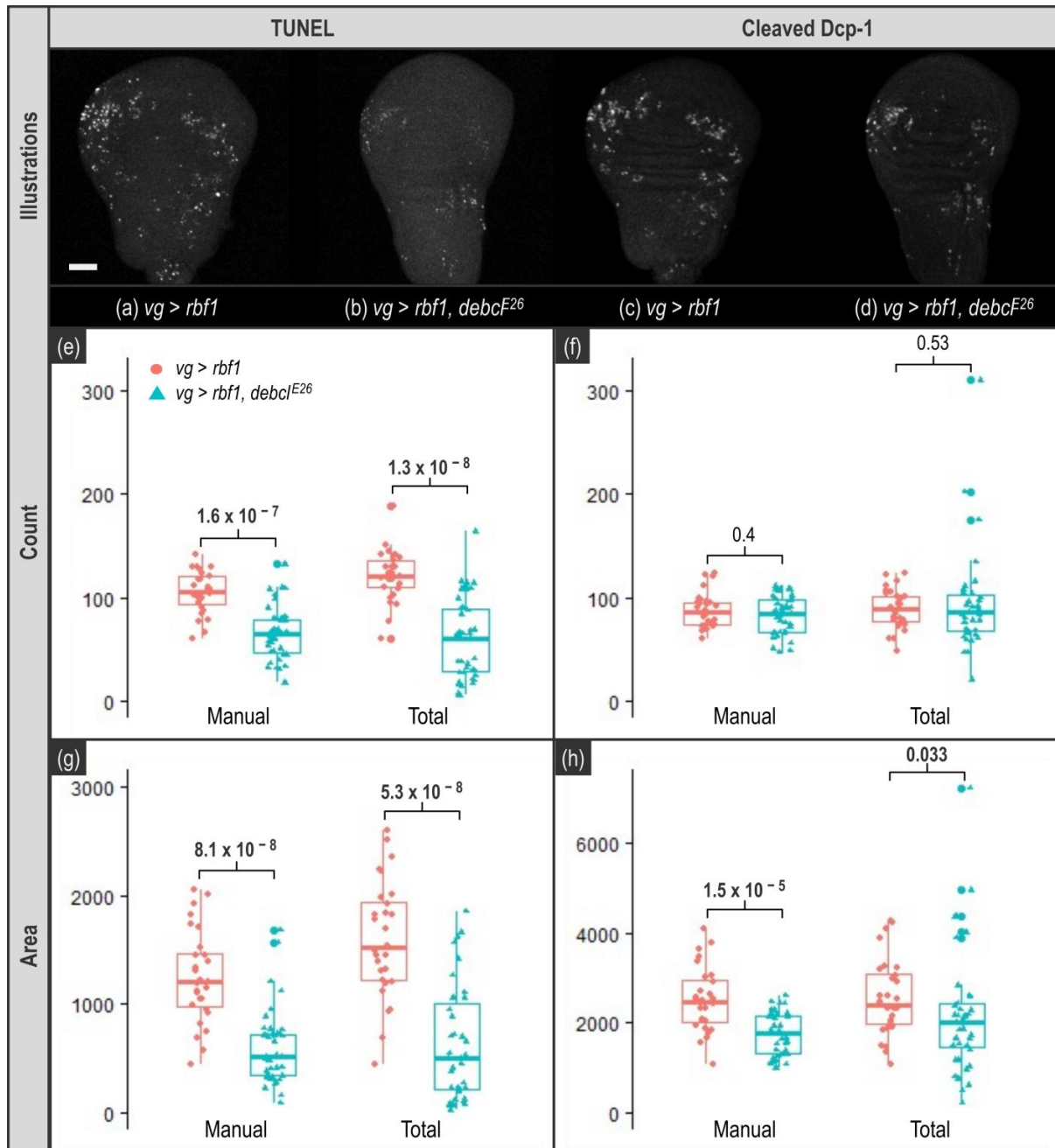
However, manually determined threshold values display some variability (Figure 4, Exp). This can justify using a distinct individual, manually determined, threshold value for each image as it might enable a more accurate segmentation and subsequent quantification. In order to assess to which extent these two thresholding methods (unique threshold value for every image or individual, manually determined threshold value per image) can affect quantification and detection of our biological effect, we tried both (Figure 5).

In the Manual condition, each image was binarized using its own manually determined threshold value (determined by experimenter 1, measure 3). From these individual threshold values, we calculated the median value per genotype and then the median of these medians. This latter value was used as the unique threshold value to binarize all images in the Total condition (1128 for TUNEL and 777 for anti-cleaved Dcp-1). We decided to use the median of the medians per genotype rather than the global median (calculated from the whole of the images independently of their genotype) to avoid giving more weight to a genotype (that might have a larger headcount for instance).

For TUNEL stained images, both thresholding approaches give a workable quantification both for count and area readouts (Figure 5e,g). Indeed, a significant decrease in apoptosis between the two genotypes is detected in all cases. However, even if Total or Manual thresholding methods enable to detect the biological effect, we noticed that quantification is still somehow altered when a unique threshold is used (Supplementary Figure S3).

As for anti-cleaved Dcp-1-stained images, the first observation is that the count readout was not usable. Indeed, we knew that apoptotic cells clusters might alter quantification as these clusters might be considered as a single object. Moreover, such underestimation is enhanced when the apoptosis rate rises, eventually leading to the flattening of the difference between two samples. However, we chose to keep this readout, as it was not possible to anticipate the extent of this phenomenon in our samples. As shown in Figure 5f, with this readout, the difference in the apoptosis rate between the two genotypes becomes undetectable. This indicates that the level of apoptosis induced by *rbf1* generates apoptotic cells clusters frequently enough to significantly alter quantification, and this, whatever the thresholding method, thus prohibiting the use of the count readout. By contrast, when anti-cleaved Dcp-1 staining is quantified using the area readout, the difference between the two thresholding methods (unique versus individual thresholds) becomes obvious. As shown in Figure 5h, when a unique threshold value is used for all images (Total), the difference of apoptosis rate between the two genotypes is barely detectable ( $p = 0.033$ ). Moreover, even with an individual, manually determined threshold value per image, all the quantification results show a noticeable variability (Figure 5, Manual) due to differences between specimen, variation of the UAS/gal4 system intensity and small asynchrony in the development of the larvae. However, use of a unique threshold value to process every image tends to increase this variability and induces the appearance of extreme values compatible with an overestimation due to inadequate segmentation (see highest values for *vg > rbf1, debcl<sup>E26</sup>* genotype in Figure 5h and also Supplementary Figure S3). On the contrary, the use of individual specific threshold values (Manual) enables to readily detect the difference of apoptosis rate between the two genotypes ( $p = 1.5 \times 10^{-5}$ ).

In the end, this analysis shows that, in our case, using an individual threshold per image is more appropriate and turns out to be the safest option for accurate segmentation, and thus relevant quantification.



**Figure 5.** Impact of thresholding on quantification. TUNEL and anti-cleaved Dcp-1 staining quantifications. Upper panel shows representative images of wing imaginal discs stained by TUNEL (a,b) or anti-cleaved Dcp-1 (c,d) for *vg > rbf1* (a,c) and *vg > rbf1, debcl<sup>E26</sup>* (b,d) genotypes. White bar corresponds to a 50  $\mu\text{m}$  scale. “Count” panel presents quantification of TUNEL (e) and anti-cleaved Dcp-1 (f) signal based on the counting of the number of objects according to the different thresholding methods (Manual and Total) for *vg > rbf1* (in pink) and *vg > rbf1, debcl<sup>E26</sup>* (in blue). “Area” panel presents quantification of TUNEL (g) and anti-cleaved Dcp-1 (h) signal based on the number of white pixels (stained area) according to the different thresholding methods (Manual, or Total) for *vg > rbf1* (in pink) and *vg > rbf1, debcl<sup>E26</sup>* (in blue). Manual indicates a custom dedicated threshold was used for each image whereas Total indicates that the same common unique threshold value was used to process all images. This unique threshold value is the median of the median per genotype of the manually determined values. *p*-values displayed above results were obtained using Wilcoxon tests.

### 3.3. Quantifications of TUNEL or Anti-Cleaved Dcp-1 Stainings Do Not Have the Same Requirements

TUNEL and anti-cleaved caspase stainings are widely used to assess the level of apoptosis in tissues. However, depending on the experimental set-up, the quantification step may become delicate.

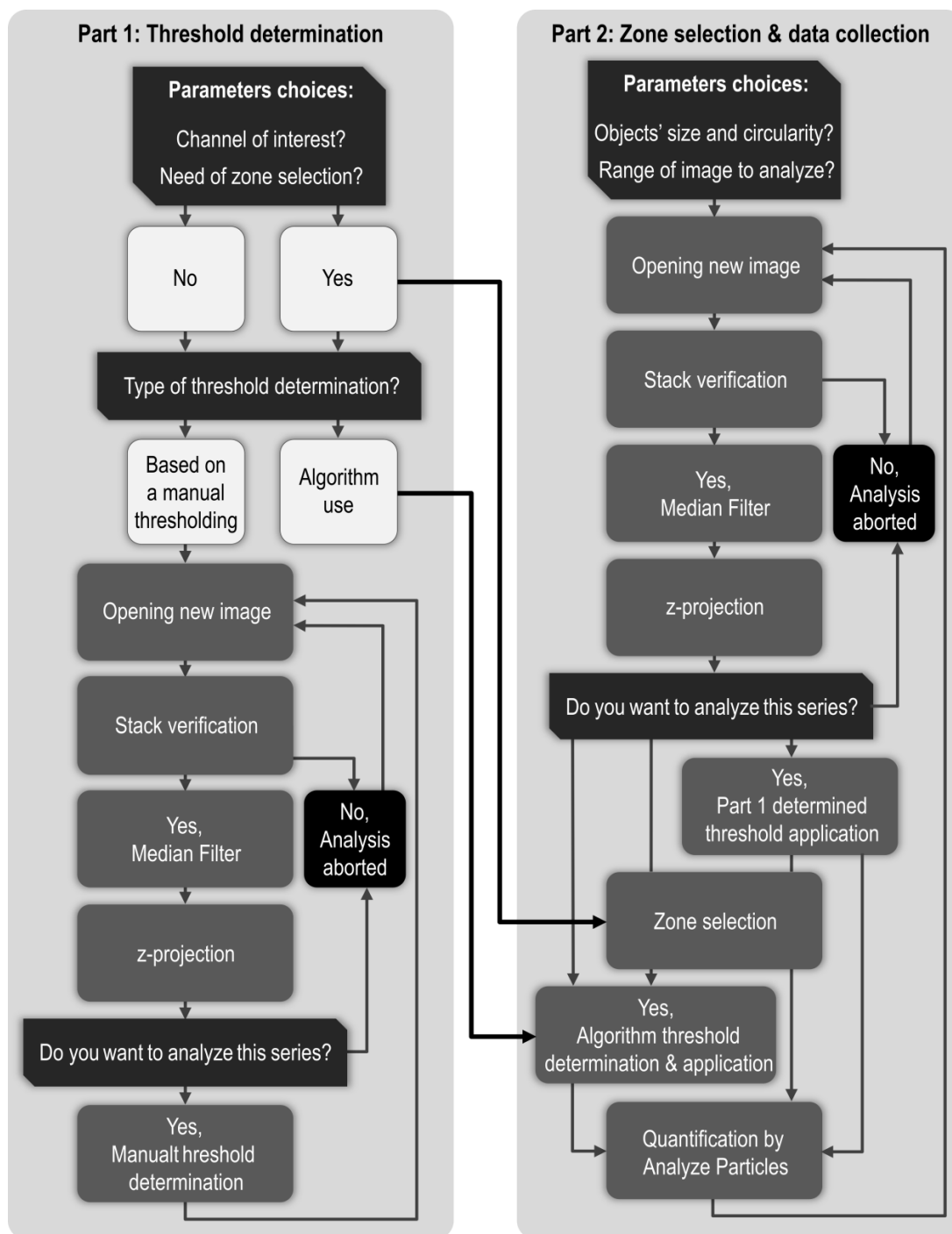
TUNEL appears as a quite robust apoptosis detection technique. Indeed, it allows to quantify apoptosis and to detect differences in apoptosis rate whatever the thresholding approach, and with both count and area readouts (Figure 5e,g). This was not totally expected since in our images, there was not a strong contrast between the apoptotic signal and the background (Figure 5a,b). However, as previously mentioned, TUNEL assay is costly, time consuming and lacks specificity as it also detects necrotic cells.

On the contrary, using antibodies against cleaved caspase(s) is considered as a more specific and convenient staining of apoptotic cells. By contrast with TUNEL which labels nuclei, caspase staining covers the whole volume of the cell, raising the issue of adjacent apoptotic cells when the readout is counting cells. Indeed, counting the number of apoptotic cells stained with anti-cleaved caspase antibody is perfectly possible as long as apoptotic cells are sufficiently separated from each other. In our experimental set-up, it appeared that the apoptosis rate was too high to prevent the underestimation of the signal due to fusion in a single object of clustered apoptotic cells. When the stained area was measured, it revealed that images from anti-cleaved Dcp-1 stainings should be carefully processed because determination of the threshold value to use for binarization is particularly important. Indeed, even if the difference of apoptotic rate between the two genotypes was known and easily seen by eye (Figure 5, compare (c) and (d)), its detection after quantification was not obvious. Actually, the decrease in apoptosis between the two genotypes is barely detected when a unique threshold value is used for the segmentation of all images (Figure 5h) whereas using a specific threshold value for each image enables to see it. Therefore, anti-cleaved Dcp-1 staining quantification is more affected by image processing than TUNEL and should be handled more carefully.

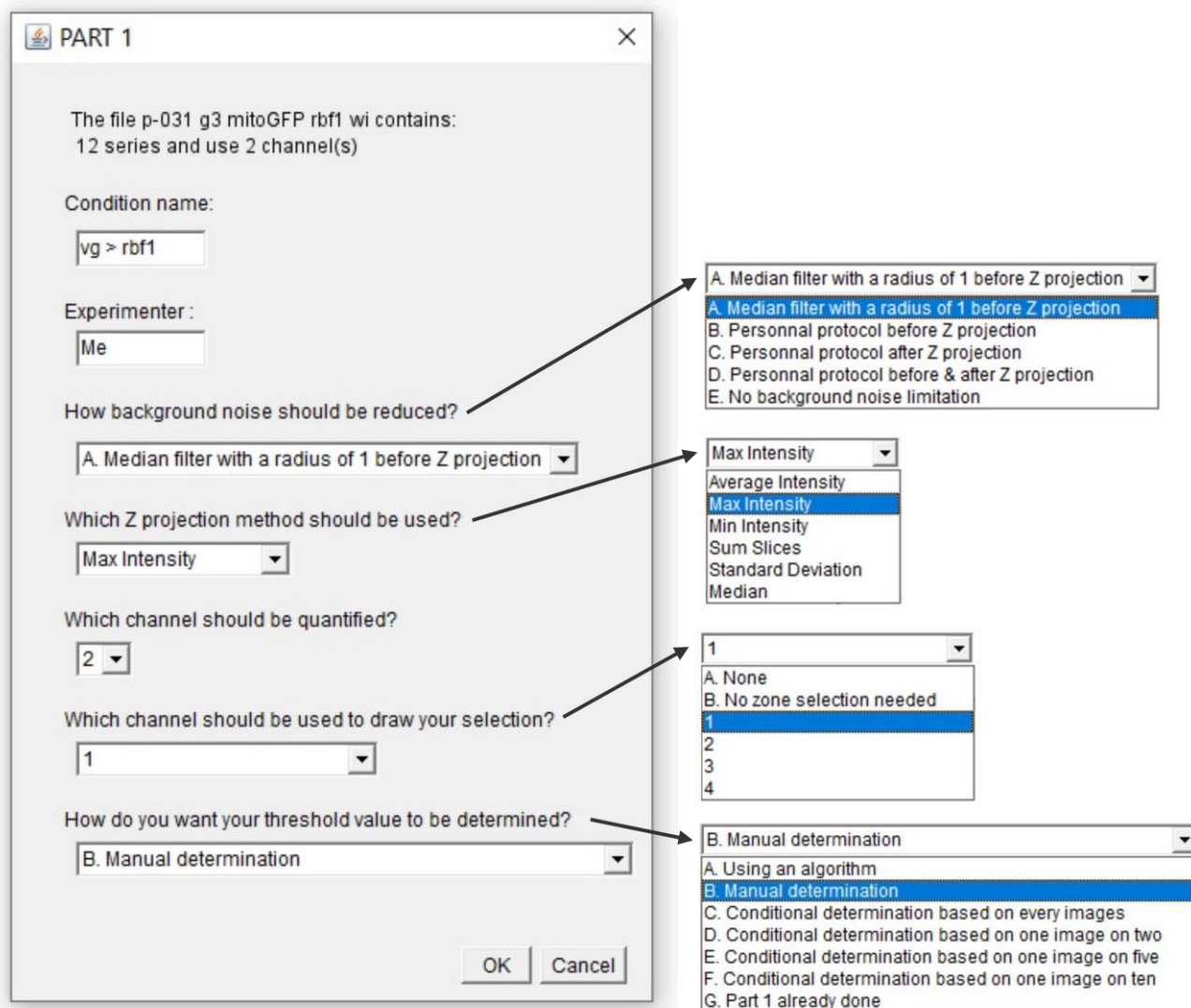
### 3.4. Macro Explanation

Once the image processing protocol was established and validated, we worked on its automatization. Indeed, doing this processing for every image manually is not only time consuming but also error-prone since it increases the probability to skip or treat twice an image or to make mistakes during data collection. We automatized this image processing protocol by developing an open-source macro called CASQITO (Computer Assisted Signal Quantification Including Threshold Options, available at <https://github.com/JdNoiron/CASQITO>, accessed on 10 October 2021). Importantly, if the image protocol presented here can be manually performed both with ImageJ and Fiji, the CASQITO macro only runs with Fiji. This macro limits experimenter's involvement to threshold determination. As we only worked on Leica microscopes, this macro only supports .lif files and should be adapted for other formats. During processing of the images, the macro generates several files for every given .lif file, which will be stored in the same folder as the parental .lif file. We thus recommend recording images from different conditions (genotypes or treatments) in separate .lif projects. The first file generated is a .txt file that recapitulates data from the Log window, which conserves settings associated with the treatment of the .lif file. Two .xls files containing results are also generated, respectively, compiling threshold determination and quantification results. It is worth noting that the latter provides all possible quantifications obtainable with the "Analyze Particles" function whatever the chosen readout may be. A .png file is also generated to display a histogram representing the distribution of the obtained threshold values. If a zone selection is required, another .png file is generated for each image to display experimenter zone selection. Lastly, the macro is not able to process multiple regions of interest on the same image, thus, even if two objects (in our case two wing imaginal discs) are in the same field and can be captured in the same image, it is highly important to capture this field twice.

The macro consists of two parts described in Figure 6, the first part allows determination of threshold values, and the second part allows zone selection and quantification.



**Figure 6.** Fiji macro workflow. Giving our image processing protocol, the macro is divided in two major parts. Part 1 (left panel) is dedicated to manual threshold determination while Part 2 (right panel) is dedicated to zone selection and quantification after application of the previously determined threshold. In Part 1, images are opened, processed according to parameters set in the window presented in Figure 7 and presented to user for threshold determination. Once every image of a .lif file has been treated in Part 1, threshold results are recapitulated before starting Part 2. In Part 2, images are opened again and treated as in Part 1, chosen threshold is applied and the resulting image is presented to user for zone selection before quantification. When every relevant image has been treated, quantification results are superficially analyzed to yield mean, standard deviation, min, and max.



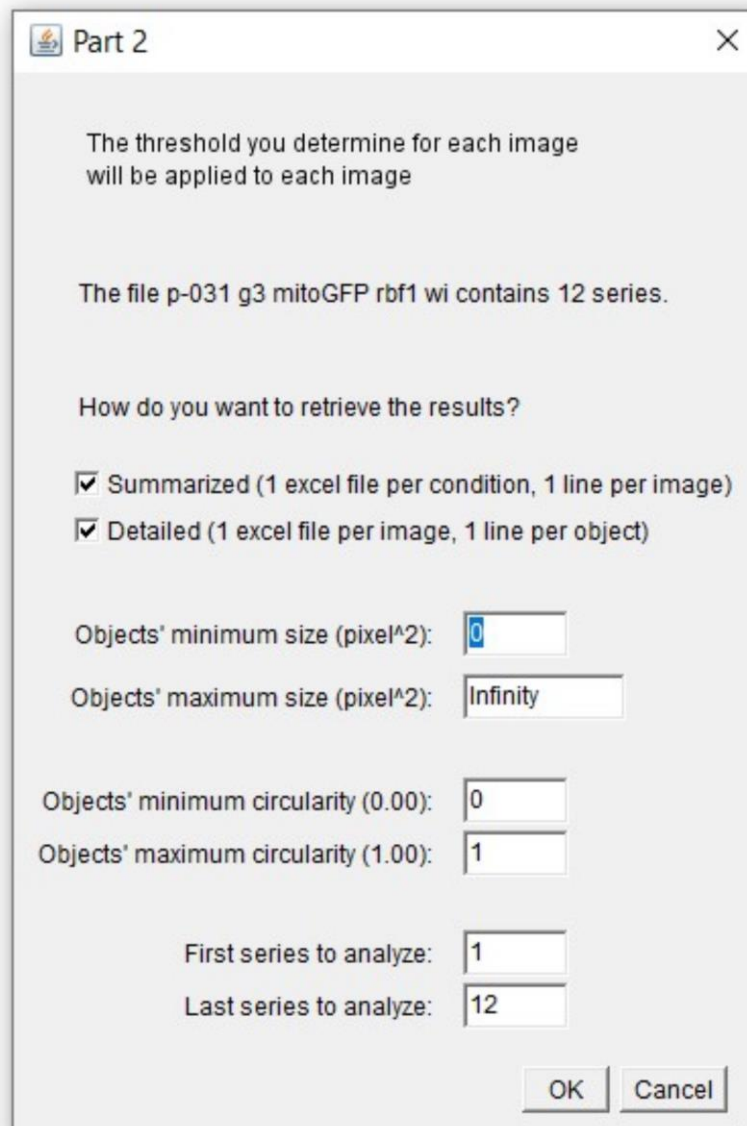
**Figure 7.** Macro Part 1. The macro starts by collecting few parameters. Importantly, the chosen “Condition name” will end up in the name of the files associated with the .lif file. The default background noise reduction method is a “Median Filter” with a radius of 1 but user can apply its own protocol before and/or after z-projection if needed. As a thresholded image is in black and white, it can hardly be used to define region of interest when a zone selection is needed. Therefore, the macro offers the possibility to define region of interest (i.e., vestigial domain in our case) on another channel or an unthresholded version of channel of interest. If the whole image has to be analyzed, a “No selection needed” option is available. Concerning threshold determination, “A. Using an algorithm” and “B. Manual determination” options will lead to application of an individual threshold value per image. If use of a unique, representative, threshold value to treat every image of a .lif file is wanted, user can make its choice after treating every image (option “C”), 50% of the images (option “D”), 25% of the images (option “E”) or 10% of the images (option “F”). If manual determination has already been carried out, the user may skip Part 1 with option “G”. Once chosen, those parameters are recorded in the Log window which content is ultimately saved in a .txt file.

Once a .lif file has been chosen, the window presented in Figure 7 opens to set a few parameters such as selecting the channels of interest and the methods to use for background noise reduction and z-projection. A more detailed explanation of the macro is provided in the Supplementary Materials.

Once these parameters are set, the macro opens the first image (that should be a stack) and process it according to the chosen parameters for background noise reduction and z-projection. In our case, a “Median Filter” with a radius of 1 is applied and the Max Intensity method is used for z-projection. At this point, threshold can be manually determined

on the processed images. Part 1 is over when all images have been processed and as it generates a table of the determined thresholds, Part 2 can be run right after or later.

Part 2 begins with the opening of the window presented in Figure 8, which offers the possibility to use the threshold value previously determined in Part 1 and to apply additional limitations to quantification such as a size limitation.



**Figure 8.** Macro Part 2. Starting window of Part 2 lightly differs according to the choice of threshold method made in Part 1. The one presented here corresponds to option “B. Manual determination”. If option “A. Using an algorithm” was chosen, the window would start by asking which algorithm should be used. If any of options “C”, “D”, “E”, or “F” were chosen, the window would start by asking which unique threshold value should be used. The following parameters (how to retrieve the results, objects’ size, objects’ circularity and range of series to analyze) are always present. Once chosen, those parameters are recorded in the Log window which content is ultimately saved in a .txt file.

Depending on the study, the quantification needed can be more or less detailed. Indeed, one might need to compare objects size within an image whereas in our case the total stained area per image is sufficient. Both options are available on the macro and are not mutually exclusive. Once these settings are carried out, images are processed as in

Part 1 except that the chosen threshold is applied. After thresholding, Part 2 will quantify but quantification can be restricted to a region of interest. This region can be drawn on the image of the channel indicated in the Part 1 window. If the thresholding is carried out using an algorithm, the zone selection is carried out before running the algorithm. Indeed, this prevents a highly illuminated artifact outside of the region of interest to weigh in the threshold value determination by the algorithm. To allow a wide range of application for this macro, we have chosen to ask the “Analyze Particles” function to quantify all the possible readouts.

#### 4. Conclusions

Apoptosis quantification in a tissue is usually indirect as it generally relies on imaging techniques. There are many ways to analyze an image and once a readout has been chosen, many processing protocols are possible. Here, we describe a semi-automatic protocol running on Fiji for quantification of apoptosis on *Drosophila* wing imaginal discs after TUNEL or activated-caspase labelings. During the development of this protocol, we paid particular attention to the weight of specific steps to obtain a realistic segmentation, which underlies an accurate quantification. As with many image processing protocols, determination of the threshold for binarization turned out to be a critical step. In our case, none of the algorithms we tested was satisfying to determine relevant thresholds.

We also considered using the same threshold value to process several images but, in the end, the best option for our data, was to use a specific threshold manually determined for each image. Indeed, this method proved to carry out a proper segmentation for all images resulting in valid quantification and subsequent detection of biological effects. Even if one could be concerned about the bias that might be induced by this approach, the bias is in fact limited as we observed that threshold values obtained by experimenters are actually highly consistent both for a given experimenter and between experimenters. Moreover, an appropriate processing of the images can facilitate this determination of a threshold value. In this sense, association of a “Median Filter” with a radius of 1 and a Max Intensity z-projection proved to be highly efficient. It would also be interesting to try the “Sum Slices” projection that adds up all pixels’ intensity of a z-axis which should enhance contrast even more (particularly after a median filter) and thus facilitate threshold determination.

The protocol presented here should not affect other readouts available in the “Analyze Particles” function such as: bounding rectangle, shape descriptors, centroid, perimeter, Feret’s diameter, or stack position. Furthermore, we kept the options implemented in CASQITO macro rather open offering a possible use of this tool for a great variety of readouts, stainings, and biological questions.

**Supplementary Materials:** The following are available online at <https://www.mdpi.com/article/10.3390/biom11101523/s1>, Figure S1: Individual experimenters’ choice of threshold values, Figure S2: Statistical analysis of threshold values consistency, Figure S3: Pairwise comparison of thresholding methods, Supplementary information: User manual of the macro.

**Author Contributions:** Conceptualization, J.d.N., J.C. and I.G.; methodology, J.d.N., J.C. and I.G.; software, J.d.N.; validation, J.d.N., M.H., J.C. and I.G.; formal analysis, J.d.N. and J.C.; investigation, J.d.N. and M.H.; writing—original draft preparation, J.d.N.; writing—review and editing, J.d.N., M.H., J.C. and I.G.; visualization, J.d.N., J.C. and I.G.; supervision, J.C. and I.G.; project administration, I.G.; funding acquisition, I.G. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by UVSQ, EPHE and private funding collected by the UVSQ Foundation.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** The data that support the findings of this study are available from the corresponding authors, I.G. and J.C., upon reasonable request.

**Acknowledgments:** We thank Vincent Scarcelli for his participation to thresholds determination, Aude Jobart Malfait for helpful discussions, Christine Wintz for technical assistance and Vincent Rincheval for his critical reading of the manuscript. Confocal microscopy was performed at the CYMAGES imaging facility (UVSQ/University of Paris-Saclay).

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

## References

1. Nirmala, J.G.; Lopus, M. Cell Death Mechanisms in Eukaryotes. *Cell Biol. Toxicol.* **2020**, *36*, 145–164. [CrossRef] [PubMed]
2. Duan, W.R.; Garner, D.S.; Williams, S.D.; Funckes-Shippy, C.L.; Spath, I.S.; Blomme, E.A. Comparison of Immunohistochemistry for Activated Caspase-3 and Cleaved Cytokeratin 18 with the TUNEL Method for Quantification of Apoptosis in Histological Sections of PC-3 Subcutaneous Xenografts. *J. Pathol.* **2003**, *199*, 221–228. [CrossRef] [PubMed]
3. Li, M.; Sun, S.; Priest, J.; Bi, X.; Fan, Y. Characterization of TNF-Induced Cell Death in *Drosophila* Reveals Caspase- and JNK-Dependent Necrosis and Its Role in Tumor Suppression. *Cell Death Dis* **2019**, *10*, 1–14. [CrossRef] [PubMed]
4. Courtoy, G.E.; Donnez, J.; Marbaix, E.; Dolmans, M.-M. In Vivo Mechanisms of Uterine Myoma Volume Reduction with Ulipristal Acetate Treatment. *Fertil. Steril.* **2015**, *104*, 426–434. [CrossRef] [PubMed]
5. Almuedo-Castillo, M.; Crespo, X.; Seebeck, F.; Bartscherer, K.; Salò, E.; Adell, T. JNK Controls the Onset of Mitosis in Planarian Stem Cells and Triggers Apoptotic Cell Death Required for Regeneration and Remodeling. *PLoS Genet.* **2014**, *10*, e1004400. [CrossRef] [PubMed]
6. Huang, Y.; Lai, H.; Xu, H.; Wu, W.; Lai, X.; Ho, G.; Ma, L.; Chen, Y. Impact of Perinatal Systemic Hypoxic–Ischemic Injury on the Brain of Male Offspring Rats: An Improved Model of Neonatal Hypoxic–Ischemic Encephalopathy in Early Preterm Newborns. *PLoS ONE* **2013**, *8*, e82502. [CrossRef]
7. Jadid, K.D.; Davidsson, J.; Lidin, E.; Hånell, A.; Angéria, M.; Mathiesen, T.; Risling, M.; Günther, M. COX-2 Inhibition by Diclofenac Is Associated With Decreased Apoptosis and Lesion Area After Experimental Focal Penetrating Traumatic Brain Injury in Rats. *Front Neurol.* **2019**, *10*, 811. [CrossRef] [PubMed]
8. Daniel, B.; DeCoster, M.A. Quantification of SPLA2-Induced Early and Late Apoptosis Changes in Neuronal Cell Cultures Using Combined TUNEL and DAPI Staining. *Brain Res. Protoc.* **2004**, *13*, 144–150. [CrossRef] [PubMed]
9. Ahmadpour, S.; Heravi, M.A.Y. Quantification of TUNEL Assay in Hippocampus of Diabetic Rats by MAT LAB: Comparison with Stereological Method. *J. Clin. Exp. Pathol.* **2012**, *2*, 1–3. [CrossRef]
10. Schindelin, J.; Arganda-Carreras, I.; Frise, E.; Kaynig, V.; Longair, M.; Pietzsch, T.; Preibisch, S.; Rueden, C.; Saalfeld, S.; Schmid, B.; et al. Fiji: An Open-Source Platform for Biological-Image Analysis. *Nat. Methods* **2012**, *9*, 676–682. [CrossRef] [PubMed]
11. Schneider, C.A.; Rasband, W.S.; Eliceiri, K.W. NIH Image to ImageJ: 25 Years of Image Analysis. *Nat. Methods* **2012**, *9*, 671–675. [CrossRef] [PubMed]
12. Rasband, W.S. ImageJ, U. S. National Institutes of Health, Bethesda, Maryland, USA. 1997–2018. Available online: <https://imagej.nih.gov/ij/> (accessed on 10 October 2021).
13. Milet, C.; Rincheval-Arnold, A.; Mignotte, B.; Guénal, I. The *Drosophila* Retinoblastoma Protein Induces Apoptosis in Proliferating but Not in Post-Mitotic Cells. *Cell Cycle* **2010**, *9*, 97–103. [CrossRef] [PubMed]
14. Clavier, A.; Ruby, V.; Rincheval-Arnold, A.; Mignotte, B.; Guénal, I. The *Drosophila* Retinoblastoma Protein, Rbf1, Induces a Drp1- and Drp1-Dependent Mitochondrial Apoptosis. *J. Cell Sci.* **2015**, *128*, 3239–3249. [CrossRef] [PubMed]
15. Otsu, N. A Threshold Selection Method from Gray-Level Histograms. *IEEE Trans. Syst. Man Cybern.* **1979**, *9*, 62–66. [CrossRef]
16. Tsai, W.-H. Moment-Preserving Thresholding: A New Approach. *Comput. Vis. Graph. Image Process.* **1985**, *29*, 377–393. [CrossRef]



## Annexe II : Script d'analyses de données RNA-seq, du contrôle qualité au comptage des reads pour l'échantillon CD8\_1.

```
#Trimming/QC des reads
fastp -h CD8_1_qc4.html \
-i ../fastq_compressed/220704_NB501850_A_L1-4_AVWR-1_R1.fastq.gz \
-I ../fastq_compressed/220704_NB501850_A_L1-4_AVWR-1_R2.fastq.gz \
-o CD8_1_R1_trimmed.fastq.gz \
-O CD8_1_R2_trimmed.fastq.gz

#Création du génome
STAR \
--runMode genomeGenerate \
--genomeSAindexNbases 12 \
--genomeDir ./analysis2/genome_assem/gen \
--genomeFastaFiles ./analysis2/genome_assem/seq/dm6chr2L.fa
./analysis2/genome_assem/seq/dm6chr2R.fa ./analysis2/genome_assem/seq/dm6chr3L.fa
./analysis2/genome_assem/seq/dm6chr3R.fa ./analysis2/genome_assem/seq/dm6chr4.fa
./analysis2/genome_assem/seq/dm6chrmito.fa ./analysis2/genome_assem/seq/dm6chrY.fa
./analysis2/genome_assem/seq/dm6chrX.fa \
--sjdbGTFfile ./analysis2/genome_assem/dm6.gtf \
--sjdbOverhang 74

#Alignement avec STAR
STAR \
--genomeDir ../analysis2/genome_assem/gen \
--readFilesIn ./trimmed_qc/CD8_1_R1_trimmed.fastq.gz
./trimmed_qc/CD8_1_R2_trimmed.fastq.gz \
--readFilesCommand zcat \
--outSAMtype BAM SortedByCoordinate \
--outFileNamePrefix CD8_1_aligned

#Vérification du run : cat AVWR1_aligneLog.final.out

#Création index
samtools index ./picard/CD8_1_sorted_picard.bam

#Retirer les duplicats
java -jar ../../withoutduplicates/quality/picard.jar MarkDuplicates -
REMOVE_DUPLICATES=true \
-I ./STAR/CD8_1_alignedAligned.sortedByCoord.out.bam \
-O CD8_1_sorted_picard.bam \
```

## ANNEXES

```
-M CD8_1_marked_dup_metrics.txt
```

```
#Voir le nombre de reads restants
```

```
samtools flagstat ./picard/CD8_1_sorted_picard.bam
```

```
#Conversion des BAM en raw counts (HTseq)
```

```
htseq-count -f bam -r pos -s no ./picard/CD8_1_sorted_picard.bam
```

```
../Drosophila_melanogaster.BDGP6.32.107.gtf.gz >
```

```
CD8_1_filtered_sorted_counts_withoutdup.counts
```

## Annexe III : Liste des gènes identifiés comme activés sans retrait des duplicats.

Les gènes sont considérés activés quand la p-value ajustée (padj) est inférieure à 0,05 et le log2 Fold Change (log2FC) est supérieur à 1,5. Les gènes grisés ne sont pas communs aux listes avec et sans retrait des duplicats.

Identifiant Flybase	Nom du gène	Chromosome	padj	log2FC
FBgn0267763	lncRNA:CR460 94	3R	2,76E-14	9,282
FBgn0003863	alphaTry	2R	2,92E-09	7,93
FBgn0262816	CG43187	2R	2,21E-08	7,84
FBgn0013772	Cyp6a8	2R	2,70E-08	7,624
FBgn0262480	CG43070	2R	9,31E-08	7,432
FBgn0262479	CG43069	2R	1,14E-10	7,23
FBgn0037436	Hr83	3R	1,36E-27	7,099
FBgn0030836	CG8664	X	8,35E-26	7,096
FBgn0039873	Smt	3R	2,76E-07	7,029
FBgn0265912	lncRNA:CR447 01	2L	4,17E-07	7,013
FBgn0034899	CG13560	2R	5,58E-07	6,981
FBgn0031579	CG15422	2L	6,43E-07	6,949
FBgn0031010	CG8028	X	5,74E-06	6,759
FBgn0052475	mthl8	3L	5,46E-09	6,758
FBgn0033826	CG4734	2R	7,90E-08	6,646
FBgn0039811	CG15550	3R	2,91E-06	6,63
FBgn0039474	CG6283	3R	2,89E-06	6,576
FBgn0004429	LysP	3L	1,38E-05	6,553
FBgn0264437	lncRNA:CR438 55	2L	1,46E-29	6,476
FBgn0034413	CG15115	2R	1,91E-05	6,412
FBgn0036712	brv2	3L	3,35E-05	6,381
FBgn0032609	CG13280	2L	1,19E-41	6,377
FBgn0035847	mthl7	3L	1,17E-05	6,364
FBgn0053282	CG33282	2L	2,03E-05	6,34
FBgn0034294	Muc55B	2R	1,84E-05	6,336
FBgn0035638	Tektin-C	3L	1,30E-78	6,21
FBgn0010294	ng2	X	4,65E-05	6,17
FBgn0267230	lncRNA:CR456 70	3L	3,59E-05	6,169
FBgn0262481	CG43071	2R	1,38E-05	6,124
FBgn0040261	Ugt37E1	2L	6,48E-05	6,099

## ANNEXES

FBgn0038354	CG5404	3R	7,71E-05	6,092
FBgn0038516	P5cr-2	3R	7,71E-05	6,089
FBgn0032088	CG13102	2L	0,00011521	5,992
FBgn0031520	CG8837	2L	0,00010604	5,944
FBgn0015714	Cyp6a17	2R	2,05E-25	5,909
FBgn0031522	CG3285	2L	0,00015808	5,895
FBgn0028514	CG4793	2L	8,71E-05	5,852
FBgn0031412	CG16995	2L	0,00024565	5,844
FBgn0002940	ninaE	3R	0,00031693	5,788
FBgn0053301	CG33301	2L	6,64E-05	5,782
FBgn0036877	CG9452	3L	0,00034827	5,696
FBgn0260971	CG42594	X	0,00035158	5,677
FBgn0261505	CR42653	2R	0,00030267	5,676
FBgn0038717	CG17751	3R	0,00031505	5,669
FBgn0040262	Ugt37C2	2L	0,00040257	5,657
FBgn0033355	CG13748	2R	0,00044313	5,651
FBgn0045761	CHKov1	3R	3,27E-07	5,54
FBgn0036713	Mip	3L	5,37E-11	5,458
FBgn0262683	CG43153	2L	2,79E-06	5,422
FBgn0039203	CG13618	3R	1,09E-06	5,286
FBgn0259237	CG42335	3R	0,001138	5,222
FBgn0264296	CG43774	2L	0,00140671	5,195
FBgn0035192	galene	3L	0,00146769	5,133
FBgn0035750	CG14826	3L	0,00142926	5,125
FBgn0031801	CG9498	2L	0,000447	5,015
FBgn0033728	Cpr49Ae	2R	0,00189193	5,012
FBgn0085330	CG34301	3R	1,19E-17	4,964
FBgn0031514	OtopLb	2L	4,02E-06	4,916
FBgn0261682	CG42729	3L	0,00288996	4,914
FBgn0052564	CG32564	X	3,19E-10	4,908
FBgn0024943	PIP82	X	7,03E-06	4,791
FBgn0034292	CG5767	2R	5,10E-20	4,777
FBgn0259736	CG42390	3R	1,20E-09	4,727
FBgn0038752	CG4462	3R	3,61E-20	4,695
FBgn0054026	CG34026	X	0,00394952	4,67
FBgn0028950	CG15255	2L	0,001138	4,64
FBgn0051219	CG31219	3R	1,04E-10	4,635
FBgn0036361	CG10154	3L	0,00465636	4,611
FBgn0050462	ste24c	2R	0,00366651	4,576
FBgn0028856	CG18063	2L	0,00441506	4,451
FBgn0039519	Cyp6a18	3R	0,00018178	4,435
FBgn0037974	CG12224	3R	1,61E-10	4,429

FBgn0259738	CG42392	2R	0,00546217	4,346
FBgn0266227	lncRNA:CR449 22	3R	0,00031505	4,342
FBgn0031654	Jon25Bii	2L	1,57E-08	4,336
FBgn0032283	CG7296	2L	0,00533831	4,307
FBgn0039307	CR13656	3R	0,00551619	4,304
FBgn0267730	lncRNA:CR460 61	3R	1,07E-05	4,283
FBgn0030257	SmydA-4	X	0,00240203	4,281
FBgn0038960	CG13855	3R	1,57E-05	4,21
FBgn0053128	CG33128	2L	2,72E-15	4,149
FBgn0037683	CG18473	3R	2,99E-15	4,144
FBgn0015799	Rbf	X	6,75E-73	4,127
FBgn0004240	DptA	2R	3,28E-05	4,107
FBgn0000277	CecA2	3R	5,10E-20	4,105
FBgn0051776	CG31776	2L	0,00733835	4,095
FBgn0038750	CG4465	3R	6,50E-09	4,094
FBgn0010357	betaTry	2R	9,05E-11	4,091
FBgn0250815	Jon65Aiv	3L	1,98E-20	4,049
FBgn0050203	CG30203	2R	8,85E-15	4,043
FBgn0053462	CG33462	2R	1,19E-07	4,018
FBgn0013995	Calx	3R	3,57E-11	3,954
FBgn0040759	CG13177	2R	1,37E-06	3,942
FBgn0040299	Myo28B1	2L	0,00011128	3,931
FBgn0034712	Alp8	2R	1,14E-14	3,867
FBgn0036264	CG11529	3L	0,0016933	3,859
FBgn0038966	pinta	3R	1,85E-41	3,859
FBgn0034010	CG8157	2R	5,21E-06	3,825
FBgn0034159	CG15615	2R	0,001138	3,825
FBgn0033748	vis	2R	9,77E-07	3,77
FBgn0039085	Ugt303B3	3R	0,00988393	3,766
FBgn0036952	CG6933	3L	1,66E-09	3,751
FBgn0025643	CG3588	X	0,00078875	3,735
FBgn0034972	CG10339	2R	0,00589248	3,725
FBgn0033980	Cyp6a20	2R	8,35E-26	3,714
FBgn0020906	Jon25Bi	2L	4,89E-12	3,704
FBgn0051100	CG31100	3R	2,36E-06	3,667
FBgn0053459	CG33459	2R	0,00142926	3,663
FBgn0039818	CG11318	3R	1,41E-14	3,659
FBgn0011822	Pgcl	X	0,01089924	3,657
FBgn0036106	CG6409	3L	0,00057181	3,643
FBgn0039328	CHKov2	3R	2,91E-24	3,603

## ANNEXES

FBgn0259236	comm3	3L	0,00213372	3,603
FBgn0036940	obst-J	3L	0,00073091	3,599
FBgn0036563	CG13075	3L	0,00039153	3,557
FBgn0044048	Ilp5	3L	0,00493712	3,54
FBgn0015008	Actn3	3R	0,01275747	3,535
FBgn0004580	Cbp53E	2R	1,53E-23	3,519
FBgn0010425	epsilonTry	2R	1,38E-05	3,51
FBgn0031969	pes	2L	6,16E-44	3,5
FBgn0033885	DJ-1alpha	2R	0,00442621	3,472
FBgn0010042	GstD6	3R	0,01013622	3,447
FBgn0039787	CG9702	3R	0,00627326	3,44
FBgn0032495	CG16820	2L	9,19E-05	3,433
FBgn0028374	Hug	3R	0,00182526	3,417
FBgn0031561	IM33	2L	5,83E-08	3,416
FBgn0032282	CG7299	2L	1,34E-06	3,414
FBgn0033978	Cyp6a23	2R	5,89E-10	3,407
FBgn0036949	CG7290	3L	9,15E-06	3,406
FBgn0035187	Trh	3L	0,00072977	3,402
FBgn0039081	Irk2	3R	1,61E-10	3,398
FBgn0033720	CG13160	2R	0,00206547	3,341
FBgn0033733	CG8834	2R	0,00713115	3,333
FBgn0050360	Mal-A6	2R	0,01573367	3,321
FBgn0036226	CG7252	3L	0,00843652	3,319
FBgn0037975	CG3397	3R	0,00855555	3,298
FBgn0263390	Nha2	3R	3,06E-07	3,292
FBgn0034221	CG10764	2R	1,66E-05	3,278
FBgn0034052	CG8299	2R	0,01442645	3,276
FBgn0035665	Jon65Aiii	3L	1,66E-09	3,264
FBgn0053460	CG33460	2R	0,00855555	3,249
FBgn0038337	CG6125	3R	2,09E-07	3,242
FBgn0038241	CG8087	3R	0,00580224	3,235
FBgn0038242	CG14852	3R	0,01479782	3,222
FBgn0032911	tadr	2L	0,00027297	3,198
FBgn0033302	Cyp6a14	2R	4,86E-05	3,17
FBgn0050090	CG30090	2R	6,87E-10	3,167
FBgn0030040	CG15347	X	8,53E-07	3,157
FBgn0050287	CG30287	2R	0,00019106	3,154
FBgn0040260	Ugt37D1	2L	9,50E-07	3,145
FBgn0035359	CG1143	3L	1,31E-09	3,142
FBgn0033541	CG12934	2R	1,95E-06	3,141
FBgn0267987	lncRNA:CR462 54	2L	0,00719633	3,108

FBgn0015575	alpha-Est7	3R	0,00957817	3,105
FBgn0034289	CG10910	2R	1,52E-07	3,098
FBgn0039124	tbrd-1	3R	0,01896917	3,075
FBgn0039313	CG11892	3R	8,01E-07	3,025
FBgn0039297	CG11852	3R	0,0019004	3,013
FBgn0263234	Phae1	2L	0,00435254	3,002
FBgn0263776	CG43693	3L	1,81E-05	2,99
FBgn0020376	Sr-CIII	2L	0,013052	2,989
FBgn0035193	CG9192	3L	4,40E-18	2,982
FBgn0037238	CG1090	3R	5,48E-05	2,957
FBgn0028533	CG7953	2L	2,27E-10	2,952
FBgn0031560	CG16713	2L	0,00262377	2,932
FBgn0033469	CG12133	2R	0,02081767	2,916
FBgn0037996	CG4830	3R	0,00266816	2,887
FBgn0052302	CG32302	3L	0,00018284	2,871
FBgn0266446	CG45076	3R	0,01963272	2,862
FBgn0039073	CG4408	3R	1,76E-05	2,851
FBgn0039101	CG16710	3R	3,33E-05	2,838
FBgn0040959	Peritrophin-15a	2L	1,26E-05	2,815
FBgn0029810	CG12239	X	3,54E-09	2,814
FBgn0036381	CG8745	3L	6,38E-05	2,813
FBgn0028699	Rh50	3L	0,01636646	2,802
FBgn0028920	CG8997	2L	8,13E-08	2,793
FBgn0046302	CG10650	2L	0,00236265	2,783
FBgn0052284	CG32284	3L	0,01628088	2,775
FBgn0011555	thetaTry	2R	0,01072347	2,746
FBgn0038846	CG5697	3R	0,00221112	2,737
FBgn0037724	Fst	3R	9,31E-14	2,725
FBgn0051087	CG31087	3R	0,01718333	2,713
FBgn0039480	Cpr97Ea	3R	7,26E-06	2,712
FBgn0038973	Pebp1	3R	1,85E-06	2,697
FBgn0016123	Alp4	3R	1,89E-06	2,691
FBgn0030574	sbm	X	3,90E-16	2,69
FBgn0031580	CG15423	2L	0,00039207	2,69
FBgn0002522	lab	3R	0,02618618	2,686
FBgn0035293	CG5687	3L	4,15E-16	2,675
FBgn0038260	CG14855	3R	0,00456734	2,656
FBgn0051789	CG31789	2L	0,00069326	2,644
FBgn0034710	Alp7	2R	1,42E-06	2,629
FBgn0034156	CG5348	2R	0,0236908	2,628
FBgn0034716	Oatp58Dc	2R	0,00394952	2,611

## ANNEXES

FBgn0054031	CG34031	3L	0,02791033	2,606
FBgn0030097	CG12115	X	0,00528809	2,604
FBgn0034275	CG5002	2R	1,50E-08	2,602
FBgn0287618	asRNA:CR4648 7	2R	0,02791033	2,598
FBgn0027929	NimB1	2L	0,02560208	2,589
FBgn0000406	Cyt-b5-r	2L	0,00039153	2,574
FBgn0040060	yip7	3L	1,55E-05	2,573
FBgn0038465	Irc	3R	5,44E-06	2,572
FBgn0036931	CG14183	3L	1,24E-05	2,564
FBgn0042104	CG18747	3R	0,0186279	2,547
FBgn0032048	Dh31	2L	0,00055704	2,539
FBgn0028400	Syt4	3R	3,27E-20	2,537
FBgn0039325	CG10560	3R	0,00946779	2,537
FBgn0039915	Gat	4	5,21E-06	2,533
FBgn0034406	Jheh3	2R	0,02784108	2,518
FBgn0040364	CG11378	X	0,00038456	2,504
FBgn0013763	Idgf6	2R	6,94E-22	2,48
FBgn0036752	Adgf-A	3L	7,30E-09	2,47
FBgn0032373	Vha100-5	2L	0,02362804	2,451
FBgn0037895	CG6723	3R	0,01628088	2,446
FBgn0031703	CG12512	2L	0,02075162	2,431
FBgn0016650	Lgr1	3R	0,00843652	2,43
FBgn0031089	CG9572	X	0,00024627	2,416
FBgn0004430	LysS	3L	0,00530457	2,41
FBgn0038753	CG4459	3R	0,0145359	2,407
FBgn0034935	Orcokinin	2R	0,00080664	2,379
FBgn0034501	CG13868	2R	4,65E-24	2,337
FBgn0031313	CG5080	2L	1,03E-07	2,33
FBgn0030999	Mur18B	X	3,12E-05	2,321
FBgn0034428	AANATL6	2R	0,03076629	2,313
FBgn0004428	LysE	3L	0,0351148	2,312
FBgn0033730	Cpr49Ag	2R	0,01243876	2,306
FBgn0034901	CG11300	2R	0,00267809	2,278
FBgn0038959	CG13856	3R	0,03033732	2,265
FBgn0264907	CG44098	3R	1,36E-06	2,262
FBgn0041156	exex	3L	0,03333663	2,254
FBgn0266211	CR44906	2L	0,01628088	2,237
FBgn0031747	CG9021	2L	0,00598807	2,224
FBgn0038941	CG7080	3R	0,03523005	2,214
FBgn0265836	CG44625	2R	0,03791253	2,205
FBgn0002789	Mp20	2R	0,02588718	2,197



FBgn0259683	Ir40a	2L	0,02949805	2,186
<b>FBgn0267408</b>	<b>AOX1</b>	<b>3R</b>	<b>1,90E-05</b>	<b>2,185</b>
FBgn0262366	CG43064	3L	0,03791253	2,179
<b>FBgn0010549</b>	<b>l(2)03659</b>	<b>2R</b>	<b>3,32E-12</b>	<b>2,174</b>
FBgn0035010	CG13579	2R	0,01750323	2,169
<b>FBgn0032264</b>	<b>Lip4</b>	<b>2L</b>	<b>3,19E-09</b>	<b>2,168</b>
FBgn0034117	CG7997	2R	0,00221474	2,149
FBgn0265375	lncRNA:CR443 16	2R	0,02667706	2,14
FBgn0036474	Or71a	3L	0,0402333	2,122
<b>FBgn0012037</b>	<b>Ance</b>	<b>2L</b>	<b>2,40E-45</b>	<b>2,121</b>
FBgn0030837	CG8661	X	0,01390714	2,113
FBgn0045038	Proc	2L	0,00011805	2,111
FBgn0028490	CG31705	2L	1,06E-27	2,107
FBgn0039051	CG17109	3R	0,0004536	2,107
FBgn0051098	CG31098	3R	2,27E-05	2,107
FBgn0038391	GATAe	3R	0,03327192	2,102
FBgn0024315	Picot	2R	0,00010159	2,089
FBgn0003308	ry	3R	0,00645639	2,078
FBgn0264343	CG43799	2L	0,03630125	2,078
<b>FBgn0035968</b>	<b>Slc45-1</b>	<b>3L</b>	<b>0,00798757</b>	<b>2,077</b>
FBgn0002571	Mal-A3	2R	0,02490472	2,071
<b>FBgn0003896</b>	<b>tup</b>	<b>2L</b>	<b>9,66E-07</b>	<b>2,066</b>
FBgn0031562	CG3604	2L	0,02784108	2,064
FBgn0037782	Npc2d	3R	0,04229105	2,064
FBgn0015591	AstA	3R	0,00957817	2,063
FBgn0038897	CG5849	3R	0,00798757	2,063
FBgn0267181	asRNA:CR4562 1	X	0,04260457	2,063
<b>FBgn0262945</b>	<b>lncRNA:CR432 64</b>	<b>X</b>	<b>1,80E-05</b>	<b>2,058</b>
FBgn0036179	CG7368	3L	0,02289536	2,049
<b>FBgn0041183</b>	<b>Tep1</b>	<b>2L</b>	<b>0,02791033</b>	<b>2,047</b>
FBgn0033304	Cyp6a13	2R	0,02972239	2,034
<b>FBgn0039756</b>	<b>CG9743</b>	<b>3R</b>	<b>3,41E-10</b>	<b>2,034</b>
FBgn0261453	scaRNA:MeU2- C41	3R	0,04383056	2,033
FBgn0038705	CG11626	3R	0,04399697	2,027
FBgn0038647	CG14302	3R	0,00116921	2,024
<b>FBgn0034490</b>	<b>CG9864</b>	<b>2R</b>	<b>0,00024744</b>	<b>2,02</b>
FBgn0030331	CG15221	X	0,04129547	2,018
<b>FBgn0037000</b>	<b>ZnT77C</b>	<b>3L</b>	<b>4,73E-05</b>	<b>2,015</b>

## ANNEXES

FBgn0020414	ldgf3	2L	3,06E-07	2,009
FBgn0051431	CG31431	3R	0,00019473	1,99
FBgn0039486	CAH9	3R	0,03384246	1,986
FBgn0034076	Jhedup	2R	0,04596704	1,973
FBgn0036181	Muc68Ca	3L	0,00353176	1,96
FBgn0051547	NKCC	3R	2,00E-07	1,956
FBgn0034288	CG5084	2R	0,03339592	1,941
FBgn0012344	Dh44	3R	0,00189075	1,925
FBgn0038343	Trissin	3R	0,04823491	1,925
FBgn0036831	CG6839	3L	0,03658767	1,92
FBgn0037548	CG7900	3R	2,09E-10	1,92
FBgn0024989	CG3777	X	6,81E-20	1,907
FBgn0050148	CG30148	2R	0,02476494	1,9
FBgn0031461	daw	2L	1,18E-09	1,894
FBgn0260006	drd	X	3,30E-05	1,882
FBgn0262985	CG43293	3L	0,04992515	1,871
FBgn0036927	Gabat	3L	7,71E-05	1,866
FBgn0031910	CG15818	2L	0,03790084	1,863
FBgn0015011	AhcyL2	3R	9,66E-06	1,855
FBgn0085407	Pvf3	2L	1,48E-07	1,854
FBgn0035969	CG4476	3L	0,02939852	1,851
FBgn0033448	hebe	2R	0,00292577	1,844
FBgn0263111	cac	X	9,53E-13	1,833
FBgn0000114	bru1	2L	0,02675763	1,831
FBgn0039312	CG10514	3R	0,01044706	1,829
FBgn0265048	cv-d	3R	2,89E-06	1,829
FBgn0028518	CG18480	2L	0,02571283	1,827
FBgn0034301	CG5756	2R	8,66E-06	1,821
FBgn0026315	Ugt35A1	3R	0,00267809	1,814
FBgn0037387	CG1213	3R	0,03207718	1,811
FBgn0042106	CG18754	3R	0,00404378	1,808
FBgn0034709	Swim	2R	7,88E-06	1,791
FBgn0262570	CG43110	2R	0,02939852	1,79
FBgn0034335	GstE1	2R	0,00792264	1,772
FBgn0031306	CG4577	2L	0,00023055	1,77
FBgn0261625	GLS	2R	1,32E-08	1,753
FBgn0044050	llp3	3L	0,03243165	1,745
FBgn0053542	upd3	X	0,00050161	1,743
FBgn0039109	CG10365	3R	2,27E-07	1,734
FBgn0037432	CG10298	3R	0,04205879	1,725
FBgn0031678	Nepl3	2L	9,53E-13	1,723
FBgn0267692	lncRNA:CR46028	3R	0,00838755	1,718

FBgn0038407	CG6126	3R	0,00374431	1,715
FBgn0032022	CG14275	2L	0,01058122	1,711
FBgn0036394	CG9040	3L	0,04048067	1,703
FBgn0003117	pnr	3R	1,59E-05	1,696
FBgn0030912	CG6023	X	1,36E-05	1,687
FBgn0038170	CG14367	3R	0,04236694	1,661
FBgn0032304	CG17134	2L	0,0296701	1,647
FBgn0000659	fkx	3R	0,00322485	1,641
FBgn0052373	CG32373	3L	2,94E-08	1,639
FBgn0036046	llp2	3L	0,00625823	1,627
FBgn0037431	CG17917	3R	0,03635346	1,622
FBgn0053513	Nmdar2	X	0,01118534	1,62
FBgn0027556	CG4928	X	0,00026667	1,614
FBgn0030310	PGRP-SA	X	0,00051223	1,608
FBgn0259140	CG42255	3L	0,03804459	1,604
FBgn0034279	CG18635	2R	0,04928428	1,594
FBgn0002723	Met	X	4,43E-11	1,582
FBgn0045852	ham	2L	0,0090178	1,581
FBgn0040091	Ugt317A1	2R	0,00036003	1,577
FBgn0052198	CG32198	3L	0,02081605	1,575
FBgn0039777	Jon99Fii	3R	0,01697056	1,569
FBgn0287423	Nplp2	3L	0,04554153	1,553
FBgn0035779	CG8562	3L	0,04905491	1,532
FBgn0263118	tx	3R	0,00057245	1,528
FBgn0050339	CG30339	2R	0,0092558	1,527
FBgn0010387	Acbp2	3L	0,03243165	1,521
FBgn0037503	CG14598	3R	0,00202682	1,515
FBgn0052512	CG32512	X	5,53E-07	1,512
FBgn0063494	GstE6	2R	0,00262352	1,51
FBgn0028704	Nckx30C	2L	0,00076371	1,507
FBgn0085359	CG34330	X	0,02275508	1,502

## Annexe IV : Liste des gènes identifiés comme réprimés sans retrait des duplicats.

Les gènes sont considérés réprimés quand la p-value ajustée (padj) est inférieure à 0,05 et le log<sub>2</sub> Fold Change (log<sub>2</sub>FC) est inférieur à -1,5. Les gènes grisés ne sont pas communs aux listes avec et sans retrait des duplicats.

Identifiant Flybase	Nom du gène	Chromosome	padj	log <sub>2</sub> FC
<b>FBgn0030313</b>	Reep1	X	5,81E-23	-9,035
FBgn0015037	Cyp4p1	2R	4,98E-08	-6,975
FBgn0036044	Zasp67	3L	4,87E-16	-6,94
FBgn0033395	Cyp4p2	2R	2,40E-12	-6,916
FBgn0286030	lncRNA:CR463 50	3R	9,66E-06	-6,457
FBgn0035861	CG7213	3L	8,84E-05	-6,098
FBgn0040965	lncRNA:CR131 30	2L	3,97E-05	-5,622
FBgn0034263	CG10934	2R	0,00027972	-5,521
FBgn0034352	CG17669	2R	0,00625823	-4,297
FBgn0266649	lncRNA:CR451 56	3L	0,00852419	-3,821
FBgn0035543	CG15020	3L	0,00857849	-3,444
FBgn0033521	CG12896	2R	3,98E-13	-3,29
FBgn0038018	Tim17a1	3R	0,00618352	-3,093
FBgn0263219	Dscam4	3L	1,36E-29	-3,06
FBgn0036492	Best3	3L	0,02297719	-2,826
FBgn0000449	dib	3L	6,43E-07	-2,519
FBgn0058469	lncRNA:CR404 69	X	2,57E-06	-2,399
FBgn0034605	Ugt49C1	2R	0,00052226	-2,325
FBgn0286204	ich	3R	0,00011149	-2,005
FBgn0039459	IntS12	3R	6,23E-09	-1,927
FBgn0027259	Kmn1	X	5,78E-10	-1,925
FBgn0039795	Spn100A	3R	2,54E-05	-1,835
FBgn0013275	Hsp70Aa	3R	0,0100956	-1,816
FBgn0038339	CG6118	3R	0,01058122	-1,801
FBgn0039411	dysf	3R	8,71E-05	-1,725
FBgn0036677	CG13023	3L	0,00066715	-1,643
FBgn0039170	CG13609	3R	0,00034827	-1,639
FBgn0051785	CG31785	2L	0,03362414	-1,578
FBgn0003996	w	X	4,96E-13	-1,539

## Annexe V : Liste des gènes identifiés comme activés avec retrait des duplicats.

Les gènes sont considérés activés quand la p-value ajustée (padj) est inférieure à 0,05 et le log2 Fold Change (log2FC) est supérieur à 1,5. Les gènes grisés ne sont pas communs aux listes avec et sans retrait des duplicats.

Identifiant Flybase	Nom du gène	Chromosome	padj	log2FC
<b>FBgn0052475</b>	mthl8	3L	2,43E-09	6,753
FBgn0267763	lncRNA:CR46094	3R	1,05E-07	6,342
FBgn0032609	CG13280	2L	3,04E-16	6,295
FBgn0028514	CG4793	2L	7,60E-13	6,229
FBgn0264437	lncRNA:CR43855	2L	4,03E-09	5,958
FBgn0037436	Hr83	3R	2,89E-09	5,939
FBgn0015714	Cyp6a17	2R	1,17E-08	5,763
FBgn0035638	Tektin-C	3L	6,25E-31	5,762
FBgn0045761	CHKov1	3R	1,13E-23	5,739
FBgn0030836	CG8664	X	6,27E-10	5,411
FBgn0262816	CG43187	2R	4,17E-05	5,245
FBgn0039787	CG9702	3R	8,65E-12	5,221
FBgn0034406	Jheh3	2R	1,24E-05	5,196
FBgn0262479	CG43069	2R	2,26E-05	5,042
FBgn0039203	CG13618	3R	3,25E-22	4,901
FBgn0034292	CG5767	2R	8,65E-12	4,895
FBgn0003863	alphaTry	2R	0,00014046	4,892
FBgn0038242	CG14852	3R	0,00031718	4,735
FBgn0051219	CG31219	3R	2,69E-06	4,541
FBgn0034899	CG13560	2R	0,00043827	4,514
FBgn0013772	Cyp6a8	2R	0,00053371	4,464
FBgn0036713	Mip	3L	4,57E-05	4,409
FBgn0262480	CG43070	2R	0,00065366	4,4
FBgn0028699	Rh50	3L	9,17E-06	4,367
FBgn0038752	CG4462	3R	5,53E-10	4,299
FBgn0053460	CG33460	2R	7,92E-14	4,298
FBgn0010357	betaTry	2R	2,44E-07	4,111
FBgn0000277	CecA2	3R	6,20E-12	4,075
FBgn0265912	lncRNA:CR44701	2L	0,00132034	4,029
FBgn0039328	CHKov2	3R	4,99E-18	4,005
FBgn0085330	CG34301	3R	3,31E-06	3,933
FBgn0031579	CG15422	2L	0,00179701	3,919
FBgn0039873	Smtv	3R	0,00164471	3,909

## ANNEXES

FBgn0015799	Rbf	X	2,38E-73	3,885
FBgn0053128	CG33128	2L	2,09E-11	3,84
FBgn0034712	Alp8	2R	1,06E-10	3,834
FBgn0033826	CG4734	2R	0,0018786	3,828
FBgn0038750	CG4465	3R	7,51E-06	3,819
FBgn0027929	NimB1	2L	3,63E-06	3,742
FBgn0020906	Jon25Bi	2L	1,72E-09	3,718
FBgn0031561	IM33	2L	2,38E-07	3,706
FBgn0039670	CG7567	3R	0,00269882	3,706
FBgn0037974	CG12224	3R	9,03E-07	3,683
FBgn0034156	CG5348	2R	0,00111586	3,68
FBgn0004240	DptA	2R	0,00029144	3,658
FBgn0250815	Jon65Aiv	3L	8,93E-13	3,651
FBgn0052564	CG32564	X	0,00018909	3,642
FBgn0036106	CG6409	3L	4,94E-05	3,624
FBgn0037683	CG18473	3R	7,22E-09	3,619
FBgn0035789	mthl6	3L	5,78E-05	3,614
FBgn0013995	Calx	3R	5,18E-08	3,585
FBgn0039474	CG6283	3R	0,0032587	3,546
FBgn0039818	CG11318	3R	2,31E-06	3,521
FBgn0259736	CG42390	3R	0,00024568	3,47
FBgn0033748	vis	2R	2,82E-05	3,419
FBgn0031654	Jon25Bii	2L	6,60E-05	3,418
FBgn0051100	CG31100	3R	1,61E-05	3,393
FBgn0036831	CG6839	3L	6,30E-05	3,391
FBgn0050203	CG30203	2R	2,57E-07	3,383
FBgn0034289	CG10910	2R	2,59E-07	3,378
FBgn0010425	epsilonTry	2R	0,00022261	3,371
FBgn0039811	CG15550	3R	0,00478108	3,332
FBgn0031562	CG3604	2L	1,08E-05	3,325
FBgn0038966	pinta	3R	6,57E-19	3,31
FBgn0035665	Jon65Aiii	3L	1,44E-09	3,307
FBgn0033302	Cyp6a14	2R	7,51E-06	3,288
FBgn0031969	pes	2L	2,90E-26	3,268
FBgn0033980	Cyp6a20	2R	2,68E-10	3,263
FBgn0267730	lncRNA:CR46061	3R	0,00069804	3,25
FBgn0031514	OtopLb	2L	0,00499766	3,242
FBgn0024943	PIP82	X	0,00400811	3,225
FBgn0029810	CG12239	X	2,99E-12	3,219
FBgn0032282	CG7299	2L	0,00034144	3,218
FBgn0050090	CG30090	2R	6,05E-08	3,19
FBgn0263390	Nha2	3R	2,21E-05	3,187

FBgn0040759	CG13177	2R	0,00018574	3,186
FBgn0034221	CG10764	2R	4,43E-06	3,154
FBgn0036952	CG6933	3L	1,06E-05	3,144
FBgn0038337	CG6125	3R	5,26E-05	3,129
FBgn0004580	Cbp53E	2R	3,08E-15	3,092
FBgn0033978	Cyp6a23	2R	1,19E-08	3,07
FBgn0035293	CG5687	3L	6,56E-18	3,055
FBgn0053462	CG33462	2R	0,00160224	2,982
FBgn0051789	CG31789	2L	0,00048737	2,973
FBgn0039081	Irk2	3R	1,62E-05	2,965
FBgn0028533	CG7953	2L	2,79E-07	2,944
FBgn0033730	Cpr49Ag	2R	0,00097646	2,941
FBgn0039313	CG11892	3R	0,00031718	2,938
FBgn0033541	CG12934	2R	0,00019386	2,907
FBgn0035359	CG1143	3L	7,26E-07	2,906
FBgn0034294	Muc55B	2R	0,00929072	2,876
FBgn0262683	CG43153	2L	0,00864231	2,871
FBgn0004429	LysP	3L	0,00972229	2,854
FBgn0025643	CG3588	X	0,0049569	2,85
FBgn0040060	yip7	3L	2,31E-06	2,822
FBgn0016650	Lgr1	3R	0,00248277	2,81
FBgn0044050	Ilp3	3L	0,00735648	2,803
FBgn0035193	CG9192	3L	2,28E-12	2,791
FBgn0052302	CG32302	3L	0,0004549	2,785
FBgn0065098	RNaseMRP:RNA	3R	3,14E-06	2,785
FBgn0028920	CG8997	2L	4,57E-07	2,756
FBgn0034010	CG8157	2R	0,00077962	2,756
FBgn0031010	CG8028	X	0,0107987	2,75
FBgn0267230	lncRNA:CR45670	3L	0,01058886	2,75
FBgn0038846	CG5697	3R	0,00161722	2,734
FBgn0050287	CG30287	2R	0,00107606	2,699
FBgn0034159	CG15615	2R	0,00992456	2,682
FBgn0037724	Fst	3R	4,99E-10	2,682
FBgn0259236	comm3	3L	0,00806795	2,669
FBgn0030574	sbm	X	6,30E-11	2,668
FBgn0031747	CG9021	2L	2,97E-06	2,648
FBgn0016123	Alp4	3R	4,36E-07	2,642
FBgn0263776	CG43693	3L	0,00011197	2,635
FBgn0034275	CG5002	2R	1,97E-07	2,633
FBgn0035847	mthl7	3L	0,01216836	2,631
FBgn0038973	Pebp1	3R	2,24E-05	2,617
FBgn0040260	Ugt37D1	2L	0,00325808	2,589

## ANNEXES

FBgn0004430	LysS	3L	0,00140476	2,562
FBgn0038465	Irc	3R	1,63E-05	2,548
FBgn0262481	CG43071	2R	0,01300652	2,537
FBgn0028400	Syt4	3R	2,18E-16	2,526
FBgn0013763	ldgf6	2R	2,30E-20	2,516
FBgn0030040	CG15347	X	0,00017067	2,513
FBgn0036949	CG7290	3L	0,00055311	2,497
FBgn0267408	AOX1	3R	3,50E-07	2,496
FBgn0015524	otp	2R	0,01029685	2,45
FBgn0032495	CG16820	2L	0,00880718	2,427
FBgn0034716	Oatp58Dc	2R	0,01235062	2,426
FBgn0039101	CG16710	3R	0,00011602	2,419
FBgn0035969	CG4476	3L	0,00830218	2,417
FBgn0010294	ng2	X	0,01566376	2,41
FBgn0031089	CG9572	X	0,00034006	2,397
FBgn0030837	CG8661	X	0,01013441	2,387
FBgn0267692	lncRNA:CR46028	3R	0,00037288	2,387
FBgn0010549	l(2)03659	2R	8,78E-16	2,326
FBgn0034901	CG11300	2R	0,00472911	2,303
FBgn0031313	CG5080	2L	3,38E-06	2,298
FBgn0036752	Adgf-A	3L	8,56E-05	2,292
FBgn0035968	Slc45-1	3L	0,00316602	2,28
FBgn0015011	AhcyL2	3R	1,87E-06	2,275
FBgn0039073	CG4408	3R	0,00565143	2,273
FBgn0032048	Dh31	2L	0,0032587	2,27
FBgn0012344	Dh44	3R	0,00020766	2,267
FBgn0034710	Alp7	2R	0,00071245	2,267
FBgn0039480	Cpr97Ea	3R	0,00191737	2,266
FBgn0032264	Lip4	2L	4,49E-09	2,261
FBgn0034291	CG5770	2R	0,003823	2,246
FBgn0038241	CG8087	3R	0,01837329	2,244
FBgn0264907	CG44098	3R	9,02E-05	2,241
FBgn0030097	CG12115	X	0,01087172	2,205
FBgn0024315	Picot	2R	0,00012958	2,192
FBgn0044048	Ilp5	3L	0,02003214	2,189
FBgn0038845	Alp5	3R	0,00872553	2,184
FBgn0028518	CG18480	2L	0,0176747	2,169
FBgn0050098	CG30098	2R	0,00951368	2,161
FBgn0036927	Gabat	3L	1,83E-07	2,16
FBgn0036931	CG14183	3L	0,004915	2,131
FBgn0032603	CG17928	2L	0,0156484	2,115
FBgn0039297	CG11852	3R	0,01353773	2,113



FBgn0011581	Ms	3R	0,00403415	2,112
FBgn0031580	CG15423	2L	0,00923652	2,107
FBgn0026315	Ugt35A1	3R	0,00179716	2,085
FBgn0085359	CG34330	X	0,00722387	2,053
FBgn0267081	asRNA:CR45525	X	0,02308818	2,05
FBgn0034501	CG13868	2R	3,77E-13	2,034
FBgn0012037	Ance	2L	2,16E-38	2,026
FBgn0028490	CG31705	2L	5,24E-20	2,024
FBgn0037548	CG7900	3R	6,30E-11	2,016
FBgn0260006	drd	X	4,30E-05	2,007
FBgn0041183	Tep1	2L	0,0203476	1,995
FBgn0265048	cv-d	3R	3,41E-08	1,989
FBgn0051098	CG31098	3R	0,00123662	1,98
FBgn0052476	mthl14	3L	0,00158392	1,973
FBgn0032087	CG9568	2L	0,02003214	1,964
FBgn0053282	CG33282	2L	0,02606308	1,961
FBgn0030912	CG6023	X	5,88E-08	1,96
FBgn0038260	CG14855	3R	0,01414189	1,955
FBgn0045038	Proc	2L	0,00410801	1,942
FBgn0031520	CG8837	2L	0,02661742	1,933
FBgn0266095	lncRNA:CR44832	X	0,01795663	1,926
FBgn0020414	ldgf3	2L	1,05E-07	1,915
FBgn0032218	CG5381	2L	0,0172018	1,906
FBgn0030747	CG4301	X	0,01240905	1,902
FBgn0039109	CG10365	3R	2,79E-07	1,902
FBgn0034490	CG9864	2R	0,00371636	1,89
FBgn0036563	CG13075	3L	0,02282812	1,889
FBgn0266227	lncRNA:CR44922	3R	0,02800792	1,879
FBgn0020416	ldgf1	2L	0,00219557	1,868
FBgn0262945	lncRNA:CR43264	X	0,0008792	1,851
FBgn0287423	Nplp2	3L	0,0224375	1,847
FBgn0030999	Mur18B	X	0,00314566	1,84
FBgn0263118	tx	3R	3,27E-05	1,837
FBgn0034709	Swim	2R	2,58E-05	1,827
FBgn0039756	CG9743	3R	1,34E-05	1,824
FBgn0039051	CG17109	3R	0,00473611	1,807
FBgn0085407	Pvf3	2L	5,15E-08	1,793
FBgn0038658	CG14292	3R	0,02661742	1,792
FBgn0261625	GLS	2R	3,95E-08	1,787
FBgn0040364	CG11378	X	0,01071175	1,777
FBgn0000036	nAChRalpha1	3R	0,00375299	1,775
FBgn0263111	cac	X	1,22E-08	1,771

## ANNEXES

FBgn0031678	Nepl3	2L	6,84E-15	1,759
FBgn0037000	ZnT77C	3L	0,00268615	1,757
FBgn0036953	CG17145	3L	0,01867104	1,755
FBgn0264296	CG43774	2L	0,03162904	1,755
FBgn0050339	CG30339	2R	0,0147775	1,747
FBgn0028542	NimB4	2L	0,00928564	1,745
FBgn0024989	CG3777	X	5,66E-17	1,737
FBgn0051431	CG31431	3R	0,00364601	1,73
FBgn0034413	CG15115	2R	0,03269627	1,727
FBgn0038960	CG13855	3R	0,03104512	1,725
FBgn0031461	daw	2L	7,72E-08	1,704
FBgn0003896	tup	2L	2,52E-05	1,702
FBgn0040091	Ugt317A1	2R	9,52E-07	1,685
FBgn0040299	Myo28B1	2L	0,03104512	1,678
FBgn0051547	NKCC	3R	3,96E-06	1,674
FBgn0039312	CG10514	3R	0,02697587	1,656
FBgn0053459	CG33459	2R	0,03220545	1,656
FBgn0037996	CG4830	3R	0,02266406	1,654
FBgn0031560	CG16713	2L	0,02661742	1,645
FBgn0036712	brv2	3L	0,03581257	1,643
FBgn0040261	Ugt37E1	2L	0,03608874	1,641
FBgn0053301	CG33301	2L	0,03635702	1,636
FBgn0038391	GATAe	3R	0,03609761	1,632
FBgn0032022	CG14275	2L	0,01216673	1,627
FBgn0052373	CG32373	3L	5,44E-07	1,626
FBgn0002723	Met	X	5,12E-10	1,615
FBgn0027556	CG4928	X	0,00043827	1,602
FBgn0052198	CG32198	3L	0,02015952	1,578
FBgn0046302	CG10650	2L	0,02638934	1,567
FBgn0039108	CG10232	3R	0,02907609	1,566
FBgn0039927	CG11155	4	7,80E-06	1,557
FBgn0032304	CG17134	2L	0,02810416	1,555
FBgn0033448	hebe	2R	0,0112939	1,546
FBgn0019985	mGluR	4	0,00351044	1,543
FBgn0034935	Orcokinin	2R	0,01919399	1,541
FBgn0034335	GstE1	2R	0,01270093	1,536
FBgn0052512	CG32512	X	1,97E-07	1,526
FBgn0051999	CG31999	4	4,05E-08	1,519
FBgn0037503	CG14598	3R	0,00094293	1,518
FBgn0036690	Ilp8	3L	0,02614116	1,508
FBgn0045852	ham	2L	0,01441377	1,507

## Annexe VI : Liste des gènes identifiés comme réprimés avec retrait des duplicats.

Les gènes sont considérés réprimés quand la p-value ajustée (padj) est inférieure à 0,05 et le log<sub>2</sub> Fold Change (log<sub>2</sub>FC) est inférieur à -1,5. Les gènes grisés ne sont pas communs aux listes avec et sans retrait des duplicats.

Identifiant Flybase	Nom du gène	Chromosome	padj	log <sub>2</sub> FC
FBgn0030313	Reep1	X	6,49E-47	-8,256
FBgn0034352	CG17669	2R	2,14E-22	-7,825
FBgn0015037	Cyp4p1	2R	1,18E-22	-6,828
FBgn0033395	Cyp4p2	2R	1,58E-10	-6,535
FBgn0036044	Zasp67	3L	9,59E-12	-6,014
FBgn0286030	lncRNA:CR463 50	3R	0,00123662	-4,073
FBgn0040965	lncRNA:CR131 30	2L	0,00254779	-3,616
FBgn0035861	CG7213	3L	0,00377425	-3,512
FBgn0039510	CG3339	3R	0,00159124	-3,37
FBgn0033521	CG12896	2R	2,34E-10	-3,333
FBgn0263219	Dscam4	3L	2,34E-24	-3,013
FBgn0000449	dib	3L	5,12E-10	-2,785
FBgn0038694	CG5217	3R	0,00548457	-2,247
FBgn0058469	lncRNA:CR404 69	X	0,00038774	-2,118
FBgn0034263	CG10934	2R	0,02413419	-2,017
FBgn0027259	Kmn1	X	1,11E-10	-1,994
FBgn0039459	IntS12	3R	8,62E-10	-1,972
FBgn0286204	ich	3R	0,00230839	-1,788
FBgn0039795	Spn100A	3R	0,00023967	-1,761
FBgn0036677	CG13023	3L	0,00069804	-1,717
FBgn0034605	Ugt49C1	2R	0,01367442	-1,59
FBgn0003996	w	X	8,83E-12	-1,533
FBgn0039411	dysf	3R	0,00059985	-1,51

# **BIBLIOGRAPHIE**

1. Kerr, J.F.R.; Wyllie, A.H.; Currie, A.R. Apoptosis: A Basic Biological Phenomenon with Wide-Ranging Implications in Tissue Kinetics. *Br J Cancer* **1972**, *26*, 239–257.
2. Crawford, A.M.; Kerr, J.F.; Currie, A.R. The Relationship of Acute Mesodermal Cell Death to the Teratogenic Effects of 7-OHM-12-MBA in the Foetal Rat. *Br J Cancer* **1972**, *26*, 498–503.
3. Kerr, J.F.R. Shrinkage Necrosis: A Distinct Mode of Cellular Death. *The Journal of Pathology* **1971**, *105*, 13–20, doi:10.1002/path.1711050103.
4. Elmore, S. Apoptosis: A Review of Programmed Cell Death. *Toxicol Pathol* **2007**, *35*, 495–516, doi:10.1080/01926230701320337.
5. Kerr, J.F.R. History of the Events Leading to the Formulation of the Apoptosis Concept. *Toxicology* **2002**, *181–182*, 471–474, doi:10.1016/S0300-483X(02)00457-2.
6. Majno, G.; Joris, I. Apoptosis, Oncosis, and Necrosis. An Overview of Cell Death. *Am J Pathol* **1995**, *146*, 3–15.
7. Glücksmann, A. Cell Deaths in Normal Vertebrate Ontogeny. *Biological Reviews* **1951**, *26*, 59–86, doi:10.1111/j.1469-185X.1951.tb00774.x.
8. Farber, J.L. Biology of Disease: Membrane Injury and Calcium Homeostasis in the Pathogenesis of Coagulative Necrosis. *Lab Invest* **1982**, *47*, 114–123.
9. Sulston, J.E.; Schierenberg, E.; White, J.G.; Thomson, J.N. The Embryonic Cell Lineage of the Nematode *Caenorhabditis Elegans*. *Developmental Biology* **1983**, *100*, 64–119, doi:10.1016/0012-1606(83)90201-4.
10. Sulston, J.E.; Horvitz, H.R. Post-Embryonic Cell Lineages of the Nematode, *Caenorhabditis Elegans*. *Dev Biol* **1977**, *56*, 110–156, doi:10.1016/0012-1606(77)90158-0.
11. Ellis, H.M.; Horvitz, H.R. Genetic Control of Programmed Cell Death in the Nematode *C. Elegans*. *Cell* **1986**, *44*, 817–829, doi:10.1016/0092-8674(86)90004-8.
12. Hengartner, M.O.; Ellis, R.; Horvitz, R. *Caenorhabditis Elegans* Gene *Ced-9* Protects Cells from Programmed Cell Death. *Nature* **1992**, *356*, 494–499, doi:10.1038/356494a0.
13. Hengartner, M.O.; Horvitz, H.R. *C. Elegans* Cell Survival Gene *Ced-9* Encodes a Functional Homolog of the Mammalian Proto-Oncogene *Bcl-2*. *Cell* **1994**, *76*, 665–676, doi:10.1016/0092-8674(94)90506-1.
14. Vaux, D.L.; Weissman, I.L.; Kim, S.K. Prevention of Programmed Cell Death in *Caenorhabditis Elegans* by Human *Bcl-2*. *Science* **1992**, *258*, 1955–1957, doi:10.1126/science.1470921.
15. Arvanitis, M.; Li, D.-D.; Lee, K.; Mylonakis, E. Apoptosis in *C. Elegans*: Lessons for Cancer and Immunity. *Front Cell Infect Microbiol* **2013**, *3*, 67, doi:10.3389/fcimb.2013.00067.
16. Rubin, G.M. *Drosophila Melanogaster* as an Experimental Organism. *Science* **1988**, *240*, 1453–1459, doi:10.1126/science.3131880.
17. Rubin, G.M.; Lewis, E.B. A Brief History of *Drosophila*'s Contributions to Genome Research. *Science* **2000**, *287*, 2216–2218, doi:10.1126/science.287.5461.2216.
18. Adams, M.D.; Celniker, S.E.; Holt, R.A.; Evans, C.A.; Gocayne, J.D.; Amanatides, P.G.; Scherer, S.E.; Li, P.W.; Hoskins, R.A.; Galle, R.F.; et al. The Genome Sequence of *Drosophila Melanogaster*. *Science* **2000**, *287*, 2185–2195, doi:10.1126/science.287.5461.2185.
19. Reiter, L.T.; Potocki, L.; Chien, S.; Gribskov, M.; Bier, E. A Systematic Analysis of Human Disease-Associated Gene Sequences In *Drosophila Melanogaster*. *Genome Res.* **2001**, *11*, 1114–1125, doi:10.1101/gr.169101.
20. Brand, A.H.; Perrimon, N. Targeted Gene Expression as a Means of Altering Cell Fates and Generating Dominant Phenotypes. *Development* **1993**, *118*, 401–415, doi:10.1242/dev.118.2.401.
21. Fleisher, T.A. Apoptosis. *Annals of Allergy, Asthma & Immunology* **1997**, *78*, 245–250, doi:10.1016/S1081-1206(10)63176-6.
22. Zakeri, Z.F.; Ahuja, H.S. Cell Death/Apoptosis: Normal, Chemically Induced, and Teratogenic Effect. *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis* **1997**, *396*, 149–161, doi:10.1016/S0027-5107(97)00181-4.
23. Nijhawan, D.; Honarpour, N.; Wang, X. Apoptosis in Neural Development and Disease. *Annu Rev Neurosci* **2000**, *23*, 73–87, doi:10.1146/annurev.neuro.23.1.73.

## BIBLIOGRAPHIE

24. Tilly, J.L.; Kowalski, K.I.; Johnson, A.L.; Hsueh, A.J.W. INVOLVEMENT OF APOPTOSIS IN OVARIAN FOLLICULAR ATRESIA AND POSTOVULATORY REGRESSION. *Endocrinology* **1991**, *129*, 2799–2801, doi:10.1210/endo-129-5-2799.
25. Opferman, J.T.; Korsmeyer, S.J. Apoptosis in the Development and Maintenance of the Immune System. *Nat Immunol* **2003**, *4*, 410–415, doi:10.1038/ni0503-410.
26. Renehan, A.G.; Booth, C.; Potten, C.S. What Is Apoptosis, and Why Is It Important? *BMJ* **2001**, *322*, 1536–1538.
27. Ethell, D.W.; Buhler, L.A. Fas Ligand-Mediated Apoptosis in Degenerative Disorders of the Brain. *J Clin Immunol* **2003**, *23*, 439–446, doi:10.1023/B:JOCI.0000010420.96419.a8.
28. Li, C.J.; Friedman, D.J.; Wang, C.; Meteleev, V.; Pardee, A.B. Induction of Apoptosis in Uninfected Lymphocytes by HIV-1 Tat Protein. *Science* **1995**, *268*, 429–431, doi:10.1126/science.7716549.
29. Kerr, J.F.R.; Winterford, C.M.; Harmon, B.V. Apoptosis. Its Significance in Cancer and Cancer Therapy. *Cancer* **1994**, *73*, 2013–2026, doi:10.1002/1097-0142(19940415)73:8<2013::AID-CNCR2820730802>3.0.CO;2-J.
30. Watanabe–Fukunaga, R.; Brannan, C.I.; Copeland, N.G.; Jenkins, N.A.; Nagata, S. Pillars Article: Lymphoproliferation Disorder in Mice Explained by Defects in Fas Antigen That Mediates Apoptosis. *Nature*. 1992. 356: 314–317. *The Journal of Immunology* **2012**, *189*, 5101–5104.
31. Dickman, M.; Williams, B.; Li, Y.; de Figueiredo, P.; Wolpert, T. Reassessing Apoptosis in Plants. *Nature Plants* **2017**, *3*, 773–779, doi:10.1038/s41477-017-0020-x.
32. Yuan, J.; Horvitz, H.R. The Caenorhabditis Elegans Genes Ced-3 and Ced-4 Act Cell Autonomously to Cause Programmed Cell Death. *Developmental Biology* **1990**, *138*, 33–41, doi:10.1016/0012-1606(90)90174-H.
33. Alnemri, E.S.; Livingston, D.J.; Nicholson, D.W.; Salvesen, G.; Thornberry, N.A.; Wong, W.W.; Yuan, J. Human ICE/CED-3 Protease Nomenclature. *Cell* **1996**, *87*, 171, doi:10.1016/S0092-8674(00)81334-3.
34. Van Opendenbosch, N.; Lamkanfi, M. Caspases in Cell Death, Inflammation, and Disease. *Immunity* **2019**, *50*, 1352–1364, doi:10.1016/j.immuni.2019.05.020.
35. Sprick, M.R.; Weigand, M.A.; Rieser, E.; Rauch, C.T.; Juo, P.; Blenis, J.; Krammer, P.H.; Walczak, H. FADD/MORT1 and Caspase-8 Are Recruited to TRAIL Receptors 1 and 2 and Are Essential for Apoptosis Mediated by TRAIL Receptor 2. *Immunity* **2000**, *12*, 599–609, doi:10.1016/S1074-7613(00)80211-3.
36. Zou, H.; Li, Y.; Liu, X.; Wang, X. An APAF-1-Cytochrome c Multimeric Complex Is a Functional Apoptosome That Activates Procaspase-9\*. *Journal of Biological Chemistry* **1999**, *274*, 11549–11556, doi:10.1074/jbc.274.17.11549.
37. Rathmell, J.C.; Thompson, C.B. The Central Effectors of Cell Death in the Immune System. *Annu Rev Immunol* **1999**, *17*, 781–828, doi:10.1146/annurev.immunol.17.1.781.
38. Fuentes-Prior, P.; Salvesen, G.S. The Protein Structures That Shape Caspase Activity, Specificity, Activation and Inhibition. *Biochem J* **2004**, *384*, 201–232, doi:10.1042/BJ20041142.
39. Slee, E.A.; Harte, M.T.; Kluck, R.M.; Wolf, B.B.; Casiano, C.A.; Newmeyer, D.D.; Wang, H.-G.; Reed, J.C.; Nicholson, D.W.; Alnemri, E.S.; et al. Ordering the Cytochrome c-Initiated Caspase Cascade: Hierarchical Activation of Caspases-2, -3, -6, -7, -8, and -10 in a Caspase-9-Dependent Manner. *J Cell Biol* **1999**, *144*, 281–292.
40. Wang, X.-J.; Cao, Q.; Liu, X.; Wang, K.-T.; Mi, W.; Zhang, Y.; Li, L.-F.; LeBlanc, A.C.; Su, X.-D. Crystal Structures of Human Caspase 6 Reveal a New Mechanism for Intramolecular Cleavage Self-Activation. *EMBO Rep* **2010**, *11*, 841–847, doi:10.1038/embor.2010.141.
41. Delbridge, A.R.D.; Grabow, S.; Strasser, A.; Vaux, D.L. Thirty Years of BCL-2: Translating Cell Death Discoveries into Novel Cancer Therapies. *Nat Rev Cancer* **2016**, *16*, 99–109, doi:10.1038/nrc.2015.17.
42. Oltval, Z.N.; Milliman, C.L.; Korsmeyer, S.J. Bcl-2 Heterodimerizes in Vivo with a Conserved Homolog, Bax, That Accelerates Programed Cell Death. *Cell* **1993**, *74*, 609–619, doi:10.1016/0092-8674(93)90509-O.

43. Kluck, R.M.; Bossy-Wetzel, E.; Green, D.R.; Newmeyer, D.D. The Release of Cytochrome c from Mitochondria: A Primary Site for Bcl-2 Regulation of Apoptosis. *Science* **1997**, *275*, 1132–1136, doi:10.1126/science.275.5303.1132.
44. Czabotar, P.E.; Westphal, D.; Dewson, G.; Ma, S.; Hockings, C.; Fairlie, W.D.; Lee, E.F.; Yao, S.; Robin, A.Y.; Smith, B.J.; et al. Bax Crystal Structures Reveal How BH3 Domains Activate Bax and Nucleate Its Oligomerization to Induce Apoptosis. *Cell* **2013**, *152*, 519–531, doi:10.1016/j.cell.2012.12.031.
45. Wolf, P.; Schoeniger, A.; Edlich, F. Pro-Apoptotic Complexes of BAX and BAK on the Outer Mitochondrial Membrane. *Biochim Biophys Acta Mol Cell Res* **2022**, *1869*, 119317, doi:10.1016/j.bbamcr.2022.119317.
46. Estaquier, J.; Vallette, F.; Vayssiere, J.-L.; Mignotte, B. The Mitochondrial Pathways of Apoptosis. In *Advances in Mitochondrial Medicine*; Scatena, R., Bottoni, P., Giardina, B., Eds.; Advances in Experimental Medicine and Biology; Springer Netherlands: Dordrecht, 2012; Vol. 942, pp. 157–183 ISBN 978-94-007-2868-4.
47. Ho, T.; Tan, B.X.; Lane, D. How the Other Half Lives: What P53 Does When It Is Not Being a Transcription Factor. *Int J Mol Sci* **2019**, *21*, 13, doi:10.3390/ijms21010013.
48. Czabotar, P.E.; Lessene, G.; Strasser, A.; Adams, J.M. Control of Apoptosis by the BCL-2 Protein Family: Implications for Physiology and Therapy. *Nat Rev Mol Cell Biol* **2014**, *15*, 49–63, doi:10.1038/nrm3722.
49. Oda, E.; Ohki, R.; Murasawa, H.; Nemoto, J.; Shibue, T.; Yamashita, T.; Tokino, T.; Taniguchi, T.; Tanaka, N. Noxa, a BH3-Only Member of the Bcl-2 Family and Candidate Mediator of P53-Induced Apoptosis. *Science* **2000**, *288*, 1053–1058, doi:10.1126/science.288.5468.1053.
50. Nakano, K.; Vousden, K.H. PUMA, a Novel Proapoptotic Gene, Is Induced by P53. *Molecular Cell* **2001**, *7*, 683–694, doi:10.1016/S1097-2765(01)00214-3.
51. Letai, A.; Bassik, M.C.; Walensky, L.D.; Sorcinelli, M.D.; Weiler, S.; Korsmeyer, S.J. Distinct BH3 Domains Either Sensitize or Activate Mitochondrial Apoptosis, Serving as Prototype Cancer Therapeutics. *Cancer Cell* **2002**, *2*, 183–192, doi:10.1016/S1535-6108(02)00127-7.
52. Deveraux, Q.L.; Roy, N.; Stennicke, H.R.; Van Arsdale, T.; Zhou, Q.; Srinivasula, S.M.; Alnemri, E.S.; Salvesen, G.S.; Reed, J.C. IAPs Block Apoptotic Events Induced by Caspase-8 and Cytochrome c by Direct Inhibition of Distinct Caspases. *The EMBO Journal* **1998**, *17*, 2215–2223, doi:10.1093/emboj/17.8.2215.
53. Deveraux, Q.L.; Reed, J.C. IAP Family Proteins—Suppressors of Apoptosis. *Genes Dev.* **1999**, *13*, 239–252.
54. Duckett, C.S.; Li, F.; Wang, Y.; Tomaselli, K.J.; Thompson, C.B.; Armstrong, R.C. Human IAP-Like Protein Regulates Programmed Cell Death Downstream of Bcl-XL and Cytochrome c. *Mol Cell Biol* **1998**, *18*, 608–615.
55. Du, C.; Fang, M.; Li, Y.; Li, L.; Wang, X. Smac, a Mitochondrial Protein That Promotes Cytochrome c-Dependent Caspase Activation by Eliminating IAP Inhibition. *Cell* **2000**, *102*, 33–42, doi:10.1016/S0092-8674(00)00008-8.
56. Verhagen, A.M.; Ekert, P.G.; Pakusch, M.; Silke, J.; Connolly, L.M.; Reid, G.E.; Moritz, R.L.; Simpson, R.J.; Vaux, D.L. Identification of DIABLO, a Mammalian Protein That Promotes Apoptosis by Binding to and Antagonizing IAP Proteins. *Cell* **2000**, *102*, 43–53, doi:10.1016/S0092-8674(00)00009-X.
57. van Loo, G.; van Gurp, M.; Depuydt, B.; Srinivasula, S.M.; Rodriguez, I.; Alnemri, E.S.; Gevaert, K.; Vandekerckhove, J.; Declercq, W.; Vandenabeele, P. The Serine Protease Omi/HtrA2 Is Released from Mitochondria during Apoptosis. Omi Interacts with Caspase-Inhibitor XIAP and Induces Enhanced Caspase Activity. *Cell Death Differ* **2002**, *9*, 20–26, doi:10.1038/sj.cdd.4400970.
58. Igney, F.H.; Krammer, P.H. Death and Anti-Death: Tumour Resistance to Apoptosis. *Nat Rev Cancer* **2002**, *2*, 277–288, doi:10.1038/nrc776.
59. White, K.; Grether, M.E.; Abrams, J.M.; Young, L.; Farrell, K.; Steller, H. Genetic Control of Programmed Cell Death in *Drosophila*. *Science* **1994**, *264*, 677–683, doi:10.1126/science.8171319.

## BIBLIOGRAPHIE

60. Hay, B.A.; Wassarman, D.A.; Rubin, G.M. Drosophila Homologs of Baculovirus Inhibitor of Apoptosis Proteins Function to Block Cell Death. *Cell* **1995**, *83*, 1253–1262, doi:10.1016/0092-8674(95)90150-7.
61. Hay, B.A.; Wolff, T.; Rubin, G.M. Expression of Baculovirus P35 Prevents Cell Death in Drosophila. *Development* **1994**, *120*, 2121–2129, doi:10.1242/dev.120.8.2121.
62. Clavier, A.; Rincheval-Arnold, A.; Mignotte, B.; Guénel, I. Le Come-Back de La Mitochondrie Dans l'apoptose Chez La Drosophile. *médecine/sciences* **2016**, *32*, 478–484, doi:10.1051/medsci/20163205014.
63. Kornbluth, S.; White, K. Apoptosis in Drosophila: Neither Fish nor Fowl (nor Man, nor Worm). *Journal of Cell Science* **2005**, *118*, 1779–1787, doi:10.1242/jcs.02377.
64. Kanuka, H.; Sawamoto, K.; Inohara, N.; Matsuno, K.; Okano, H.; Miura, M. Control of the Cell Death Pathway by Dapaf-1, a Drosophila Apaf-1/CED-4-Related Caspase Activator. *Molecular Cell* **1999**, *4*, 757–769, doi:10.1016/S1097-2765(00)80386-X.
65. Kumar, S.; Dumanis, J. The Fly Caspases. *Cell Death Differ* **2000**, *7*, 1039–1044, doi:10.1038/sj.cdd.4400756.
66. Wilson, R.; Goyal, L.; Ditzel, M.; Zachariou, A.; Baker, D.A.; Agapite, J.; Steller, H.; Meier, P. The DIAP1 RING Finger Mediates Ubiquitination of Dronc and Is Indispensable for Regulating Apoptosis. *Nat Cell Biol* **2002**, *4*, 445–450, doi:10.1038/ncb799.
67. Colussi, P.A.; Quinn, L.M.; Huang, D.C.S.; Coombe, M.; Read, S.H.; Richardson, H.; Kumar, S. Debcl, a Proapoptotic Bcl-2 Homologue, Is a Component of the Drosophila Melanogaster Cell Death Machinery. *Journal of Cell Biology* **2000**, *148*, 703–714, doi:10.1083/jcb.148.4.703.
68. Quinn, L.; Coombe, M.; Mills, K.; Daish, T.; Colussi, P.; Kumar, S.; Richardson, H. Buffy, a Drosophila Bcl-2 Protein, Has Anti-Apoptotic and Cell Cycle Inhibitory Functions. *The EMBO Journal* **2003**, *22*, 3568–3579, doi:10.1093/emboj/cdg355.
69. Senoo-Matsuda, N.; Igaki, T.; Miura, M. Bax-like Protein Drob-1 Protects Neurons from Expanded Polyglutamine-Induced Toxicity in Drosophila. *EMBO J.* **2005**, *24*, 2700–2713, doi:10.1038/sj.emboj.7600721.
70. Zamzami, N.; Kroemer, G. The Mitochondrion in Apoptosis: How Pandora's Box Opens. *Nat Rev Mol Cell Biol* **2001**, *2*, 67–71, doi:10.1038/35048073.
71. Igaki, T.; Kanuka, H.; Inohara, N.; Sawamoto, K.; Núñez, G.; Okano, H.; Miura, M. Drob-1, a Drosophila Member of the Bcl-2/CED-9 Family That Promotes Cell Death. *Proceedings of the National Academy of Sciences* **2000**, *97*, 662–667, doi:10.1073/pnas.97.2.662.
72. Clavier, A. Caractérisation des événements moléculaires et cellulaires de l'apoptose induite par rbf1, l'homologue de drosophile du gène suppresseur de tumeur rb.
73. Varkey, J.; Chen, P.; Jemmerson, R.; Abrams, J.M. Altered Cytochrome c Display Precedes Apoptotic Cell Death in Drosophila. *Journal of Cell Biology* **1999**, *144*, 701–710, doi:10.1083/jcb.144.4.701.
74. Hawkins, C.J.; Yoo, S.J.; Peterson, E.P.; Wang, S.L.; Vernooy, S.Y.; Hay, B.A. The Drosophila Caspase DRONC Cleaves Following Glutamate or Aspartate and Is Regulated by DIAP1, HID, and GRIM\*. *Journal of Biological Chemistry* **2000**, *275*, 27084–27093, doi:10.1016/S0021-9258(19)61483-3.
75. Fraser, A.G.; McCarthy, N.J.; Evan, G.I. DrICE Is an Essential Caspase Required for Apoptotic Activity in Drosophila Cells. *The EMBO Journal* **1997**, *16*, 6192–6199, doi:10.1093/emboj/16.20.6192.
76. Olson, M.R.; Holley, C.L.; Gan, E.C.; Colón-Ramos, D.A.; Kaplan, B.; Kornbluth, S. A GH3-like Domain in Reaper Is Required for Mitochondrial Localization and Induction of IAP Degradation\*. *Journal of Biological Chemistry* **2003**, *278*, 44758–44768, doi:10.1074/jbc.M308055200.
77. Osellame, L.D.; Singh, A.P.; Stroud, D.A.; Palmer, C.S.; Stojanovski, D.; Ramachandran, R.; Ryan, M.T. Cooperative and Independent Roles of the Drp1 Adaptors Mff, MiD49 and MiD51 in Mitochondrial Fission. *Journal of Cell Science* **2016**, *129*, 2170–2181, doi:10.1242/jcs.185165.
78. Thomenius, M.; Freel, C.D.; Horn, S.; Krieser, R.; Abdelwahid, E.; Cannon, R.; Balasundaram, S.; White, K.; Kornbluth, S. Mitochondrial Fusion Is Regulated by Reaper to Modulate



- Drosophila Programmed Cell Death. *Cell Death Differ* **2011**, *18*, 1640–1650, doi:10.1038/cdd.2011.26.
79. Tout savoir sur le rétinoblastome - Hôpital ophtalmique Jules-Gonin Available online: <https://www.ophtalmique.ch/hopital/maladies-de-la-vue/reтинoblastome/tout-savoir-sur-le-reтинoblastome/> (accessed on 10 October 2022).
  80. Albert, D.M. Historic Review of Retinoblastoma. *Ophthalmology* **1987**, *94*, 654–662, doi:10.1016/S0161-6420(87)33407-4.
  81. Ortiz, M.V.; Dunkel, I.J. Retinoblastoma. *J Child Neurol* **2016**, *31*, 227–236, doi:10.1177/0883073815587943.
  82. Cancers de l'enfant : le rétinoblastome, une tumeur de la rétine | Institut Curie Available online: <https://curie.fr/dossier-pedagogique/cancers-de-lenfant-le-reтинoblastome-une-tumeur-de-la-reтine> (accessed on 10 October 2022).
  83. Rao, R.; Honavar, S.G. Retinoblastoma. *Indian J Pediatr* **2017**, *84*, 937–944, doi:10.1007/s12098-017-2395-0.
  84. Dunkel, I.J.; Khakoo, Y.; Kernan, N.A.; Gershon, T.; Gilheeny, S.; Lyden, D.C.; Wolden, S.L.; Orjuela, M.; Gardner, S.L.; Abramson, D.H. Intensive Multimodality Therapy for Patients with Stage 4a Metastatic Retinoblastoma. *Pediatric Blood & Cancer* **2010**, *55*, 55–59, doi:10.1002/pbc.22504.
  85. Knudson, A.G. Mutation and Cancer: Statistical Study of Retinoblastoma. *Proceedings of the National Academy of Sciences* **1971**, *68*, 820–823, doi:10.1073/pnas.68.4.820.
  86. Francke, U.; Kung, F. Sporadic Bilateral Retinoblastoma and 13q- Chromosomal Deletion. *Med Pediatr Oncol* **1976**, *2*, 379–385, doi:10.1002/mpo.2950020404.
  87. Knudson, A.G.; Meadows, A.T.; Nichols, W.W.; Hill, R. Chromosomal Deletion and Retinoblastoma. *N Engl J Med* **1976**, *295*, 1120–1123, doi:10.1056/NEJM197611112952007.
  88. Lee, W.-H.; Bookstein, R.; Hong, F.; Young, L.-J.; Shew, J.-Y.; Lee, E.Y.-H.P. Human Retinoblastoma Susceptibility Gene: Cloning, Identification, and Sequence. *Science* **1987**, doi:10.1126/science.3823889.
  89. Huang, H.-J.S.; Yee, J.-K.; Shew, J.-Y.; Chen, P.-L.; Bookstein, R.; Friedmann, T.; Lee, E.Y.-H.P.; Lee, W.-H. Suppression of the Neoplastic Phenotype by Replacement of the RB Gene in Human Cancer Cells. *Science* **1988**, *242*, 1563–1566, doi:10.1126/science.3201247.
  90. Classon, M.; Harlow, E. The Retinoblastoma Tumour Suppressor in Development and Cancer. *Nature Reviews Cancer* **2002**, *2*, 910, doi:10.1038/nrc950.
  91. Burkhardt, D.L.; Sage, J. Cellular Mechanisms of Tumour Suppression by the Retinoblastoma Gene. *Nat Rev Cancer* **2008**, *8*, 671–682, doi:10.1038/nrc2399.
  92. Chatterjee, S.J.; Datar, R.; Youssefzadeh, D.; George, B.; Goebell, P.J.; Stein, J.P.; Young, L.; Shi, S.-R.; Gee, C.; Groshen, S.; et al. Combined Effects of P53, P21, and PRb Expression in the Progression of Bladder Transitional Cell Carcinoma. *Journal of Clinical Oncology* **2004**, *22*, 1007–1013, doi:10.1200/JCO.2004.05.174.
  93. Ittmann, M.M.; Wiczorek, R. Alterations of the Retinoblastoma Gene in Clinically Localized, Stage B Prostate Adenocarcinomas. *Hum. Pathol.* **1996**, *27*, 28–34.
  94. Lipinski, M.M.; Jacks, T. The Retinoblastoma Gene Family in Differentiation and Development. *Oncogene* **1999**, *18*, 7873–7882, doi:10.1038/sj.onc.1203244.
  95. Du, W.; Pogoriler, J. Retinoblastoma Family Genes. *Oncogene* **2006**, *25*, 5190–5200, doi:10.1038/sj.onc.1209651.
  96. Kato, J.; Matsushime, H.; Hiebert, S.W.; Ewen, M.E.; Sherr, C.J. Direct Binding of Cyclin D to the Retinoblastoma Gene Product (PRb) and PRb Phosphorylation by the Cyclin D-Dependent Kinase CDK4. *13*.
  97. Du, W.; Vidal, M.; Xie, J.E.; Dyson, N. RBF, a Novel RB-Related Gene That Regulates E2F Activity and Interacts with Cyclin E in Drosophila. *Genes Dev.* **1996**, *10*, 1206–1218, doi:10.1101/gad.10.10.1206.
  98. Sanidas, I.; Morris, R.; Fella, K.A.; Rumde, P.H.; Boukhali, M.; Tai, E.C.; Ting, D.T.; Lawrence, M.S.; Haas, W.; Dyson, N.J. A Code of Mono-Phosphorylation Modulates the Function of RB. *Molecular Cell* **2019**, *73*, 985–1000.e6, doi:10.1016/j.molcel.2019.01.004.
  99. Jacks, T.; Fazeli, A.; Schmitt, E.M.; Bronson, R.T.; Goodell, M.A.; Weinberg, R.A. Effects of an Rb Mutation in the Mouse. *Nature* **1992**, *359*, 295–300, doi:10.1038/359295a0.

## BIBLIOGRAPHIE

100. Keller, S.A.; Ullah, Z.; Buckley, M.S.; Henry, R.W.; Arnosti, D.N. Distinct Developmental Expression of Drosophila Retinoblastoma Factors. *Gene Expr. Patterns* **2005**, *5*, 411–421, doi:10.1016/j.modgep.2004.09.005.
101. Cobrinik, D.; Lee, M.H.; Hannon, G.; Mulligan, G.; Bronson, R.T.; Dyson, N.; Harlow, E.; Beach, D.; Weinberg, R.A.; Jacks, T. Shared Role of the PRB-Related P130 and P107 Proteins in Limb Development. *Genes Dev* **1996**, *10*, 1633–1644, doi:10.1101/gad.10.13.1633.
102. Morris, E.J.; Dyson, N.J. Retinoblastoma Protein Partners. In *Advances in Cancer Research*; Academic Press, 2001; Vol. 82, pp. 1–54.
103. Frolov, M.V.; Dyson, N.J. Molecular Mechanisms of E2F-Dependent Activation and PRB-Mediated Repression. *J Cell Sci* **2004**, *117*, 2173–2181, doi:10.1242/jcs.01227.
104. Cao, L.; Peng, B.; Yao, L.; Zhang, X.; Sun, K.; Yang, X.; Yu, L. The Ancient Function of RB-E2F Pathway: Insights from Its Evolutionary History. *Biology Direct* **2010**, *5*, 55, doi:10.1186/1745-6150-5-55.
105. Kirienko, N.V.; Fay, D.S. Transcriptome Profiling of the C. Elegans Rb Ortholog Reveals Diverse Developmental Roles. *Developmental Biology* **2007**, *305*, 674–684, doi:10.1016/j.ydbio.2007.02.021.
106. Desvoyes, B.; Gutierrez, C. Roles of Plant Retinoblastoma Protein: Cell Cycle and Beyond. *The EMBO Journal* **2020**, *39*, e105802, doi:10.15252/embj.2020105802.
107. Frolov, M.V.; Huen, D.S.; Stevaux, O.; Dimova, D.; Balczarek-Strang, K.; Elsdon, M.; Dyson, N.J. Functional Antagonism between E2F Family Members. *Genes Dev.* **2001**, *15*, 2146–2160, doi:10.1101/gad.903901.
108. Dynlacht, B.D.; Brook, A.; Dembski, M.; Yenush, L.; Dyson, N. DNA-Binding and Trans-Activation Properties of Drosophila E2F and DP Proteins. *Proceedings of the National Academy of Sciences* **1994**, *91*, 6359–6363, doi:10.1073/pnas.91.14.6359.
109. Stevaux, O.; Dimova, D.; Frolov, M.V.; Taylor-Harding, B.; Morris, E.; Dyson, N. Distinct Mechanisms of E2F Regulation by Drosophila RBF1 and RBF2. *The EMBO journal* **2002**, *21*, 4927–4937.
110. Chicas, A.; Wang, X.; Zhang, C.; McCurrach, M.; Zhao, Z.; Mert, O.; Dickins, R.A.; Narita, M.; Zhang, M.; Lowe, S.W. Dissecting the Unique Role of the Retinoblastoma Tumor Suppressor during Cellular Senescence. *Cancer Cell* **2010**, *17*, 376–387, doi:10.1016/j.ccr.2010.01.023.
111. Sanidas, I.; Lee, H.; Rumde, P.H.; Boulay, G.; Morris, R.; Golczer, G.; Stanzione, M.; Hajizadeh, S.; Zhong, J.; Ryan, M.B.; et al. Chromatin-Bound RB Targets Promoters, Enhancers, and CTCF-Bound Loci and Is Redistributed by Cell-Cycle Progression. *Molecular Cell* **2022**, *82*, 3333–3349.e9, doi:10.1016/j.molcel.2022.07.014.
112. Korenjak, M.; Anderssen, E.; Ramaswamy, S.; Whetstone, J.R.; Dyson, N.J. RBF Binding to Both Canonical E2F Targets and Noncanonical Targets Depends on Functional DE2F/DDP Complexes. *Mol Cell Biol* **2012**, *32*, 4375–4387, doi:10.1128/MCB.00536-12.
113. Acharya, P.; Negre, N.; Johnston, J.; Wei, Y.; White, K.P.; Henry, R.W.; Arnosti, D.N. Evidence for Autoregulation and Cell Signaling Pathway Regulation from Genome-Wide Binding of the Drosophila Retinoblastoma Protein. *G3 (Bethesda)* **2012**, *2*, 1459–1472, doi:10.1534/g3.112.004424.
114. Dimova, D.K.; Stevaux, O.; Frolov, M.V.; Dyson, N.J. Cell Cycle-Dependent and Cell Cycle-Independent Control of Transcription by the Drosophila E2F/RB Pathway. *Genes Dev* **2003**, *17*, 2308–2320, doi:10.1101/gad.1116703.
115. Korenjak, M.; Kwon, E.; Morris, R.T.; Anderssen, E.; Amzallag, A.; Ramaswamy, S.; Dyson, N.J. DREAM Co-Operates with Insulator-Binding Proteins and Regulates Expression at Divergently Paired Genes. *Nucleic Acids Res.* **2014**, *42*, 8939–8953, doi:10.1093/nar/gku609.
116. Mäkelä, J.-A.; Toppari, J. Retinoblastoma-E2F Transcription Factor Interplay Is Essential for Testicular Development and Male Fertility. *Frontiers in Endocrinology* **2022**, *13*.
117. Payankulam, S.; Hickey, S.L.; Arnosti, D.N. Cell Cycle Expression of Polarity Genes Features Rb Targeting of Vang. *Cells & Development* **2022**, *169*, 203747, doi:10.1016/j.cdev.2021.203747.

118. Rustgi, A.K.; Dyson, N.; Bernards, R. Amino-Terminal Domains of c-Myc and N-Myc Proteins Mediate Binding to the Retinoblastoma Gene Product. *Nature* **1991**, *352*, 541–544, doi:10.1038/352541a0.
119. Chen, P.L.; Riley, D.J.; Chen, Y.; Lee, W.H. Retinoblastoma Protein Positively Regulates Terminal Adipocyte Differentiation through Direct Interaction with C/EBPs. *Genes Dev* **1996**, *10*, 2794–2804, doi:10.1101/gad.10.21.2794.
120. Helin, K.; Harlow, E.; Fattaey, A. Inhibition of E2F-1 Transactivation by Direct Binding of the Retinoblastoma Protein. *Molecular and Cellular Biology* **1993**, *13*, 6501–6508, doi:10.1128/mcb.13.10.6501-6508.1993.
121. Tevosian, S.G.; Shih, H.H.; Mendelson, K.G.; Sheppard, K.A.; Paulson, K.E.; Yee, A.S. HBP1: A HMG Box Transcriptional Repressor That Is Targeted by the Retinoblastoma Family. *Genes Dev* **1997**, *11*, 383–396, doi:10.1101/gad.11.3.383.
122. Dynlacht, B.D. Regulation of Transcription by Proteins That Control the Cell Cycle. *Nature* **1997**, *389*, 149–152, doi:10.1038/38225.
123. Ross, J.F.; Liu, X.; Dynlacht, B.D. Mechanism of Transcriptional Repression of E2F by the Retinoblastoma Tumor Suppressor Protein. *Molecular Cell* **1999**, *3*, 195–205, doi:10.1016/S1097-2765(00)80310-X.
124. Wei, Y.; Mondal, S.S.; Mouawad, R.; Wilczyński, B.; Henry, R.W.; Arnosti, D.N. Genome-Wide Analysis of Drosophila RBF2 Protein Highlights the Diversity of RB Family Targets and Possible Role in Regulation of Ribosome Biosynthesis. *G3 (Bethesda)* **2015**, *5*, 1503–1515, doi:10.1534/g3.115.019166.
125. Longworth, M.S.; Herr, A.; Ji, J.-Y.; Dyson, N.J. RBF1 Promotes Chromatin Condensation through a Conserved Interaction with the Condensin II Protein DCAP-D3. *Genes Dev.* **2008**, *22*, 1011–1024, doi:10.1101/gad.1631508.
126. Ishak, C.A.; Coschi, C.H.; Roes, M.V.; Dick, F.A. Disruption of CDK-Resistant Chromatin Association by PRB Causes DNA Damage, Mitotic Errors, and Reduces Condensin II Recruitment. *Cell Cycle* **2017**, *16*, 1430–1439, doi:10.1080/15384101.2017.1338984.
127. Dick, F.A.; Goodrich, D.W.; Sage, J.; Dyson, N.J. Non-Canonical Functions of the RB Protein in Cancer. *Nat Rev Cancer* **2018**, *18*, 442–451, doi:10.1038/s41568-018-0008-5.
128. Narlikar, G.J.; Fan, H.-Y.; Kingston, R.E. Cooperation between Complexes That Regulate Chromatin Structure and Transcription. *Cell* **2002**, *108*, 475–487, doi:10.1016/S0092-8674(02)00654-2.
129. Dunaief, J.L.; Strober, B.E.; Guha, S.; Khavari, P.A.; Alin, K.; Luban, J.; Begemann, M.; Crabtree, G.R.; Goff, S.P. The Retinoblastoma Protein and BRG1 Form a Complex and Cooperate to Induce Cell Cycle Arrest. *Cell* **1994**, *79*, 119–130, doi:10.1016/0092-8674(94)90405-7.
130. Lacoste, N.; Côté, J. Le code épigénétique des histones. *Med Sci (Paris)* **2003**, *19*, 955–959, doi:10.1051/medsci/20031910955.
131. Martínez-Balbás, M.A.; Bauer, U.-M.; Nielsen, S.J.; Brehm, A.; Kouzarides, T. Regulation of E2F1 Activity by Acetylation. *The EMBO Journal* **2000**, *19*, 662–671, doi:10.1093/emboj/19.4.662.
132. Lee, H.; Ohno, K.; Voskoboynik, Y.; Ragusano, L.; Martinez, A.; Dimova, D.K. Drosophila RB Proteins Repress Differentiation-Specific Genes via Two Different Mechanisms. *Molecular and Cellular Biology* **2010**, *30*, 2563–2577, doi:10.1128/MCB.01075-09.
133. Vandel, L.; Nicolas, E.; Vaute, O.; Ferreira, R.; Ait-Si-Ali, S.; Trouche, D. Transcriptional Repression by the Retinoblastoma Protein through the Recruitment of a Histone Methyltransferase. *Mol Cell Biol* **2001**, *21*, 6484–6494, doi:10.1128/MCB.21.19.6484-6494.2001.
134. Nielsen, S.J.; Schneider, R.; Bauer, U.M.; Bannister, A.J.; Morrison, A.; O’Carroll, D.; Firestein, R.; Cleary, M.; Jenuwein, T.; Herrera, R.E.; et al. Rb Targets Histone H3 Methylation and HP1 to Promoters. *Nature* **2001**, *412*, 561–565, doi:10.1038/35087620.
135. Couteau, F.; Guerry, F.; Müller, F.; Palladino, F. A Heterochromatin Protein 1 Homologue in Caenorhabditis Elegans Acts in Germline and Vulval Development. *EMBO reports* **2002**, *3*, 235–241, doi:10.1093/embo-reports/kvf051.

## BIBLIOGRAPHIE

136. Ishak, C.A.; Marshall, A.E.; Passos, D.T.; White, C.R.; Kim, S.J.; Cecchini, M.J.; Ferwati, S.; MacDonald, W.A.; Howlett, C.J.; Welch, I.D.; et al. An RB-EZH2 Complex Mediates Silencing of Repetitive DNA Sequences. *Mol Cell* **2016**, *64*, 1074–1087, doi:10.1016/j.molcel.2016.10.021.
137. Robertson, K.D.; Ait-Si-Ali, S.; Yokochi, T.; Wade, P.A.; Jones, P.L.; Wolffe, A.P. DNMT1 Forms a Complex with Rb, E2F1 and HDAC1 and Represses Transcription from E2F-Responsive Promoters. *Nat Genet* **2000**, *25*, 338–342, doi:10.1038/77124.
138. Isaac, C.E.; Francis, S.M.; Martens, A.L.; Julian, L.M.; Seifried, L.A.; Erdmann, N.; Binné, U.K.; Harrington, L.; Sicinski, P.; Bérubé, N.G.; et al. The Retinoblastoma Protein Regulates Pericentric Heterochromatin. *Mol Cell Biol* **2006**, *26*, 3659–3671, doi:10.1128/MCB.26.9.3659-3671.2006.
139. Montoya-Durango, D.E.; Ramos, K.A.; Bojang, P.; Ruiz, L.; Ramos, I.N.; Ramos, K.S. LINE-1 Silencing by Retinoblastoma Proteins Is Effected through the Nucleosomal and Remodeling Deacetylase Multiprotein Complex. *BMC Cancer* **2016**, *16*, 38, doi:10.1186/s12885-016-2068-9.
140. Narasimha, A.M.; Kaulich, M.; Shapiro, G.S.; Choi, Y.J.; Sicinski, P.; Dowdy, S.F. Cyclin D Activates the Rb Tumor Suppressor by Mono-Phosphorylation. *eLife* **2014**, *3*, e02872, doi:10.7554/eLife.02872.
141. Buchkovich, K.; Duffy, L.A.; Harlow, E. The Retinoblastoma Protein Is Phosphorylated during Specific Phases of the Cell Cycle. *Cell* **1989**, *58*, 1097–1105, doi:10.1016/0092-8674(89)90508-4.
142. Sherr, C.J.; McCormick, F. The RB and P53 Pathways in Cancer. *Cancer Cell* **2002**, *2*, 103–112, doi:10.1016/S1535-6108(02)00102-2.
143. Shirodkar, S.; Ewen, M.; DeCaprio, J.A.; Morgan, J.; Livingston, D.M.; Chittenden, T. The Transcription Factor E2F Interacts with the Retinoblastoma Product and a P107-Cyclin A Complex in a Cell Cycle-Regulated Manner. *Cell* **1992**, *68*, 157–166, doi:10.1016/0092-8674(92)90214-W.
144. Bosco, G.; Du, W.; Orr-Weaver, T.L. DNA Replication Control through Interaction of E2F–RB and the Origin Recognition Complex. *Nat Cell Biol* **2001**, *3*, 289–295, doi:10.1038/35060086.
145. Weng, L.; Zhu, C.; Xu, J.; Du, W. Critical Role of Active Repression by E2F and Rb Proteins in Endoreplication during Drosophila Development. *The EMBO Journal* **2003**, *22*, 3865–3875, doi:10.1093/emboj/cdg373.
146. Gordon, G.M.; Du, W. Conserved RB Functions in Development and Tumor Suppression. *Protein Cell* **2011**, *2*, 864–878, doi:10.1007/s13238-011-1117-z.
147. Ianari, A.; Natale, T.; Calo, E.; Ferretti, E.; Alesse, E.; Screpanti, I.; Haigis, K.; Gulino, A.; Lees, J.A. Proapoptotic Function of the Retinoblastoma Tumor Suppressor Protein. *Cancer Cell* **2009**, *15*, 184–194, doi:10.1016/j.ccr.2009.01.026.
148. Tripathi, B.K.; Irvine, K.D. The Wing Imaginal Disc. *Genetics* **2022**, *220*, iyac020, doi:10.1093/genetics/iyac020.
149. O’Brochta, D.A.; Bryant, P.J. A Zone of Non-Proliferating Cells at a Lineage Restriction Boundary in Drosophila. *Nature* **1985**, *313*, 138–141, doi:10.1038/313138a0.
150. Ready, D.F.; Hanson, T.E.; Benzer, S. Development of the Drosophila Retina, a Neurocrystalline Lattice. *Dev Biol* **1976**, *53*, 217–240, doi:10.1016/0012-1606(76)90225-6.
151. Ariss, M.M.; Islam, A.B.M.M.K.; Critcher, M.; Zappia, M.P.; Frolov, M.V. Single Cell RNA-Sequencing Identifies a Metabolic Aspect of Apoptosis in Rbf Mutant. *Nat Commun* **2018**, *9*, 5024, doi:10.1038/s41467-018-07540-z.
152. Roignant, J.-Y.; Treisman, J.E. Pattern Formation in the Drosophila Eye Disc. *Int J Dev Biol* **2009**, *53*, 795–804, doi:10.1387/ijdb.072483jr.
153. Slack, R.S.; Miller, F.D. Retinoblastoma gene in mouse neural development. *Developmental Genetics* **1996**, *18*, 81–91, doi:10.1002/(SICI)1520-6408(1996)18:1<81::AID-DVG9>3.0.CO;2-Y.
154. de Bruin, A.; Wu, L.; Saavedra, H.I.; Wilson, P.; Yang, Y.; Rosol, T.J.; Weinstein, M.; Robinson, M.L.; Leone, G. Rb Function in Extraembryonic Lineages Suppresses Apoptosis in the CNS of Rb-Deficient Mice. *Proc Natl Acad Sci U S A* **2003**, *100*, 6546–6551, doi:10.1073/pnas.1031853100.

155. Almasan, A.; Yin, Y.; Kelly, R.E.; Lee, E.Y.; Bradley, A.; Li, W.; Bertino, J.R.; Wahl, G.M. Deficiency of Retinoblastoma Protein Leads to Inappropriate S-Phase Entry, Activation of E2F-Responsive Genes, and Apoptosis. *Proc Natl Acad Sci U S A* **1995**, *92*, 5436–5440, doi:10.1073/pnas.92.12.5436.
156. Bosco, E.E.; Mayhew, C.N.; Hennigan, R.F.; Sage, J.; Jacks, T.; Knudsen, E.S. RB Signaling Prevents Replication-Dependent DNA Double-Strand Breaks Following Genotoxic Insult. *Nucleic Acids Res* **2004**, *32*, 25–34, doi:10.1093/nar/gkg919.
157. Knudsen, K.E.; Weber, E.; Arden, K.C.; Cavenee, W.K.; Feramisco, J.R.; Knudsen, E.S. The Retinoblastoma Tumor Suppressor Inhibits Cellular Proliferation through Two Distinct Mechanisms: Inhibition of Cell Cycle Progression and Induction of Cell Death. *Oncogene* **1999**, *18*, 5239–5245, doi:10.1038/sj.onc.1202910.
158. Milet, C.; Rincheval-Arnold, A.; Mignotte, B.; GuénaI, I. The Drosophila Retinoblastoma Protein Induces Apoptosis in Proliferating but Not in Post-Mitotic Cells. *Cell Cycle* **2010**, *9*, 97–103, doi:10.4161/cc.9.1.10251.
159. Clavier, A.; Baillet, A.; Rincheval-Arnold, A.; Coléno-Costes, A.; Lasbleiz, C.; Mignotte, B.; GuénaI, I. The Pro-Apoptotic Activity of Drosophila Rbf1 Involves DE2F2-Dependent Downregulation of Diap1 and Buffy mRNA. *Cell Death Dis* **2014**, *5*, e1405, doi:10.1038/cddis.2014.372.
160. Sevrioukov, E.A.; Burr, J.; Huang, E.W.; Assi, H.H.; Monserrate, J.P.; Purves, D.C.; Wu, J.N.; Song, E.J.; Brachmann, C.B. Drosophila Bcl-2 Proteins Participate in Stress-Induced Apoptosis, but Are Not Required for Normal Development. *Genesis* **2007**, *45*, 184–193, doi:10.1002/dvg.20279.
161. Clavier, A.; Ruby, V.; Rincheval-Arnold, A.; Mignotte, B.; GuénaI, I. The Drosophila Retinoblastoma Protein, Rbf1, Induces a Debcl- and Drp1-Dependent Mitochondrial Apoptosis. *Journal of Cell Science* **2015**, *128*, 3239–3249, doi:10.1242/jcs.169896.
162. Milet, C.; Rincheval-Arnold, A.; Moriéras, A.; Clavier, A.; Garrigue, A.; Mignotte, B.; GuénaI, I. Mutating RBF Can Enhance Its Pro-Apoptotic Activity and Uncovers a New Role in Tissue Homeostasis. *PLoS One* **2014**, *9*, e102902, doi:10.1371/journal.pone.0102902.
163. Clavier, A.; Rincheval-Arnold, A.; Baillet, A.; Mignotte, B.; GuénaI, I. Two Different Specific JNK Activators Are Required to Trigger Apoptosis or Compensatory Proliferation in Response to Rbf1 in Drosophila. *Cell Cycle* **2016**, *15*, 283–294, doi:10.1080/15384101.2015.1100776.
164. Rolland, S.G.; Conradt, B. New Role of the BCL2 Family of Proteins in the Regulation of Mitochondrial Dynamics. *Curr Opin Cell Biol* **2010**, *22*, 852–858, doi:10.1016/j.ceb.2010.07.014.
165. Karbowski, M.; Lee, Y.-J.; Gaume, B.; Jeong, S.-Y.; Frank, S.; Nechushtan, A.; Santel, A.; Fuller, M.; Smith, C.L.; Youle, R.J. Spatial and Temporal Association of Bax with Mitochondrial Fission Sites, Drp1, and Mfn2 during Apoptosis. *J Cell Biol* **2002**, *159*, 931–938, doi:10.1083/jcb.200209124.
166. Hilgendorf, K.I.; Leshchiner, E.S.; Nedelcu, S.; Maynard, M.A.; Calo, E.; Ianari, A.; Walensky, L.D.; Lees, J.A. The Retinoblastoma Protein Induces Apoptosis Directly at the Mitochondria. *Genes Dev* **2013**, *27*, 1003–1015, doi:10.1101/gad.211326.112.
167. Ferecatu, I.; Le Floch, N.; Bergeaud, M.; Rodríguez-Enfedaque, A.; Rincheval, V.; Oliver, L.; Vallette, F.M.; Mignotte, B.; Vayssière, J.-L. Evidence for a Mitochondrial Localization of the Retinoblastoma Protein. *BMC Cell Biol* **2009**, *10*, 50, doi:10.1186/1471-2121-10-50.
168. Khidr, L.; Chen, P.-L. RB, the Conductor That Orchestrates Life, Death and Differentiation. *Oncogene* **2006**, *25*, 5210, doi:10.1038/sj.onc.1209612.
169. Norton, J.D.; Deed, R.W.; Craggs, G.; Sablitzky, F. Id Helix—Loop—Helix Proteins in Cell Growth and Differentiation. *Trends in Cell Biology* **1998**, *8*, 58–65, doi:10.1016/S0962-8924(98)80013-5.
170. Chen, D.; Opavsky, R.; Pacal, M.; Tanimoto, N.; Wenzel, P.; Seeliger, M.W.; Leone, G.; Bremner, R. Rb-Mediated Neuronal Differentiation through Cell-Cycle-Independent Regulation of E2f3a. *PLoS Biol* **2007**, *5*, e179, doi:10.1371/journal.pbio.0050179.
171. Nicolay, B.N.; Bayarmagnai, B.; Moon, N.S.; Benevolenskaya, E.V.; Frolov, M.V. Combined Inactivation of PRB and Hippo Pathways Induces Dedifferentiation in the Drosophila Retina. *PLOS Genetics* **2010**, *6*, e1000918, doi:10.1371/journal.pgen.1000918.

## BIBLIOGRAPHIE

172. Gu, W.; Schneider, J.W.; Condorelli, G.; Kaushal, S.; Mahdavi, V.; Nadal-Ginard, B. Interaction of Myogenic Factors and the Retinoblastoma Protein Mediates Muscle Cell Commitment and Differentiation. *Cell* **1993**, *72*, 309–324, doi:10.1016/0092-8674(93)90110-C.
173. Campisi, J.; d’Adda di Fagagna, F. Cellular Senescence: When Bad Things Happen to Good Cells. *Nat Rev Mol Cell Biol* **2007**, *8*, 729–740, doi:10.1038/nrm2233.
174. Narita, M.; Nuñez, S.; Heard, E.; Narita, M.; Lin, A.W.; Hearn, S.A.; Spector, D.L.; Hannon, G.J.; Lowe, S.W. Rb-Mediated Heterochromatin Formation and Silencing of E2F Target Genes during Cellular Senescence. *Cell* **2003**, *113*, 703–716, doi:10.1016/S0092-8674(03)00401-X.
175. Levine, B.; Kroemer, G. Autophagy in the Pathogenesis of Disease. *Cell* **2008**, *132*, 27–42, doi:10.1016/j.cell.2007.12.018.
176. Qu, X.; Yu, J.; Bhagat, G.; Furuya, N.; Hibshoosh, H.; Troxel, A.; Rosen, J.; Eskelinen, E.-L.; Mizushima, N.; Ohsumi, Y.; et al. Promotion of Tumorigenesis by Heterozygous Disruption of the *Beclin 1* Autophagy Gene. *J Clin Invest* **2003**, *112*, 1809–1820, doi:10.1172/JCI20039.
177. Jiang, H.; Martin, V.; Gomez-Manzano, C.; Johnson, D.G.; Alonso, M.; White, E.; Xu, J.; McDonnell, T.J.; Shinjima, N.; Fueyo, J. The RB-E2F1 Pathway Regulates Autophagy. *Cancer Research* **2010**, *70*, 7882–7893, doi:10.1158/0008-5472.CAN-10-1604.
178. Manickavinayaham, S.; Dennehey, B.K.; Johnson, D.G. Direct Regulation of DNA Repair by E2F and RB in Mammals and Plants: Core Function or Convergent Evolution? *Cancers* **2021**, *13*, 934, doi:10.3390/cancers13050934.
179. Chen, J.; Zhu, F.; Weeks, R.L.; Biswas, A.K.; Guo, R.; Li, Y.; Johnson, D.G. E2F1 Promotes the Recruitment of DNA Repair Factors to Sites of DNA Double-Strand Breaks. *Cell Cycle* **2011**, *10*, 1287–1294, doi:10.4161/cc.10.8.15341.
180. Liu, K.; Lin, F.-T.; Ruppert, J.M.; Lin, W.-C. Regulation of E2F1 by BRCT Domain-Containing Protein TopBP1. *Molecular and Cellular Biology* **2003**, *23*, 3287–3304, doi:10.1128/MCB.23.9.3287-3304.2003.
181. Vélez-Cruz, R.; Manickavinayaham, S.; Biswas, A.K.; Clary, R.W.; Premkumar, T.; Cole, F.; Johnson, D.G. RB Localizes to DNA Double-Strand Breaks and Promotes DNA End Resection and Homologous Recombination through the Recruitment of BRG1. *Genes Dev.* **2016**, *30*, 2500–2512, doi:10.1101/gad.288282.116.
182. Biedermann, S.; Harashima, H.; Chen, P.; Heese, M.; Bouyer, D.; Sofroni, K.; Schnittger, A. The Retinoblastoma Homolog RBR1 Mediates Localization of the Repair Protein RAD51 to DNA Lesions in Arabidopsis. *The EMBO Journal* **2017**, *36*, 1279–1297, doi:10.15252/embj.201694571.
183. Jiang, Y.; Yam, J.C.; Tham, C.C.; Pang, C.P.; Chu, W.K. RB Regulates DNA Double Strand Break Repair Pathway Choice by Mediating CtIP Dependent End Resection. *International Journal of Molecular Sciences* **2020**, *21*, 9176, doi:10.3390/ijms21239176.
184. Ogiwara, H.; Kohno, T. CBP and P300 Histone Acetyltransferases Contribute to Homologous Recombination by Transcriptionally Activating the BRCA1 and RAD51 Genes. *PLOS ONE* **2012**, *7*, e52810, doi:10.1371/journal.pone.0052810.
185. Trouche, D.; Cook, A.; Kouzarides, T. The CBP Co-Activator Stimulates E2F1/DP1 Activity. *Nucleic Acids Research* **1996**, *24*, 4139–4145, doi:10.1093/nar/24.21.4139.
186. Desvoyes, B.; Gutierrez, C. Roles of Plant Retinoblastoma Protein: Cell Cycle and Beyond. *EMBO J* **2020**, *39*, doi:10.15252/embj.2020105802.
187. Loubiere, V.; Delest, A.; Schuettengruber, B.; Martinez, A.-M.; Cavalli, G. Chromatin Immunoprecipitation Experiments from Whole Drosophila Embryos or Larval Imaginal Discs. *Bio Protoc* **2017**, *7*, e2327, doi:10.21769/BioProtoc.2327.
188. de Noiron, J.; Hoareau, M.; Colin, J.; Guénel, I. Apoptosis Quantification in Tissue: Development of a Semi-Automatic Protocol and Assessment of Critical Steps of Image Processing. *Biomolecules* **2021**, *11*, 1523, doi:10.3390/biom11101523.
189. Zweidler, A.; Cohen, L.H. Large-Scale Isolation and Fractionation of Organs of Drosophila Melanogaster Larvae. *J. Cell Biol.* **1971**, *51*, 240–248.
190. Fristrom, J.W.; Mitchell, H.K. The Preparative Isolation of Imaginal Discs from Larvae of Drosophila Melanogaster. *The Journal of Cell Biology* **1965**, *27*, 445–448, doi:10.1083/jcb.27.2.445.

191. Marty, F.; Rockel-Bauer, C.; Simigdala, N.; Brunner, E.; Basler, K. Large-Scale Imaginal Disc Sorting: A Protocol for “Omics”-Approaches. *Methods* **2014**, *68*, 260–264, doi:10.1016/j.ymeth.2014.04.005.
192. Hoareau, M.; de Noiron, J.; Colin, J.; Guénel, I. Mass Purification Protocol for *Drosophila Melanogaster* Wing Imaginal Discs: An Alternative to Dissection to Obtain Large Numbers of Disc Cells. *Biology (Basel)* **2022**, *11*, 1384, doi:10.3390/biology11101384.
193. Brand, A.H.; Manoukian, A.S.; Perrimon, N. Ectopic Expression in *Drosophila*. *Methods Cell Biol* **1994**, *44*, 635–654, doi:10.1016/s0091-679x(08)60936-x.
194. Galaxy Available online: <https://usegalaxy.org/> (accessed on 28 December 2022).
195. Analysis of ChIP-Seq Data - Galaxy Community Hub Available online: <https://galaxyproject.org/tutorials/chip/> (accessed on 28 December 2022).
196. Iyer, E.P.R.; Iyer, S.C.; Sulkowski, M.J.; Cox, D.N. Isolation and Purification of *Drosophila* Peripheral Neurons by Magnetic Bead Sorting. *J Vis Exp* **2009**, doi:10.3791/1599.
197. Klebes, A.; Biehs, B.; Cifuentes, F.; Kornberg, T.B. Expression Profiling of *Drosophila* Imaginal Discs. *Genome Biol* **2002**, *3*, research0038.1-research0038.16.
198. Parekh, S.; Ziegenhain, C.; Vieth, B.; Enard, W.; Hellmann, I. The Impact of Amplification on Differential Expression Analyses by RNA-Seq. *Sci Rep* **2016**, *6*, 25533, doi:10.1038/srep25533.
199. Picard Tools - By Broad Institute Available online: <http://broadinstitute.github.io/picard/> (accessed on 19 January 2023).
200. Dobin, A.; Davis, C.A.; Schlesinger, F.; Drenkow, J.; Zaleski, C.; Jha, S.; Batut, P.; Chaisson, M.; Gingeras, T.R. STAR: Ultrafast Universal RNA-Seq Aligner. *Bioinformatics* **2013**, *29*, 15–21, doi:10.1093/bioinformatics/bts635.
201. Varet, H.; Brillet-Guéguen, L.; Coppée, J.-Y.; Dillies, M.-A. SARTools: A DESeq2- and EdgeR-Based R Pipeline for Comprehensive Differential Analysis of RNA-Seq Data. *PLoS One* **2016**, *11*, e0157022, doi:10.1371/journal.pone.0157022.
202. Georgette, D.; Ahn, S.; MacAlpine, D.M.; Cheung, E.; Lewis, P.W.; Beall, E.L.; Bell, S.P.; Speed, T.; Manak, J.R.; Botchan, M.R. Genomic Profiling and Expression Studies Reveal Both Positive and Negative Activities for the *Drosophila* Myb MuvB/DREAM Complex in Proliferating Cells. *Genes Dev.* **2007**, *21*, 2880–2896, doi:10.1101/gad.1600107.
203. Ashburner, M.; Ball, C.A.; Blake, J.A.; Botstein, D.; Butler, H.; Cherry, J.M.; Davis, A.P.; Dolinski, K.; Dwight, S.S.; Eppig, J.T.; et al. Gene Ontology: Tool for the Unification of Biology. *Nat Genet* **2000**, *25*, 25–29, doi:10.1038/75556.