



UNIVERSITAT DE  
BARCELONA

## The regulatory landscape of *Drosophila* imaginal disc regeneration

Elena Vizcaya Molina



Aquesta tesi doctoral està subjecta a la llicència **Reconeixement- NoComercial – SenseObraDerivada 4.0. Espanya de Creative Commons.**

Esta tesis doctoral está sujeta a la licencia **Reconocimiento - NoComercial – SinObraDerivada 4.0. España de Creative Commons.**

This doctoral thesis is licensed under the **Creative Commons Attribution-NonCommercial-NoDerivs 4.0. Spain License.**

Departament de Genètica, Microbiologia i Estadística  
Programa de Doctorat en Genètica  
Facultat de Biologia  
Universitat de Barcelona

# The regulatory landscape of *Drosophila* imaginal disc regeneration

Memòria presentada per l'

**Elena Vizcaya Molina**

Per optar al grau de

**Doctora**

per la Universitat de Barcelona



UNIVERSITAT DE  
BARCELONA

Tesi doctoral realitzada sota la direcció de la Dra. Montserrat Corominas Guiu.

Realitzada al Departament de Genètica, Microbiologia i Estadística  
de la Facultat de Biologia de la Universitat de Barcelona.

La directora i tutora,

L' autora,

Dra. Montserrat Corominas Guiu

Elena Vizcaya Molina

Barcelona, Setembre 2018



## FRONT COVER

---

DNA double helixes represent Damage Responsive Regulatory Elements (DRREs). The dotted and scattered shape illustrates damage and regeneration. The colors go from cyan to magenta to represent the two DRREs types found in this work according to the thesis color code, cyan corresponds to iDRREs and magenta corresponds to eDRREs.



*“Nothing in life is to be feared,  
it is only to be understood”*

Marie Curie



## ACKNOWLEDGEMENTS

---

I començo donant les gràcies a les persones que m'han guiat en aquesta aventura. **Montse**, podria agrair-te tota la ciència que m'has arribat a ensenyar, però res es compara amb com has confiat en mi; gràcies per ensenyar-me a volar i no tallar-me mai les ales. Gràcies **Florenci**, per totes les discussions, lliçons i consells (entre pomes i xocolata) que m'han ajudat a créixer tots aquests anys. Gràcies **Roderic**, per confiar en nosaltres; sense la teva ajuda aquesta tesi hagués sigut molt diferent. **Cecilia**, trobar-te en aquest camí ha sigut la meva gran sort; per formar sempre tan bon equip amb mi i fer que aquest treball sigui tant meu con teu, infinites gràcies! I finalment, gràcies **Marina i Sílvia**, per formar-me quan vaig arribar i per sempre tenir un moment per respondre a tots els meus dubtes.

Pero ¿qué sería de mis días en el lab sin los mejores compis del mundo? Por todas las juervezas que se nos han ido de las manos, por las risas, por las charlas y por hacer que estar en el lab fuera estar en casa. Gracias a **Sandy y Paula**, mis hot(dooooor) fly girls, porque todo lo llenáis de risas y alegría (y suspiros y refranes) y siempre siempre estáis ahí, habéis sido lo mejor de esta aventura! **Haritz**, mi partner in crime de principio a fin: del primer al último día y del primer al último consejo, eskerrrik asko laguna-patxi! **Qi**, porque de entre los millones que sois, seguirías siendo mi china favorita; gracias por aguantar todas mis tonterías más tontas y no huir haciendo la croqueta. **José**, gracias por saber echarle el freno a mi mundo acelerado y por enseñarme que el tiempo no corre, que esa soy yo! **Carlos, Giacomo y Martí**, gracias por vuestra paciencia infinita y por aguantar todas mis chapas, dadle larga vida al fly lab! Y a mis dos ex-mosquis, **Irene M y Irene P**, por todas los ratos de desconexión intentando arreglar el mundo entre dosis de glucosa y cafeína, gracias!

Als meus somriures més bonics: **Clau, Mariauri, Cris Canet, Cris Martin, Berti i Betty**. Com bé diu una cançó "I contra el mur morir-nos de riure, ser iguals vol dir ser igual de diferents". Gràcies per ser com sou, per omplir d'alegria les birres de divendres en les que ja no podia més, per ser-hi sempre i per fer que la vida valgui la pena. Ens queda molta ruta tortuga per recórrer juntes!

Pero si a alguien le tengo que agradecer más que a nadie es a mi familia. Porque si no fuera por vosotros no sería ni la mitad de lo que soy. Gracias **Papa**, por ser mi superhéroe con sonrisa de niño; por haberme transmitido tu pasión al hacer las cosas, por enseñarme a aprender desaprendiendo y a saber que lo importante son siempre las pequeñas cosas. Gracias **Mama**, por ser el espejo en el que siempre me quiero mirar; por enseñarme a luchar, a ser fuerte y a brillar siempre siempre con luz propia. Gracias **Co**, por ser mi pilar en el mundo; por cuidarme y ser mi refugio, por enseñarme a creer en mi misma y por no dudar nunca de mi. Gracias **Yaya**, por guiarme desde lejos; por enseñarme que el don de la cabezonería y la testarudez te pueden llevar donde tú quieras. Sé que con vosotros a mi lado puedo conseguir todo lo que me proponga. Soy la persona más afortunada del mundo por teneros cerca! Os quiero mucho!



# TABLE OF CONTENTS

---

<b>FIGURE INDEX</b> .....	1
<b>TABLE INDEX</b> .....	5
<b>ABBREVIATIONS AND ACRONYMS</b> .....	7
<b>INTRODUCTION</b> .....	11
- Regeneration	
- Dynamics of gene regulation	
- Towards the regulatory genome of regeneration	
<b>OBJECTIVES</b> .....	35
<b>MATERIALS AND METHODS</b> .....	39
<b>RESULTS</b> .....	67
- Chapter I: The transcriptome of regeneration	
- Chapter II: The regulome of regeneration	
- Chapter II: Conservation of the regeneration regulatory logic across metazoans	
<b>DISCUSSION</b> .....	113
<b>CONCLUSIONS</b> .....	129
<b>BIBLIOGRAPHY</b> .....	123
<b>ANNEX</b> .....	143
- Annex I: Experiment genotypes	
- Annex II: Statistics and replicate analysis	
- Annex III: Clusters and hotspots lists	
- Annex IV: DRREs motif discovery through time	



# FIGURES INDEX

---

## INTRODUCTION

<b>Figure 1</b> - Levels of biological organization in regeneration .....	16
<b>Figure 2</b> - Early regeneration signals .....	17
<b>Figure 3</b> - Sources of new cells in regeneration .....	18
<b>Figure 4</b> - Developmental mechanisms of regeneration .....	19
<b>Figure 5</b> - The polar coordinated model .....	19
<b>Figure 6</b> - Phylogenetic tree of regeneration .....	20
<b>Figure 7</b> - An overview of the regulatory landscape of transcription .....	23
<b>Figure 8</b> - Chromatin accessibility models .....	25
<b>Figure 9</b> - Enhancer and promoter states .....	26
<b>Figure 10</b> - Models for TF binding .....	27
<b>Figure 11</b> - Enhancer transcription .....	27
<b>Figure 12</b> - Models for close spatial proximity .....	28
<b>Figure 13</b> - <i>Drosophila</i> imaginal discs .....	31
<b>Figure 14</b> - Overview of regeneration in the wing imaginal discs .....	33

## MATERIALS AND METHODS

<b>Figure 15</b> - <i>In vivo</i> genetic ablation using binary systems .....	51
<b>Figure 16</b> - <i>Ex vivo</i> culture of imaginal discs .....	52
<b>Figure 17</b> - Test for regenerated adult wings .....	52
<b>Figure 18</b> - RNA-seq analysis workflow .....	54
<b>Figure 19</b> - ATAC-seq technique .....	55
<b>Figure 20</b> - ATAC-seq analysis workflow .....	57
<b>Figure 21</b> - ChIP-seq technique .....	58
<b>Figure 22</b> - ChiP-seq analysis workflow .....	59
<b>Figure 23</b> - Validation of DRRE activity .....	60
<b>Figure 24</b> - 3C-qPCR technique .....	62
<b>Figure 25</b> - Reusage analysis workflow .....	64
<b>Figure 26</b> - DRREs conservation analysis workflow .....	65

<b>Figure 27</b> - Regenerative genes discovery workflow.....	66
<b>RESULTS</b>	
<b>Figure 28</b> - Experimental design .....	69
<b>Figure 29</b> - Regeneration stages description .....	70
<b>Chapter I: The transcriptome of regeneration</b>	
<b>Figure 30</b> - Differentially expressed genes after induction of cell death .....	73
<b>Figure 31</b> - Validation of DE genes.....	74
<b>Figure 32</b> - Time-course GO.....	75
<b>Figure 33</b> - Expression profiles of upregulated transcription factors.....	76
<b>Figure 34</b> - iREGULON .....	77
<b>Figure 35</b> - Signaling pathways upregulated in regeneration .....	78
<b>Figure 36</b> - Expression of DE pathway genes through time .....	79
<b>Figure 37</b> - Requirement of mTOR pathway in regeneration .....	80
<b>Figure 38</b> - Genomic distribution of clusters.....	81
<b>Figure 39</b> - The GstD cluster.....	82
<b>Figure 40</b> - Cluster size analysis.....	82
<b>Figure 41</b> - Gene expression in clusters .....	83
<b>Figure 42</b> - Biological processes related to clustered genes .....	84
<b>Figure 43</b> - Clusters containing signaling pathway members .....	84
<b>Chapter II: The regulome of regeneration</b>	
<b>Figure 44</b> - Differentially accessible chromatin after induction of cell death .....	87
<b>Figure 45</b> - Correlation of RNA-seq and ATAC-seq data .....	88
<b>Figure 46</b> - Genomic distribution of DRREs .....	89
<b>Figure 47</b> - Accessible chromatin landscape after cell death induction .....	90
<b>Figure 48</b> - Accessible chromatin landscape after cell death induction through time.....	91
<b>Figure 49</b> - Chromatin features of DRREs .....	92
<b>Figure 50</b> - Chromatin features of DRREs by genomic distribution .....	93
<b>Figure 51</b> - ChIP-seq validation of DRREs.....	94
<b>Figure 52</b> - Activity validation of DRREs after physical injury.....	95

<b>Figure 53</b> - Activity validation of DRREs after cell death induction.....	96
<b>Figure 54</b> - Chromatin architecture rearrangements in regeneration.....	98
<b>Figure 55</b> - DRREs used in other tissues and at other developmental stages .....	99
<b>Figure 56</b> - Tissue usage of DRREs .....	100
<b>Figure 57</b> - Motif enrichment in DRREs .....	101

### **Chapter III: Conservation of the regeneration regulatory logic across metazoans**

<b>Figure 58</b> - Homology of fly regenerative genes .....	105
<b>Figure 59</b> - Conservation of regenerative genes .....	106
<b>Figure 60</b> - Requirement of shared TFs in fly regeneration.....	108
<b>Figure 61</b> - Conservation of DRREs .....	109
<b>Figure 62</b> - Activation of DRREs in other species.....	110
<b>Figure 63</b> - Reusage of DRREs in zebrafish.....	111

### **DISCUSSION**

<b>Figure 64</b> - Genome-wide workflow to study regeneration .....	116
<b>Figure 65</b> - DRRE types .....	119
<b>Figure 66</b> - A model for enhancer activation .....	122
<b>Figure 67</b> - Genomic clustering .....	126

### **ANNEX**

<b>Figure 68</b> - Statistics and replicate analyses of RNA-seq.....	151
<b>Figure 69</b> - Statistics and replicate analyses of ATAC-seq.....	152
<b>Figure 70</b> - Statistics and replicate analyses of third instar larval ATAC-seq.....	153
<b>Figure 71</b> - Statistics and analysis of ChIP-seq .....	153



## TABLES INDEX

---

### MATERIALS AND METHODS

<b>Table 1</b> - Drosophila strains used .....	43
<b>Table 2</b> - Reagents .....	44
<b>Table 3</b> - qPCR primers .....	45
<b>Table 4</b> - Genome-wide data used.....	46
<b>Table 5</b> - Softwares used .....	47
<b>Table 6</b> - Reporter lines features .....	60

### RESULTS

<b>Table 7</b> - List of conserved TFs upregulated in regeneration .....	107
--	-----

### ANNEX

<b>Table 8</b> - Genotypes List .....	147
<b>Table 10</b> - Motif discovery in early DRREs .....	167
<b>Table 11</b> - Motif discovery in mid DRREs .....	168
<b>Table 12</b> - Motif discovery in late DRREs .....	169



## ABBREVIATIONS AND ACRONYMS

---

**3C:** Chromosome Conformation Capture

**A compartment:** Active compartment

**AEL:** After egg laying

**ap:** *apterous*

**ATAC-seq:** Assay for Transposase-Accessible Chromatin followed by sequencing

**Atf-3:** *activating transcription factor 3*

**B compartment:** Repressive compartment

**BEAF-32:** *Boundary element-associated factor of 32kD*

**Bmp5:** *Bone Morphogenetic Protein 5*

**bp:** basepair

**cbt:** *cabut*

**ChIP-seq:** Chromatin Immunoprecipitation followed by sequencing

**ci:** *cubitus interruptus*

**CNS:** Central Nervous System

**COF:** Cofactor

**CP:** Core Promoter

**Ctrl:** Control

**D. ana:** *Drosophila ananassae*

**D. pse:** *Drosophila pseudoobscura*

**D. will:** *Drosophila willistoni*

**D. yak:** *Drosophila yakuba*

**DE:** Differentially expressed

**DEG:** Differentially expressed genes

**dia:** *diaphanous*

**Dif:** *Dorsal immunity factor*

**Dilp8:** *Drosophila Insulin-like peptide 8*

**dl:** *dorsal*

**dlg1:** *discs large 1*

**dome:** *domeless*

**down:** downregulated

**dpp:** *decapentaplegic*

**Dref:** *DNA replication-related element factor*

**DRRE:** Damage Responsive Regulatory Element

**dsh:** *dishevelled*

**dSRF:** *Drosophila Serum Response Factor, blistered*

**eDRRE:** emerging Damage Responsive Regulatory Element

**EGFR:** Epidermal Growth Factor Receptor

**eRNA:** enhancer-RNA

**FAIREseq:** Formaldehyde-Assisted Isolation of Regulatory Elements followed by sequencing

**FC:** foldchange  
**FI:** First Intron  
**FPKM:** Fragments per Kilobase Million  
*fru:* fruitless  
**Gadd45:** Growth arrest and DNA damage-inducible 45  
**Gal4:** Galactose-induced gene A  
**Gal80<sup>TS</sup>:** Gal80 temperature sensitive  
**GEO:** Gene Expression Omnibus  
**GFP:** Green Fluorescent Protein  
**GO:** Gene Ontology  
*grh:* grainy head  
**GRN:** Gene Regulatory Network  
**GstD:** Glutathione S transferase  
**H3K27ac:** Histone 3 Lysine 27 acetylation  
**H3K27me3:** Histone 3 Lysine 27 Tri-methylation  
**H3K4me1:** Histone 3 Lysine 4 Mono-methylation  
**H3K4me3:** Histone 3 Lysine 4 Tri-methylation  
**H3K9me3:** Histone 3 Lysine 9 Tri-methylation  
**HEAL:** Hartford Engineering a Limb  
**HiC:** Chromosome Conformation Capture followed by high throughput sequencing  
*hop:* hopscotch  
**iDRRE:** increasing Damage Responsive Regulatory Element  
**Jak-STAT:** Janus kinase - signal transducer and activator  
**JMJD3:** JmjC domain-containing protein 3  
**JNK:** c-Jun-NH2-terminal kinase  
*Jra:* Jun- related antigen  
*jub:* Ajuba LIM protein  
**Kb:** Kilobase  
**L3:** Larvae 3  
**LamC:** LaminC  
**LexO:** LexA operator  
**Lgr3:** Leucine-rich repeat-containing G protein-coupled receptor 3  
**LHG:** LexA- hinge- Gal4 activation domain  
*lilli:* lilliputian  
*Irch:* leucine-rich-repeats and calponin homology domain protein  
**MAPK:** Mitogen Activated Protein Kinase  
*Mer:* Merlin  
**Mmp1:** Matrix metalloproteinase 1  
**MN:** Mononucleosome  
*mol:* moladietz  
**mTOR:** mammalian Target of Rapamycin

**Myb:** *Myb oncogene-like*

**NA:** Not associated

**NCBI:** National Center for Biotechnology Information

**NDE:** not differentially expressed

**NF:** Nucleosome Free

**NOS:** Nitric Oxide Synthase

**p38a:** *p38a MAP kinase*

**P-Akt:** phospho Akt

**PcG:** Polycomb Group

**PF:** Pioneer Factor

**P-H3:** phospho Histone 3

**phol:** *pleiohomeotoic like*

**PI3K:** Phosphonositol-3 Kinase

**pigs:** *pickled eggs*

**PlexA:** *Plexin A*

**pnr:** *pannier*

**PoI-II:** RNA Polymerase II

**PoI-II ser5P:** Polymerase II phosphorylated in Serine 5

**PPI:** Protein-Protein Interaction

**PRE:** Polycomb Response Element

**Pten:** *Phosphatase and tensin homolog*

**PTM:** Post Translational Modification

**PVDF:** Polyvinylidene difluoride

**qPCR:** quantitative Polymerase Chain Reaction

**Reg:** Regeneration

**RNAi:** RNA interference

**ROS:** Reactive Oxygen Species

**rpr:** *reaper*

**RpS18:** *Ribosomal protein S18*

**S6k:** *Ribosomal protein S6 kinase*

**sal/salm:** *spalt major*

**sal<sup>E/PV</sup>:** *spalt<sup>E/PV</sup>*

**sd:** *scalloped*

**SEM:** Standard error of the mean

**SNB:** Subnuclear body

**Sply:** *Sphingosine-1-phosphate lyase*

**STARR-seq:** Self-Transcribing activatie Regulatory Region followed by sequencing

**Stat92E:** *Signal transducer and activator at 92*

**tara:** *taranis*

**TF:** Transcription factor

**TRE:** Trithorax Response Element

***trithorax***: *trx*

***trl***: *thritoxax-related, GAGA*

**TrxG**: Trithorax group

**TSS**: Transcription Start Site

**TUNEL**: Terminal deoxynucleotidyl transferase dUTP nick end labeling

**UAS**: Upstream activation sequences

**UCSC**: University of California Santa Cruz

**Up**: upregulated

***upd3***: *unpaired 3*

***utx***: *utx histone demethylase*

**VDRC**: Vienna Drosophila Stocks Center

***wg***: *wingless*

***w***: *white*

***yki***: *yorkie*

# INTRODUCTION





*Since ancient times regeneration has sparked endless curiosity in the human being.*

*There was a time when regeneration was all legends and mysticism. A time when Prometheus was condemned to see how his liver was eaten by an eagle every day and was recovered every night. The same time when a multi-headed creature called Hydra was able to grow back two heads every time one was lost.*

*Then, there was a time to find regeneration in the workshops of alchemists, always in the pursuit of the Elixir of Life and immortality. But there was also a time for controversy, when regeneration filled sophists' thoughts: "if we only have one indivisible soul, where is it going when an animal is cut in pieces?"*

*And finally there was, and still is, a time for science. Time to ask, to observe, to answer and to learn. Great scientists, like Lazzaro Spallanzani or Thomas Morgan, started what we know as the regenerative field. Over the years, great advances have been achieved; yet there are, still, many unsolved questions and long way to go to find more answers.*





# REGENERATION

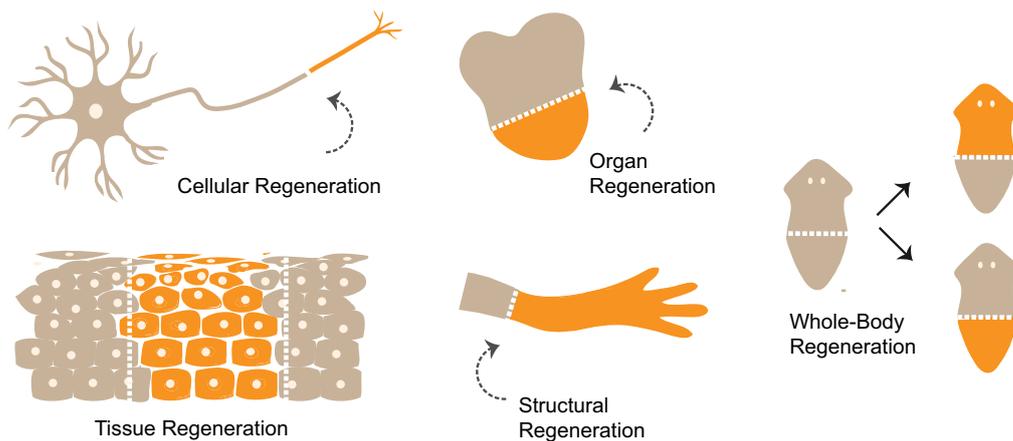
---

In 1901, Thomas Morgan defined regeneration as the replacement of missing structures following injury (Morgan 1901). Nowadays, regeneration is known as the ability to reconstruct the original shape, size, and function of body parts that have been lost or damaged to restore proper homeostasis. Commonly, regeneration is thought of as the replacement of big body parts, which is known as **reparative regeneration**. Nonetheless, there is also **physiological regeneration**, which is based on the continuous self-renewal of cells. Epithelia, organs, and tissues in general are constantly subjected to insults disrupting tissue homeostasis, which needs to be reestablished to maintain healthy individuals (Reviewed in Iisma et al. 2018).

Research in regenerative medicine seeks to unravel how both reparative and physiological regeneration work. Which are the molecules triggering regeneration? How are they orchestrated and regulated? Why do some animals regenerate better than others? What can we learn from nature to transform a non-regenerating animal into a regenerating one? We are still far from knowing how to restore missing body parts of a non-regenerating animal; however, since the first regeneration event described in 1712, in which de Réaumur described limb regeneration in crustaceans (De Réaumur 1712), we have greatly advanced.

The ability of regeneration is widely and randomly distributed in the animal kingdom (Reviewed in Sanchez-Alvarado and Tsonis 2006; Bely and Nyberg 2010). Accordingly, five main levels of organization have been proposed to describe regeneration across metazoans, ranging from a single cell type to the whole body: cellular, tissue, organ, structural and whole-body regeneration (Fig. 1) (Reviewed in Bely and Nyberg 2010; Slack 2017). **Cellular regeneration** references to the recovery of a cell part by regrowth, as it occurs in nerve axon regeneration. **Tissue regeneration** is considered as the closure of gaps in a given homogeneous cell population, which happens, for instance, during recovery of the skin epithelium after a cut. **Organ regeneration** is known as the size restoration of an organ, which often comprises multiple cell types and takes place, for example, in the liver after an hepatectomy. **Structural regeneration** refers to the appendage regeneration found in arthropods and vertebrates. Such kind of biological organization requires a pattern formation, but always in a distalwards direction. Finally, **whole-body regeneration** is the ability to reconstruct heads and tails from small body fragments as it occurs in planarian regeneration.

Moreover, regenerative capacity is regulated by a number of fundamental traits, including age, body size, life-stage, growth pattern, wound healing response and re-epithelialization among others (Reviewed in Seifert et al. 2012). For example, aging negatively affects regenerative capacity as a result of cellular senescence and telomere shortening. Also it impairs re-epithelialization, as is it evident from healing by scar formation in older mammals but not their fetal counterparts (Reviewed in Iismaa et al. 2018).



**Figure 1 - Levels of biological organization in regeneration.** Drawing depicting the five levels of biological organization in regeneration: cellular regeneration, tissue regeneration, organ regeneration, structural regeneration and whole body regeneration (Adapted from Slack 2017).

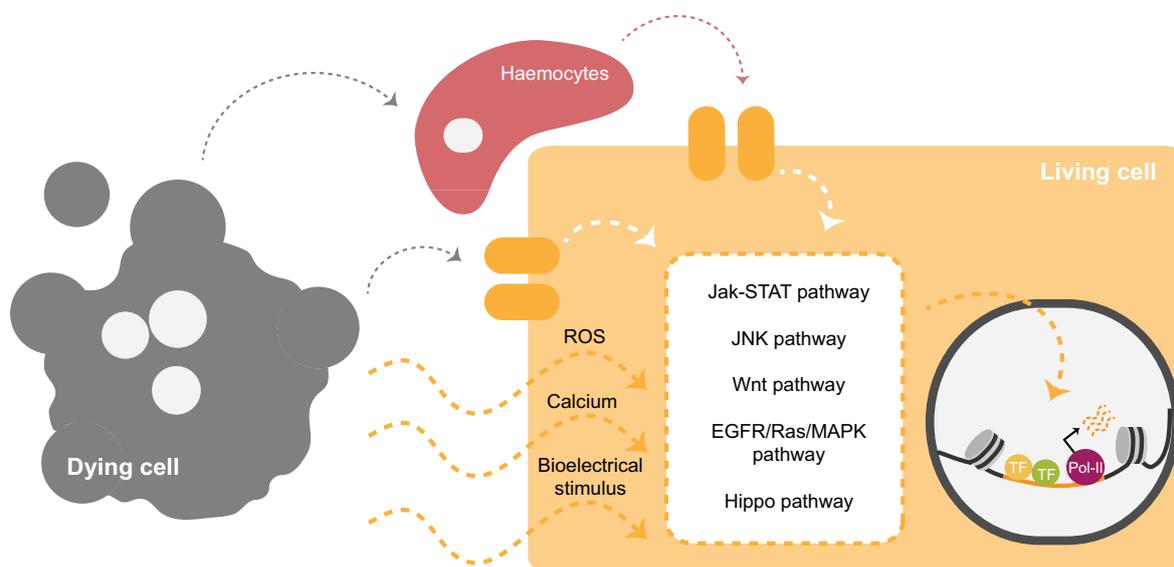
## THE MOLECULAR AND CELLULAR BASIS OF REGENERATION

Although the ability to regenerate greatly varies between tissues, organs and across species, the molecular mechanisms underlying tissue repair are remarkably conserved. Such mechanisms are common not only across different phyla but also at different levels of biological organization, (Bely and Nyberg 2010). In all instances, regeneration requires first to **sense damage** which is followed by the **wound formation**. Together, they represent the **onset** of regeneration. Subsequently, **production of new cells** is needed to recover the damaged area. The number and type of cells to be produced ultimately depends on the level of biological organization to be recovered. Finally, different **developmental mechanisms** are used to achieve the **reconstruction** of a new structure identical to the one lost previously.

### The onset of regeneration: early signals

In a matter of minutes or even seconds after damage is produced, **local responses** are released from dying or damaged cells and sensed as a pro-regenerative stimuli by the living ones. These signals include bioelectrical stimulus (Levin 2009), calcium waves (Razzell et al. 2013; Yoo et al. 2012), and propagation of reactive oxygen species (ROS) (Niethammer et al. 2009; Yoo et al. 2012; Santabárbara-Ruiz et al. 2015). At the same time, injury causes **inflammation**, which results in the recruitment of immune cells to the wounded area. These immune cells release cytokines, that are also sensed as pro-regenerative signals (Burzyn et al. 2013; Petrie et al. 2014; Wynn and Vannella 2016; Fogarty et al. 2016). The role of all these signals is to ultimately regulate the **activation of signalling pathways** such as the JNK pathway, Wnt pathway, Jak-STAT pathway, EGFR/Ras/MAPK pathway or Hippo Pathway (Bosch et al. 2005; Bergantiños et al. 2010; Blanco et al. 2010; Repiso et al. 2013; Sun and

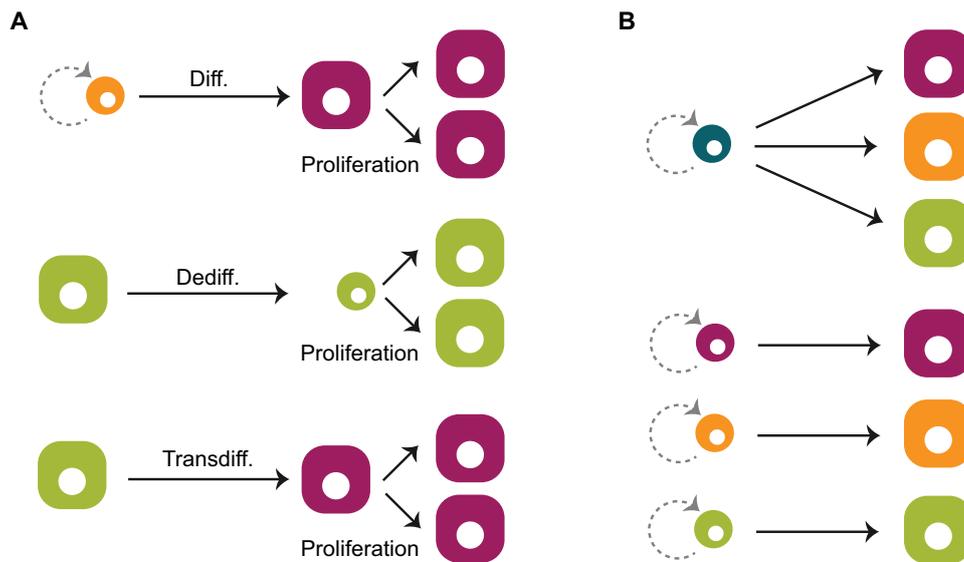
Irvine 2013; Katsuyama et al. 2015; Santabárbara-Ruiz et al. 2015; Lopez-Luque et al. 2016). Signalling is integrated in the genome by the action of effector transcription factors (TFs). Finally, the transcriptional outcome promotes the next regenerative steps (Fig. 2).



**Figure 2 - Early regeneration signals.** Drawing depicting how a living cell can sense different signals (ROS, calcium, bioelectrical stimulus, inflammatory signals and released ligands). These are released from the damaged tissue (“dying cell”), to promote regeneration. As a consequence, several signal pathways are activated. These are integrated in the nucleus to promote transcription of pro-regenerative genes.

## Production of new cells

After wounding, living cells need to proliferate to recover the damaged area. The number and type of cells to be restored, as well as the source of new cells, relies on the biological organization of the lost structure and on the species. Planarians, for example, use a population of **stem cells** called neoblasts that self-renew, generating different new cell types (Baguñà et al. 1989). Pigmented epithelial cells in the newt dorsal iris can regenerate a new lens via **transdifferentiation**: cells dedifferentiate, reenter the cell cycle, and differentiate to new lens cells (Henry and Tsonis 2010). Similarly, in *Drosophila*, committed cells from imaginal discs are able to re-specify their fate to replace the lost tissue (Repiso et al. 2013). In some cases, such as the Hydra, a combination of both, stem cells and transdifferentiation processes is required (Vogg et al. 2006). In the zebrafish heart, existing cardiomyocytes undergo **dedifferentiation** and proliferate to generate new cardiomyocytes for replacing lost heart mass (Jopling et al. 2010; Sánchez-Iranzo et al. 2018). Finally, in **compensatory proliferation**, differentiated cells simply divide, generating more cells of their kind. This occurs, for example, when hepatocytes undergo hyperplasia to recover the mammal's liver (Reviewed in Michalopoulos and DeFrances, 1997) (Fig. 3).



**Figure 3 - Sources of new cells in regeneration.** (A) New cells are formed by differentiation of stem cells and dedifferentiation or transdifferentiation of already differentiated cells. In all the cases normal proliferation occurs. (B) Different organization of stem cells. One single stem cells can give rise to many differentiated cell types (if pluripotent) or each stem cell can differentiate into one single cell type (multipotent).

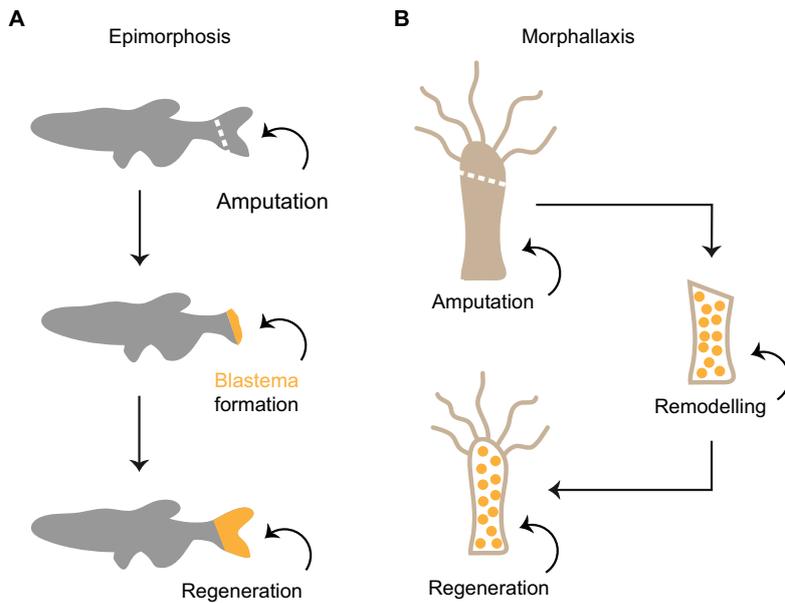
### Developmental mechanisms of regeneration

Historically, developmental regeneration mechanisms have been split in two main categories, established by Thomas Morgan in 1901: epimorphosis and morphallaxis (Reviewed in Sunderland 2010).

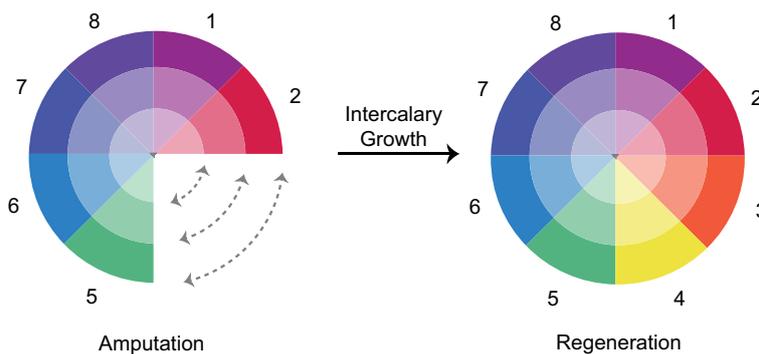
**Epimorphosis** refers to the mechanism by which the regeneration of a new part involves **proliferation**. Such proliferation occurs thanks to the **blastema** formation, a mass of morphologically undifferentiated and pluripotent cells that cover the wounded area (Fig. 4A). Epimorphosis based regeneration can be found, for instance, in zebrafish heart and fin regeneration (Poss et al. 2002). **Morphallaxis** refers to regeneration as a result of **remodelling** of existing material without proliferation, neither the formation of the blastema. At the end, morphallaxis gives rise to a smaller but well patterned organism (Fig. 4B). Hydra regeneration is one example where morphallaxis based regeneration takes place (Cummings and Bode 1984).

This classification, however, is not a real mechanistic reflection, and both, epimorphosis and morphallaxis may contribute simultaneously to a given regeneration event. This happens, for example, in **intercalary growth** mechanisms used in regeneration of amphibian limbs, cockroach legs, and *Drosophila* imaginal discs (French et al. 1978; Bryant et al. 1981; French 1981). This growth mechanism, is based on the **polar-coordinate model**, where cells have positional information for two coordinates that are continuous in a given radius. Hence, the

juxtaposition of cells from different locations can stimulate regeneration of the intervening tissue (Fig. 5).



**Figure 4 - Developmental mechanisms of regeneration.** (A) Illustration of epimorphosis upon caudal fin amputation in zebrafish. (B) Illustration of morphallaxis in Hydra regeneration.



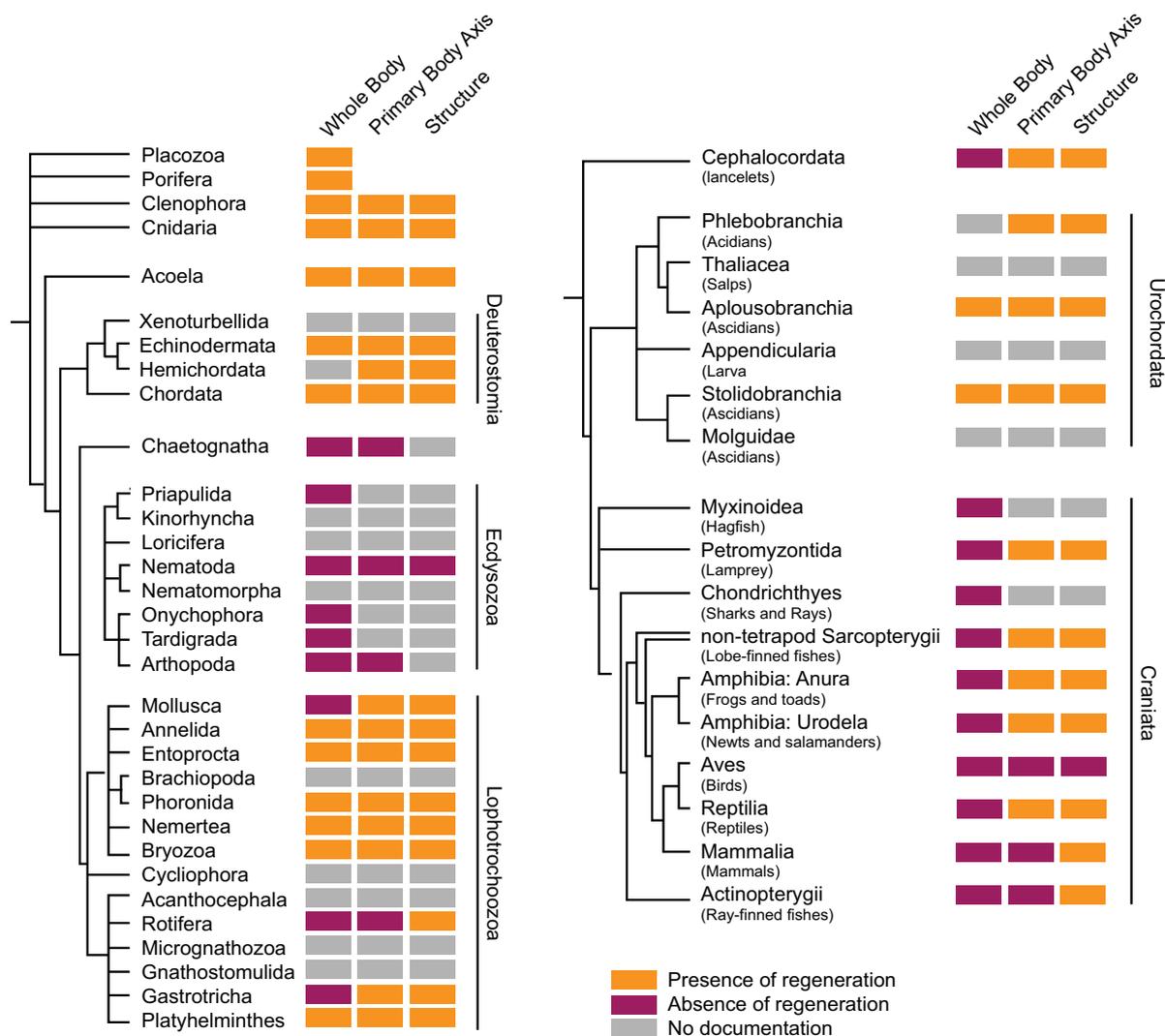
**Figure 5 - The polar coordinated model.** Series of positional values are arranged in a circle. After fragmentation, apposition of the wound edges is assumed to lead to regeneration along the values resting along the shortest path.

## MAKING THE DIFFERENCE: GOOD AND BAD REGENERATORS

Injury is unavoidable for animals, hence regeneration results in an advantageous widespread trait of survival (Brookes and Kumar 2008). Although it seems that the molecular and cellular basis of tissue repair are conserved, regeneration is not universal and greatly varies, not only between species, but also between tissues and organs or between developmental stages of the same species (Reviewed in Bely and Nyberg 2010) (Fig. 6). Moreover, it also remains unclear if regeneration involves similar molecular mechanisms that are preserved across distantly related taxa, or if the capacity to regenerate damaged tissues is a trait that has evolved repeatedly, albeit by the use of distinct regenerative pathways (Reviewed in Iismaa et al. 2018).

Planarians, for instance, are considered master-regenerators as they can reconstruct whole body animals from tiny pieces of almost any of their body parts; other platyhelminthes, however, are unable to regenerate their heads and die after head amputation (Iten and

Bryant 1973; Liu et al. 2013; Umesono et al. 2013). Such regenerative differences between close species do not only occur in high levels of biological organization, where patterning, development, and the production of many different cell types occur; regeneration following less complex levels of biological organization, as tissue regeneration, have also diverged between them. This, for example, is the case for skin regeneration in the mouse lab model (*Mus musculus*) and the african spiny mouse (*Acomys*). While the african spiny mouse perfectly regenerates skin, the mouse lab model suffers an impaired regeneration which leads to scar formation (Seifert et al. 2012). In addition, regeneration also diverges depending on the developmental stage and upon maturation of the respective species. In mammals, fetal and newborn individuals retain higher regenerative capacity, which is lost in the adult: newborn mice can heal their heart or skin better compared to adults (Porrello et al. 2011; Bullard et al. 2003). To some extent, the same occurs in some insects: the ability to regenerate specific organs at larval stages is lost in the *Drosophila* adult (Reviewed in Jaszczak and Halme 2016; Hariharan and Serras 2017).



**Figure 6 - Phylogenetic tree of regeneration.** The tree shows the presence and absence of regeneration in all phyla, ranging from whole-body to regeneration of specific structures. The cases where there is no documentation are also shown. (Adapted from Bely and Nyberg 2010).

Hence, what makes the difference between being a good or a bad regenerator? Why are some animals able to heal but not to develop a new structure? Why are some others born being good regenerators and become bad ones upon maturation? And finally, why do some individuals, lose the ability to give rise to any new structures, if they are able to do it during embryonic development?

### **Highlighting the difference: gene regulation**

Regeneration can be considered as a stepwise process in which, if one step fails, all subsequent ones fail as well (Reviewed in Roehl et al. 2018). Sensing damage is a crucial point in which many signals are activated in the living cells. The ultimate role of these signals is to be integrated into the genome to reset the transcriptional programs required in regeneration. It can be hypothesized that if an individual has the ability to develop a structure during embryonic development, it should retain the same ability during regeneration. In other words, the genes used during development are still encoded in the genome during regeneration. Nonetheless, the genome is not only composed by coding genes but also by non-coding regions, such as the regulatory elements controlling gene expression. The fact, that an individual cannot **reset the transcriptional programs** needed for regeneration could, thus, be explained by differences in the **spatial and temporal regulation** of gene expression, and not be a consequence of the genes encoded in its genome. If regenerative signals are not properly sensed or integrated into the genome, then the whole process fails. Indeed, it has been previously hypothesized that the chromatin landscape could determine the regeneration ability. Animals, that retain a flexible chromatin state could more easily reprogram gene expression to cover the emerging needs of regeneration (Reviewed in Katsuyama and Paro 2011).

Based on the hypothesis that dynamics of gene regulation and the chromatin landscape play a pivotal role in regeneration, this thesis focuses on understanding how transcription patterns can be reset upon injury.



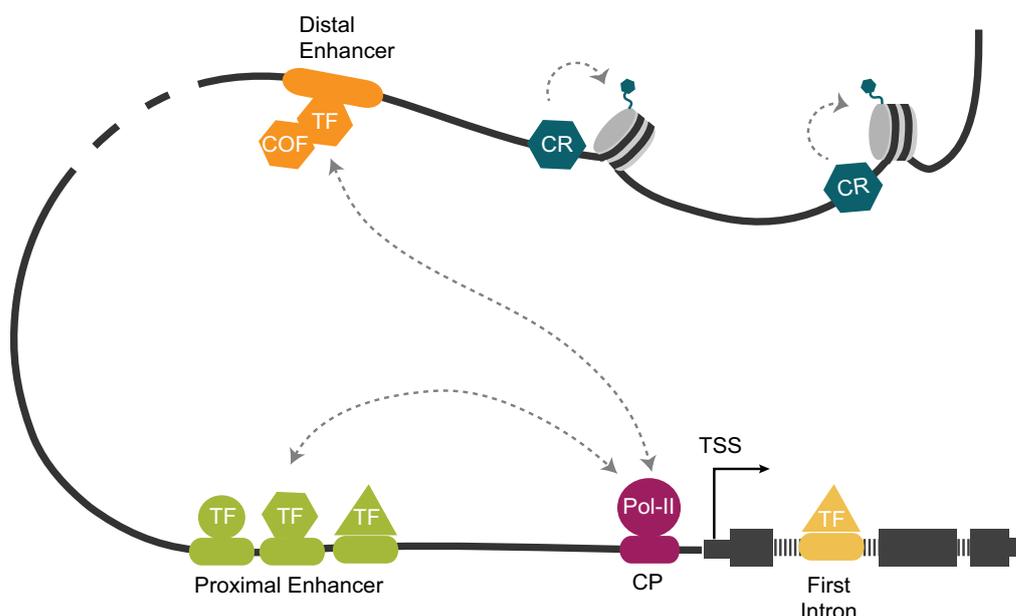
## DYNAMICS OF GENE REGULATION

In the last years many research groups have focused their efforts towards understanding how a gene is transcribed. The field has moved from studying transcription factors bound to core promoters to understanding chromatin states, non-coding elements and chromatin architecture traits, among others. Amongst this variety of elements, enhancers and their associated TFs play a leading role in the initiation of gene expression.

### GENE REGULATION BY REGULATORY ELEMENTS

Enhancers are regulatory DNA regions that, when bound by specific proteins, increase the level of transcription of an associated gene, independently of the orientation and distance to the core promoter (CP). CPs are short sequences containing the Transcription Start Site (TSS) of the gene. They indicate the transcription starting point, by recruiting the transcription machinery. Even if CPs are sufficient to recruit RNA Polymerase II (Pol-II) and drive basal levels of transcription (Orphanides et al 1996; Roeder 1996; Blackwood and Kadonaga 1998), they require regulatory elements for full activity (Banerji et al. 1981; Shlyueva et al. 2014). The regulatory information of enhancers is encoded within them as short sequences that are recognized and bound by TFs. TFs recruit cofactors (COFs) forming a complex, that ultimately mediates the recruitment of Pol-II and activation of gene transcription (Zabidi et al. 2015; Koeneke et al. 2016, Catarino and Stark 2018) (Fig. 7).

Even if the logic of gene regulation by regulatory elements is well understood, many steps and conditions are required for an enhancer to regulate a gene, ranging from chromatin opening to correct genome folding.

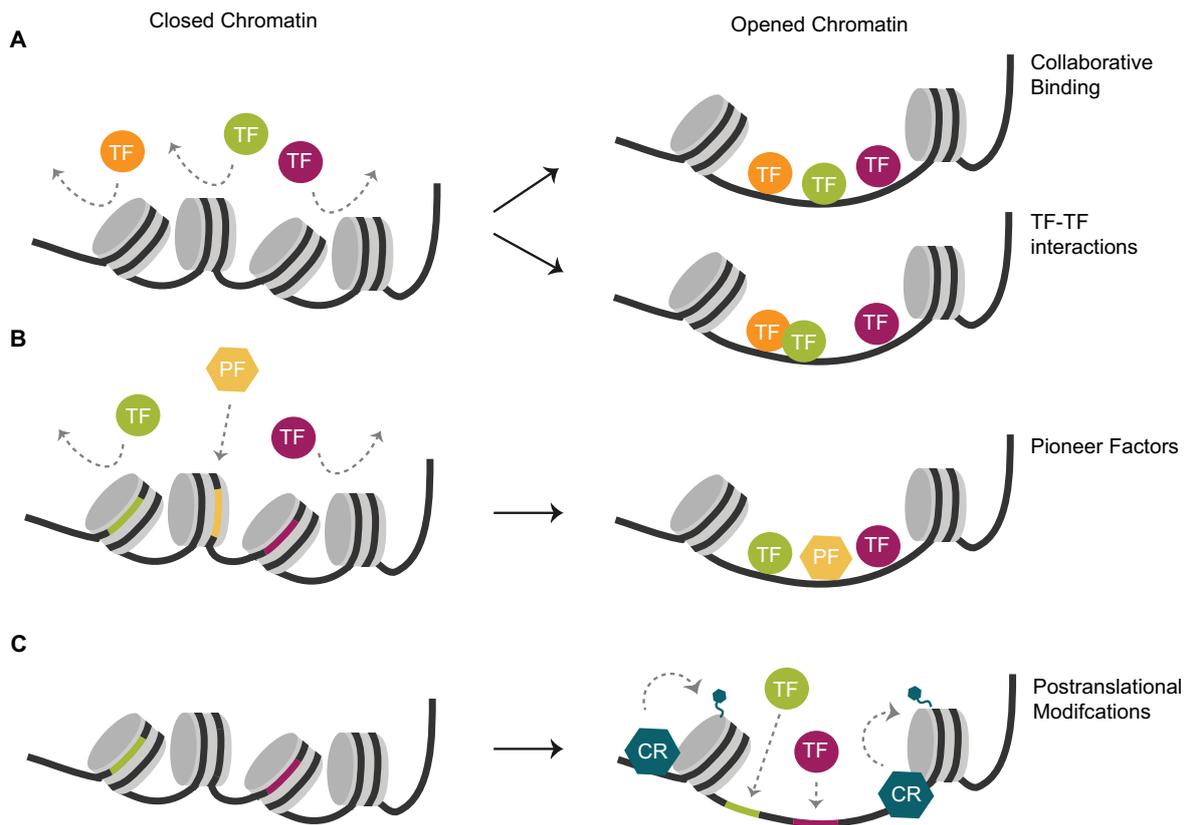


**Figure 7 - An overview of the regulatory landscape of transcription.** Transcription begins with the recruitment of the transcription machinery at the TSS. The CP, serves as a binding platform for the transcription machinery. Enhancers, from both proximal and distal positions, bind transcription factors (TFs) and cofactors (COFs) to recruit and activate Pol-II, located at the target gene promoter site. Modifications of histone tail residues surrounding regulatory elements can activate or repress gene expression by modulating the chromatin state. Finally, chromatin architecture plays a role by creating loops bringing enhancers and promoters into close contact.

### **Chromatin dynamics: enhancer accessibility**

Chromatin is a complex of macromolecules composed of DNA and histones, with the ultimate function of compacting and protecting genomic DNA (Reviewed in Venkatesh and Workman 2015). Nucleosomes are the basic chromatin unit and are formed by an octamer core of histones surrounded by 147bp of genomic DNA. They act as gatekeepers and prevent proteins, such as TFs, from accessing enhancers (Svaren et al. 1994; Walter et al. 1995). Although active enhancers are always located in an accessible position within the chromatin, enhancers *per se* are found in a default off state settled by the nucleosome positioning; they only become accessible upon given environmental conditions (Charoensawan et al. 2012; Barozzi et al. 2014). The accessibility of chromatin is, thus, a key requirement for gene regulation and is one of the most predictive features for enhancer characterization (Boyle et al. 2008).

Different mechanisms by which chromatin can become accessible have been described. One of these is the **collaborative binding**, showing a passive cooperativity between TFs that leads to eviction of nucleosomes by mass action (Reviewed in Deplanke et al. 2016). In other situations, **Protein-Protein Interactions** (PPIs), such as TF-TF or TF-COF, lead to a shift in nucleosome occupancy (Reviewed in Reiter et al. 2017) (Fig. 8A). **Pioneer factors** (PFs) are TFs that recognize and directly bind condensed chromatin, displacing the nucleosomes. The nucleosome shift allows two distinct subsequent actions: binding of other TFs and recruitment of chromatin remodelling complexes, which will lead to a more perpetuated nucleosome repositioning (Reviewed in Spitz and Furlong 2012) (Fig. 8B). Finally, **post-translational modifications** (PTMs) of histone tails, such as H3K27ac or H3K4me1, can also help to relax and open chromatin (Reviewed in Catarino and Stark 2018) (Fig. 8C).



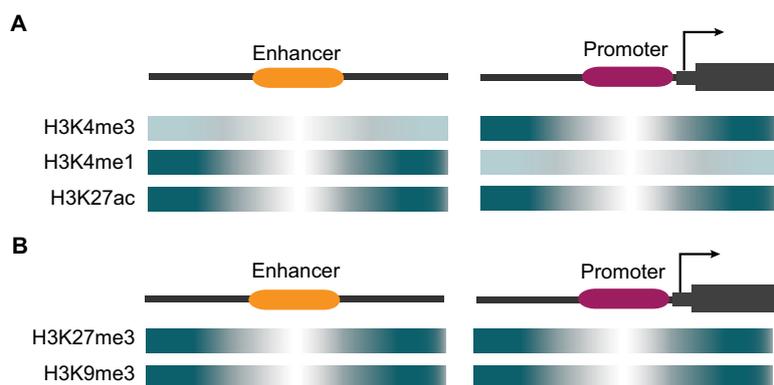
**Figure 8 - Chromatin accessibility models.** (A) Chromatin opening triggered by collaborative binding of TFs and TF-TF interactions. (B) Chromatin opening through pioneer factors. (C) Chromatin opening due to post-translational modifications triggered by chromatin remodelers (CR). (Adapted from Reiter et al. 2017)

### Enhancer activation: features and predictions

The opening of the chromatin is a requirement for enhancer activity; however, being accessible does not necessarily mean being active. Many features determining enhancer activity have been characterized, yet none of them seems to be a universal trait.

Although active regulatory elements, whether enhancers or promoters, are depleted of nucleosomes, the histones in the flanking nucleosomes often carry **PTMs**, which provide a useful readout of enhancer activity. In **active** chromatin states, promoters are usually marked with **H3K4me3**, enhancers with **H3K4me1**, and both of them with **H3K27ac** (Rada-Iglesias et al. 2011; Calo and Wysocka 2013; Shlyueva et al. 2014; Koenecke et al. 2016; Long et al. 2016) (Fig. 9A). Besides, in **silent** chromatin states, promoters and enhancers are labeled with **H3K27me3** (Reviewed in Simon and Kingston 2009; Schuettengruber et al. 2017) and **H3K9me3** is found in silent heterochromatin regions (Peters et al. 2001) (Fig. 9B). Thanks to the combinatorial action of histone marks other chromatin states have been predicted. For instance, poised bivalent enhancers are those containing both H3K4me1 and H3K27me3 (Bernstein et al. 2006) and latent enhancers are those not labeled with any type of mark,

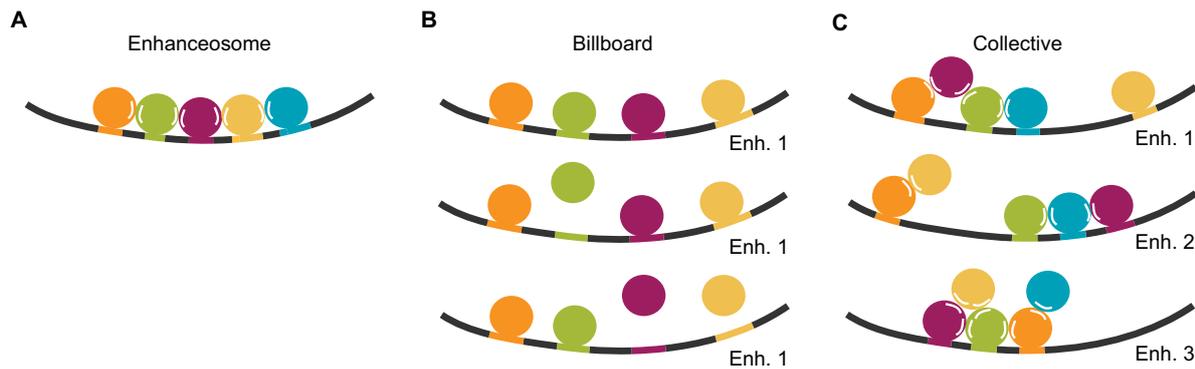
which requires them to be activated upon stimulation through signalling pathways (Ostuni et al. 2013).



**Figure 9 - Enhancer and promoter states.** (A) Chromatin marks in active enhancers and promoters. (B) Chromatin marks in silent enhancers and promoters.

Although histone modifications are one of the best predictors for enhancer activity, they present two major weak points. Usually, there is a correlation between histone marks and states, however there is no mark or combination that perfectly matches with any one state. One clear example can be found in *Drosophila* embryonic mesodermal enhancers, where 40% lack H3K27ac yet they are active (Bonn et al. 2012). Moreover, there is no evidence that such marks are sufficient nor necessary for transcription. One recent study in *Drosophila* has demonstrated that correlation does not imply causation, and that indeed, it is not the mark (H3K4me1) which is required for transcription but the histone methyltransferase, in charge of that mark (Dorigi et al. 2017; Rickels et al. 2017). Additionally it has been demonstrated that transcription can occur in the absence of histone marks in promoters of regulated genes in *Drosophila* (Pérez-Iluch et al. 2015).

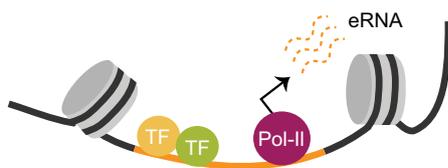
**TFs** bound to enhancer sequences are also a readout of enhancer activity. Often, each enhancer is bound to a specific TF or to a particular combination of such. This makes *in vivo* predictions difficult as they would require countless Chromatin Immunoprecipitation (ChIP) experiments. However, *in silico* prediction of motif composition and motif positioning represents a helpful tool. Some studies indicate that there could be global predictive regulatory rules, represented as **codes of motif composition**, that will ultimately determine enhancer activity (reviewed in Spitz and Furlong 2012). Three main models have been proposed to explain enhancer activity based on TFs motif composition and positioning. In the **enhanceosome** model, all TFs that bind to an enhancer are essential for the cooperative occupancy and activation of the enhancer (Merika and Thanos 2001) (Fig. 10A). In the **billboard model**, the positioning of TF binding sites at any given enhancer is flexible and subject to loose distance or organizational constraints. Only a subset of sites in the enhancer may be active at any given time (Arnosti and Kulkarni 2005) (Fig. 10B). Finally, in the **TF collective model**, the same set of TFs can, depending on the situation, bind to many distinct enhancers in different manners (Fig. 10C). Hence, the collective binding can occur using diverse motif compositions and flexible motif positioning (Junion et al. 2012).



**Figure 10 - Models for TF binding.** (A) Illustration of the enhanceosome model. This model requires the integrated activity of all TFs. The enhancer contains a fixed motif composition and positioning. (B) Illustration of the billboard model. The enhancer only requires a subset of TFs to be active. Motif composition is fixed, but motif positioning can vary. (C) Illustration of the collective model. The same combination of TFs activates different enhancers in different ways. Motif composition and positioning are variable. (Adapted from Spitz and Furlong, 2012).

Nonetheless, even if enhancer activity based on TF binding is an accepted feature, it also presents some weak points. Binding events are not necessarily correlated with activity and there is growing evidence that they might be non-functional and simply reflect chromatin accessibility (Li et al. 2011; John et al. 2011).

Together with the establishment of some universal rules for enhancer activation, other properties have been studied. Enhancers possess some inherent promoter capacity and can recruit Pol-II and TFs (Koch et al. 2011) leading to the **transcription of enhancer-RNA (eRNA)** (Tuan et al. 1992; De Santa et al. 2010; Kim et al. 2010; Lam et al. 2013) (Fig. 11). In a recent study, it has been demonstrated that the degree of enhancer or promoter activity is reflected by the level and directionality of eRNA transcription in the fly (Mikhaylichenko et al. 2018).



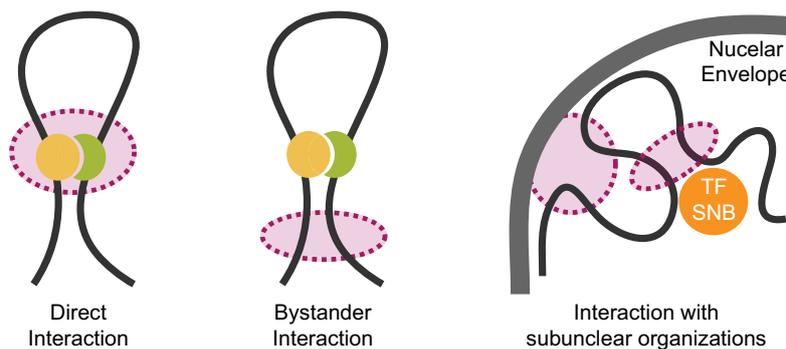
**Figure 11 - Enhancer transcription.** Illustration of transcription at the enhancer locus. TFs and Pol-II bind the enhancer and promote eRNA transcription.

### Shortening long distances: the chromatin architecture

In the last years, the flat perspective of the genome has been left behind and the **three-dimensional genome** has been incorporated as a key component influencing gene transcription. Actually, the chromatin architecture has been demonstrated to be essential for development and the response to stimuli in many eukaryotes, including yeast, worms, plants, flies, and mammals (reviewed in Rowley and Corces, 2016).

Modulation of transcription occurs, in part, through **spatial proximity** of regulatory elements and gene promoters. Enhancers are widely distributed across the whole genome, sometimes located proximal to their target promoters and sometimes located in remote regions. Despite their position, they regulate gene expression taking advantage of the chromatin architecture: enhancers can target their promoters through **chromatin loops**, which shorten long distances (Dekker et al. 2013; Rowley and Corces 2016; Schwartz and Cavalli 2017; Cubeñas-Potts et al. 2017). Chromatin loops have been proposed to assemble an **active like chromatin hub**, providing a more supportive environment for transcription, compared to the one created by TFs bound directly to their promoter alone. Indeed, many enhancer-promoter combinations usually share binding sites for common TFs, potentially leading to eRNA transcription (Reviewed in Sexton and Cavalli, 2015).

Close spatial proximity can be described as the result of direct and specific contacts between two loci, mediated by protein complexes binding these. Alternatively they can be the result of indirect co-localization of loci pairs to the same subnuclear body (SNB), such as the nuclear lamina, nucleolus or transcription factory (Reviewed in Dekker et al., 2013) (Fig. 12).



**Figure 12 - Models for close spatial proximity.** Illustration of a direct interaction, a bystander interaction, and interactions mediated by subnuclear organizations, such as transcription factors or the nuclear envelope.

Moreover, the genome architecture does not only play a role in connecting promoters and enhancers, but also in integrating the action of multiple enhancers, to modulate gene expression. The **spatiotemporal activity** of genes, for instance, usually requires the combination and interaction of multiple enhancer elements in which the genomic architecture plays a pivotal role (Reviewed in Spitz and Furlong 2012). One example is the case of the HoxD cluster, which is regulated by many regulatory elements that form a chromatin archipelago, in which all enhancers work as a single unit (Montavon et al. 2011).

# TOWARDS THE REGULATORY GENOME OF REGENERATION

---

## STATE OF THE ART

Deciphering when and how gene expression patterns are reset is probably the current main challenge of regenerative biology. Thus, a complete understanding of the process requires insight into how early signals at the onset of regeneration are integrated into the genome. There is increasing evidence that the regenerative biological outcome is dictated by how conserved genes and Gene Regulatory Networks (GRNs) are controlled. Regulatory elements, such as regeneration enhancer elements, can perform such functions. Since the appearance of genome-wide techniques, many transcriptomic studies have identified differentially expressed genes in animal models of regeneration. Their combination with forward and reverse genetic analysis, has enabled the identification of GRNs (Reviewed in Chen and Poss 2016). However, few have investigated how regulation of gene expression is achieved.

### The role of chromatin modifying factors in regeneration

Historically, two main epigenetic regulatory groups have been studied in development and, as a consequence, also in regeneration: the **Trithorax group (TrxG)** and the **Polycomb group (PcG)**. Briefly, Polycomb response elements (PREs) and Trithorax response elements (TREs) target PcG and TrxG complexes to chromatin, thus driving the epigenetic inheritance of silent or active chromatin states, respectively, throughout development (Reviewed in Schuettengruber et al. 2017). In regeneration, chromatin modifying factors belonging to the TrxG and the PcG have been proven to play role. They shift the balance between gene expression activation and silencing towards an **enhanced transcriptional state**.

Studies in mouse skin epithelium have demonstrated that the depletion of epigenetic silencing mediated by PcG proteins helps to mediate upregulation of repair genes, after physical injury. Besides, upregulation of H3K27 demethylases of the TrxG (Utx histone demethylase (UTX) and JmjC domain-containing protein 3 (JMJD3)) are required in the blastema area to promote gene expression (Shaw and Martin 2009). Similarly, studies in fly indicated, that transdifferentiation events in regeneration require an enhanced transcription state in which silencing is weakened by the coordinated action of the JNK pathway and PcG/TrxG members (Lee et al. 2005). More recently, it was discovered that the chromatin regulator Taranis (Tara), which belongs to the TrxG, stabilizes compartmental identities during the same transdifferentiation events (Schuster and Smith-Bolton 2015). Finally, a study in zebrafish demonstrated, that during regeneration, TrxG histone demethylases can turn/turn bivalent promoter domains into an active state (Stewart et al. 2009).

## Regulatory elements in regeneration

Even though characterization of regulatory elements has not been in the focus of research for many years now, a number of research groups has gained substantial insights into the field. The study of the WNT damage enhancer in *Drosophila* imaginal discs has shed light in discerning why some individuals lose the ability to regenerate upon maturation (Harris et al. 2016). In this particular case, damage induces the activation of the WNT enhancer, which is essential for activation of the Wnt pathway and to properly regenerate. However, this enhancer is repressed by the action of the PcG proteins upon maturation, and despite damage occurs it is no longer active, leading to defective regeneration.

Besides, reactivation of two embryonic enhancers in the epicardial cell layer of zebrafish and mouse hearts, explains how some epicardial genes, that are transcriptionally activated during embryonic development, can be re-induced after injury (Huang et al. 2012). Similarly, an enhancer triggering *Bone morphogenetic protein 5 (Bmp5)* expression during mouse skeletal development, is also used in bone repair. Interestingly, the same enhancer is sufficient to trigger gene expression in mesenchymal or epithelial cells in multiple tissues, suggesting it might contain an injury-responsive enhancer element (Guenther et al. 2015).

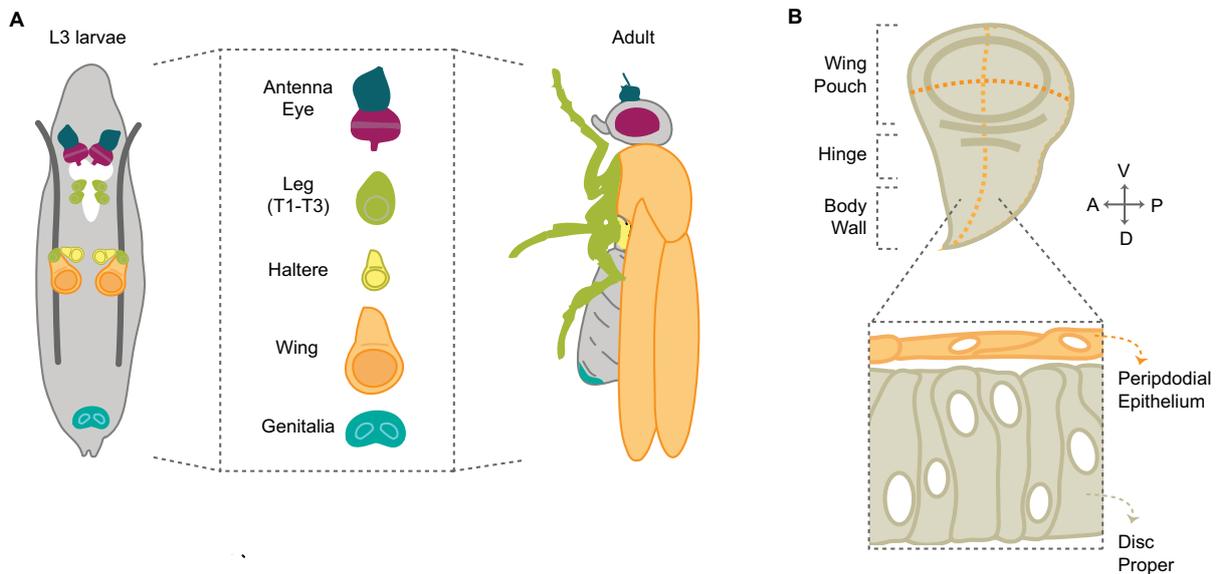
It was just two years ago, that the first genome-wide analysis searching for damage responsive-elements in zebrafish heart regeneration was published (Kang et al. 2016; Goldman et al. 2017). It allowed us to understand, that: a) enhancer elements, whether preferential or specific to regeneration, are widespread and b) several genes with induced expression during regeneration have nearby DNA elements with features expected of regeneration-activated enhancers. Among the enhancers discovered, the *leptin b* one was found not to be used in development and only required in regeneration, thus demonstrating that damage-specific enhancers do exist.

## DROSOPHILA AS A MODEL

The fruitfly *Drosophila melanogaster* is a powerful tool to investigate the regulatory genome of regeneration. *Drosophila* **imaginal discs** present great **regeneration capacity** and have been widely studied. Moreover, the fly has been extensively used to understand the landscape of gene regulation. There is countless genome-wide data describing many developmental stages, tissues, and different *Drosophila* species. This information can be used for further comparative studies giving a more integrative view to the whole regenerative process.

The imaginal discs of *Drosophila* are sacs of epithelial cells present in the larva that give rise to adult structures, such as wings and legs. Imaginal discs are specified early in embryogenesis at different locations in the embryo. They grow through development and, after metamorphosis, they give rise to the adult structures (Cohen et al. 1993) (Fig. 13A). Imaginal discs are composed of two epithelial sheets, the disc proper and the peripodial epithelium,

with their apical surfaces directed toward each other (Fristrom and Fristrom 1993) (Fig. 13B). The disc proper is composed of columnar cells and generates most of the adult structures. The peripodial epithelium is composed of flat squamous cells and is continuous with the disc proper epithelium at its edges. Moreover, some of these primordia structures, such as those of wings and legs, are subdivided into anterior-posterior and dorsal-ventral compartments (Garcia-Bellido et al. 1973).



**Figure 13 - *Drosophila* imaginal discs.** (A) Illustration depicting imaginal discs in the larvae and the structures they give rise to in the adult. (B) Illustration showing the compartmentalization of the wing disc and its epithelial composition.

## Regeneration of imaginal discs

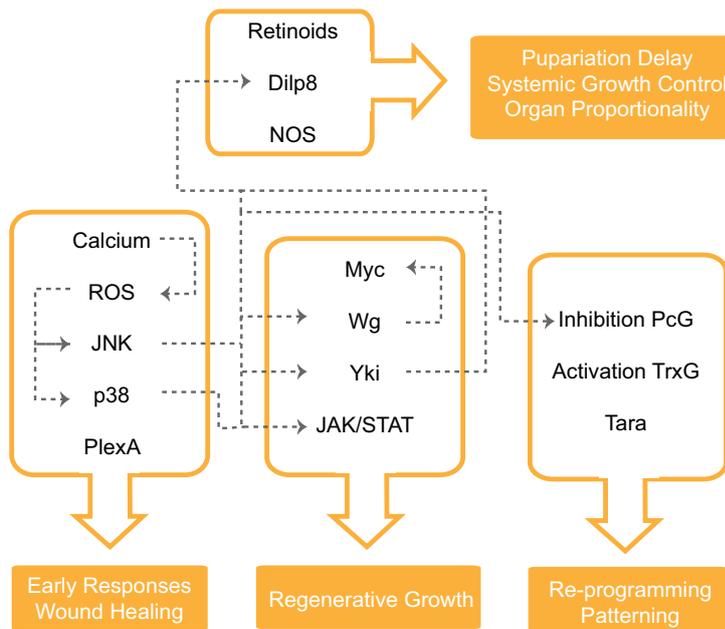
In pioneering regenerative studies beginning in the 1940s by Hadorn and colleagues, imaginal discs were fragmented into pieces, implanted, and cultured in adult female abdomens leading to regeneration (Hadorn 1963; Schubiger and Hadorn 1968). With the appearance of more sophisticated tools, such as x-ray irradiation (Haynie and Bryant, 1977) or genetic ablation systems (Smith-Bolton et al. 2009; Bergantiños et al. 2010), the molecular mechanisms behind imaginal disc regeneration started to be elucidated.

Right after damage occurs, ROS and calcium waves are produced and act as chemoattractants for macrophages (Santabárbara-Ruiz et al. 2015; Fogarty et al. 2016; Moreira et al. 2010; Razzell et al. 2013; Narciso et al. 2015; Restrepo et al. 2015). Moreover, ROS activate the JNK and p38 pathways which are required for **wound healing** and cell proliferation in imaginal disc blastemas (Santabárbara-Ruiz et al. 2015; Bosch et al. 2005; Bosch et al. 2008; Mattila et al. 2005; Lee et al. 2005). A role for Plexin A (PlexA) is also required for the proper wound healing (Yoo et al. 2016) (Fig. 14).

The pathways that drive **regenerative growth** following tissue damage seem to be the same ones as those that regulate growth during normal development. Damage to discs results in the upregulation of *wingless (wg)*, which functions by increasing Myc activity for regenerative growth (Smith-Bolton et al. 2009; Gibson and Schubiger 1999; McClure et al. 2008; Herranz et al. 2008). The activity of *wg* is indeed regulated by the before mentioned WNT damage responsive enhancer (Harris et al. 2016). The Jak-STAT pathway is also required for regenerative growth and activated via the JNK and p38 pathways (Santabárbara-Ruiz et al. 2015; Katsuyama et al. 2009; Pastor-Pareja et al. 2008; Verghese et al. 2016; La Fortezza et al. 2016). Moreover, as a response to changes in cell tension, following tissue damage, the Hippo pathway is inhibited by the Ajuba LIM protein (Jub), resulting in an increased Yorkie (Yki) activity, which is required for growth as well (Grusche et al. 2011; Sun and Irvine 2011; Meserve and Duronio 2015; Repiso et al. 2013) (Fig. 14).

The same genes that are usually used for developmental **patterning**, such as *decapentaplegic (dpp)* and *wg*, are expressed in non-physiological patterns during regeneration, which are not being normalised until regenerative growth is achieved (Smith-Bolton et al. 2009). After injury, respecification and intercalary growth are activated. For instance, cells from the hinge contribute to pouch development and vein and intervein fate changes through the process (Smith-Bolton et al. 2009; Herrera et al. 2013; Repiso et al. 2013; Verghese et al. 2016). Discs are composed by compartments that, even if lost upon damage, are rapidly re-established (Smith-Bolton et al. 2009). Cells near compartmental boundaries, however, are capable of changing their fate, adopting new compartmental identities (Herrera and Morata 2014). These changes are facilitated by PcG silencing combined with the activity of the chromatin regulator *tara*, which helps to stabilize the identities (Lee et al. 2005; Schuster and Smith-Bolton 2015) (Fig. 14).

Finally, regeneration also has **systemic effects**. For example, pupariation is delayed, indicating a strategy to elongate developmental stages, during which the fly is still capable to regenerate, before losing this ability upon maturation (Hussey et al. 1927; Simpson et al. 1980). Retinoids mediate such effects even if the mechanisms are not well understood (Halme et al. 2010). Additionally, *Drosophila* Insulin-like peptide 8 (Dilp8) is released from the discs by the direct mediation of JNK, *yki*, and, indirectly, by the chromatin modifying enzyme *trithorax (trx)* (Colombani et al. 2012; Katsuyama et al. 2015; Boone et al. 2016; Skinner et al. 2015). Dilp8 mediates the release of Nitric Oxide Synthase (NOS) through its binding to Leucine-rich repeat-containing G protein-coupled receptor 3 (Lgr3) in the prothoracic gland which also induces a pupariation delay (Jaszczak et al. 2015 and 2016) (Fig. 14).



**Figure 14 - Overview of regeneration in wing imaginal discs.** Scheme showing signals signals required for regenerative processes and their cross-talk.

Although it has been proposed that the same pathways act in developmental and regenerative growth, damage-induced signals differ from developmental ones in the mechanisms by which they are recruited. Stress signals, like ROS, can activate the JNK and p38 pathways and the coordinated action of both can trigger Jak-STAT activation, which is not likely to occur in development (Santabárbara-Ruiz et al. 2015). Another example is the aforementioned activation of wg, which, in regeneration, is ultimately triggered by the WNT damage induced enhancer (Harris et al. 2016). However, whether signal integration into the genome differs between regenerative and developmental processes, and which genes are controlled in the respective processes remains to be elucidated.



# OBJECTIVES





The regulatory genome governing regeneration has started to be elucidated as an essential element to achieve successful regeneration. Hence, the **main objective** of this thesis is to characterize the regulatory landscape of *Drosophila* wing imaginal disc regeneration. In this concern, we propose three **specific objectives**:

1. Describe the gene expression profiles throughout the recovery process after cell death induction.
2. Unravel the map of regulatory elements that respond to cell death induced regeneration throughout the recovery process.
3. Define the conserved traits of regeneration across metazoans.



# MATERIALS AND METHODS





# Materials



## Drosophila Strains

Fly strains used in this work with its resource information are depicted in Table 1. (See Annex I for a detailed list of the genotypes used).

Strain	Resource	Code
w <sup>1118</sup>	Bloomington Stock Center	5905
UAS-rpr	Wing et al. 1999	NA
LexO-rpr	Santabábara-Ruiz et al. 2015	NA
salm-Gal4	Barrio and de Celis 2004	NA
sal <sup>E/PV</sup> -LHG	Santabábara-Ruiz et al. 2015	NA
tub-Gal80 <sup>TS</sup>	McGuire et al. 2003	NA
UAS-mCD8GFP	Bloomington Stock Center	32186
GMR21F09-Gal4	Bloomington Stock Center	46164
GMR35A10-Gal4	Bloomington Stock Center	49897
GMR17D09-Gal4	Bloomington Stock Center	48766
GMR25D02-Gal4	Bloomington Stock Center	45848
VT39456-Gal4	Viena Drosophila RNAi Center (VDRC)	VT39456
GMR26G03-Gal4	Bloomington Stock Center	49169
GMR32B11-Gal4	Bloomington Stock Center	47539
GMR85E02-Gal4	Bloomington Stock Center	46801
GMR24G07-Gal4	Bloomington Stock Center	49095
GMR42G10-Gal4	Bloomington Stock Center	50168
GMR69F06-Gal4	Bloomington Stock Center	39497
GMR41E03-Gal4	Bloomington Stock Center	50126
GMR36C06-Gal4	Bloomington Stock Center	49931
GMR47D05-Gal4	Bloomington Stock Center	47605
GMR88H01-Gal4	Bloomington Stock Center	40529
ci-Gal4	Martin and Morata 2006	NA
UAS-S6K.KQ (DN)	Bloomington Stock Center	6911
UAS-PI3K[D954A] (DN)	Bloomington Stock Center	25918
RNAi-Dif	Viena Drosophila RNAi Center (VDRC)	V100537
RNAi-Stat92-E	Viena Drosophila RNAi Center (VDRC)	V106142-KK
RNAi-lilli	Viena Drosophila RNAi Center (VDRC)	V4386G-GD

**Table 1 - Drosophila strains used.** The genotype and the resource of each strain is indicated. NA (Not Associated)

## Reagents

Primary antibodies, dyes and kits used in this work with its resource information are depicted in Table 2.

Reagent	Type	Resource	Used for/ in
$\alpha$ -dSRF	Antibody	Acive-Motif	Immunohistochemistry
$\alpha$ -pH3	Antibody	Abcam	Immunohistochemistry
$\alpha$ -AKT	Antibody	Santa Cruz	Western
$\alpha$ -pAKT	Antibody	Cell Sigalling	Western and Immunohistochemistry
$\alpha$ -S6K	Antibody	Santa Cruz	Western
$\alpha$ -PTEN	Antibody	Cell Sigalling	Western
$\alpha$ -H3K4me1	Antibody	Diagenode	ChIP
$\alpha$ -H3K27ac	Antibody	Abcam	ChIP
$\alpha$ -PolII-8WG16	Antibody	Abcam	ChIP
$\alpha$ -H3K27me3	Antibody	Upstate-Millipore	ChIP
$\alpha$ -PolII P-ser5	Antibody	Abcam	ChIP
dUTP ChromaTide BODIPY	Dye	Life Technologies	TUNEL
Phalloidin-Rhodamine	Dye	Life Technologies	Actin marking
TO-PRO-3	Dye	Life Technologies	Fixed tissue imaging
NucRed	Dye	Life Technologies	Live tissue imaging
ZR RNAmicroprep	Kit	Zymo Reaserch	RNA-seq
ZR RNA clean and concentrator	Kit	Zymo Research	RNA-seq
DNA-MinElute	Kit	Qiagen	ATAC-seq and 3C
PCR cleanUP	Kit	Qiagen	ATAC-seq
Nextera Transposition Mix	Kit	Illumina	ATAC-seq

**Table 2 - Reagents.** The reagent and its source is indicated. Also it is indicated the type of experiment they were used for.

## Primers

A detailed list of the primers used for qPCR, ChIP-qPCR and 3C-qPCR is depicted in Table 3 and 4 respectively.

Region	Experiment	Forward	Reverse
<i>akt</i>	qPCR	GCCGCTTTTCCCAGTTC	ATCGATGCGAGACTTGTGG
<i>s6k</i>	qPCR	ATTCCGGTTCTAGGTCAACATC	AACTGTTGACCTGGAGCTG
<i>pten</i>	qPCR	CTCAAAAACGATTGAAATCTTG	GCGAGTCCGATGGAACAG
<i>cbt</i>	qPCR	CACTAAGGGAAACAAGTTGG	TTCTGACTCTTTTGGGCCAC
<i>rpr</i>	qPCR	GGAATCTCCACTGTGACTC	ATACCCGATCAGGCGACTCT
<i>yki</i>	qPCR	AGACCAATGATGGCCAGA	CGTGGCGATATTGGATTCTG
<i>gadd45</i>	qPCR	CACATGCACGAGTACTGCT	GTCGACTAGCTGGTTCTCGG
<i>rps18</i>	qPCR	CCTTCTGCCTGTTGAGGA	TGCACCGAGGAGAGGTC
<i>dia</i>	qPCR	CAAGTGGACGTGTGGGC	AAAGTCTTGATGTCCGCAAAG
<i>sply</i>	qPCR	CTTTCCCGATTCCCGTA	TGACGGGCTTAAGGCAATC
WNT iDRRE	ChIP-qPCR	ACAGAAACCTCGATTGCACTTT	TGCGAATTTGGAGTGATGGGTG
Proximal eDRRE	ChIP-qPCR	TGGTCAGTTGGGCTAGTGGGA	TAGACGAGGTTGGCTATAATCT
Distal eDRRE	ChIP-qPCR	TTTGACATTGGTTCGGGCCCT	CGGGCCTGCAACAGGTAATG
Proximal, 2:1	3C-qPCR	GGCCGGAATGGAGGCACTT	ACGCCTCTGATCTCTGTACCG
Proximal, 3:1	3C-qPCR	GGTGTCCGTGAGAGAGTGTGATG	CTACGCCTCTGATCTCTGTACCG
Distal, 2:1	3C-qPCR	TATACTCTGGCCTTCTGCAT	TGTGTCACGCATACGCAATAT
Distal, 2:5	3C-qPCR	ACTCTGCCTCACCGCATT	GGGATACGTACAAGAATACCATAC
Distal, 2:4	3C-qPCR	ATACTGCGACACACAGTGC	CAAAATGAGTTGGCGGGACT
Distal, 2:3	3C-qPCR	ATACTGCGACACACAGTGC	TGAGACGGAGTGGCGTAAT
Distal, 3:4	3C-qPCR	ATACTGCGACACACAGTGC	CAGGAACAGCTACGGGATT
Distal, 4:5	3C-qPCR	TCCTTGTCGATGCCCTAAA	GTGAGACGGAGTGGCGTAA
<i>ap</i> Enhancer	3C-qPCR	AACATACTCTTCTCGCCCAATCC	ACGTTTCACTTCGAGTACGACGG

**Table 3 - qPCR primers.** The region tested, the technique used and the forward and reverse primer are specified in the table.

## Genome-Wide Data

A detailed list of the genome wide data produced in this thesis and of data acquired from literature for further comparisons is depicted in Table 4.

Data Produced in this thesis					
Technique	Organism	Tissue	Condition	GEO Number	Reference
RNA-seq	<i>D. melanogaster</i>	Wing Disc, 0h ACD	Injured and Uninjured	GSE102841	Vizcaya-Molina et al. 2018
RNA-seq	<i>D. melanogaster</i>	Wing Disc, 15h ACD	Injured and Uninjured	GSE102841	Vizcaya-Molina et al. 2018
RNA-seq	<i>D. melanogaster</i>	Wing Disc, 25h ACD	Injured and Uninjured	GSE102841	Vizcaya-Molina et al. 2018
ATAC-seq	<i>D. melanogaster</i>	L3 Wing Disc	Uninjured	GSE102841	Vizcaya-Molina et al. 2018
ATAC-seq	<i>D. melanogaster</i>	Wing Disc, 0h ACD	Injured and Uninjured	GSE102841	Vizcaya-Molina et al. 2018
ATAC-seq	<i>D. melanogaster</i>	Wing Disc, 15h ACD	Injured and Uninjured	GSE102841	Vizcaya-Molina et al. 2018
ATAC-seq	<i>D. melanogaster</i>	Wing Disc, 25h ACD	Injured and Uninjured	GSE102841	Vizcaya-Molina et al. 2018
H3K4me1 ChIP-seq	<i>D. melanogaster</i>	Wing Disc, 0h ACD	Injured and Uninjured	GSE102841	Vizcaya-Molina et al. 2018
H3K27ac ChIP-seq	<i>D. melanogaster</i>	Wing Disc, 0h ACD	Injured and Uninjured	GSE102841	Vizcaya-Molina et al. 2018
RNApol-II ChIP-seq	<i>D. melanogaster</i>	Wing Disc, 0h ACD	Injured and Uninjured	GSE102841	Vizcaya-Molina et al. 2018

Literature Data Used					
Technique	Organism	Tissue	Condition	GEO Number	Reference
H3 ChIP-seq	<i>D. melanogaster</i>	L3 Wing Disc	Uninjured	GSE56551	Pérez-Lluch et al., 2015
H3K27me3 ChIP-seq	<i>D. melanogaster</i>	L3 Wing Disc	Uninjured	GSE74080	Loubière et al., 2016
FAIRE	<i>D. melanogaster</i>	Embryo 0-4h	Uninjured	GSE38727	McKay and Lieb, 2013
FAIRE	<i>D. melanogaster</i>	Embryo 6-8h	Uninjured	GSE38727	McKay and Lieb, 2013
FAIRE	<i>D. melanogaster</i>	Embryo 12-16h	Uninjured	GSE38727	McKay and Lieb, 2013
FAIRE	<i>D. melanogaster</i>	L3 haltere	Uninjured	GSE38727	McKay and Lieb, 2013
FAIRE	<i>D. melanogaster</i>	L3 CNS	Uninjured	GSE38727	McKay and Lieb, 2013
FAIRE	<i>D. melanogaster</i>	L3 Eye-Antenna	Uninjured	GSE38727	McKay and Lieb, 2013
FAIRE	<i>D. melanogaster</i>	L3 Leg	Uninjured	GSE38727	McKay and Lieb, 2013
FAIRE	<i>D. melanogaster</i>	Pharate Haltere	Uninjured	GSE38727	McKay and Lieb, 2013
FAIRE	<i>D. melanogaster</i>	Pharate CNS	Uninjured	GSE38727	McKay and Lieb, 2013
FAIRE	<i>D. melanogaster</i>	Pharate Eye-Antenna	Uninjured	GSE38727	McKay and Lieb, 2013
FAIRE	<i>D. melanogaster</i>	Pharate Leg	Uninjured	GSE38727	McKay and Lieb, 2013
STARR-seq	<i>D. ananassae</i>	Cell Line	Uninjured	GSE48251	Arnold et al., 2014
STARR-seq	<i>D. melanogaster</i>	Cell Line	Uninjured	GSE48251	Arnold et al., 2014
STARR-seq	<i>D. pseudoobscura</i>	Cell Line	Uninjured	GSE48251	Arnold et al., 2014
STARR-seq	<i>D. yakuba</i>	Cell Line	Uninjured	GSE48251	Arnold et al., 2014
STARR-seq	<i>D. willistoni</i>	Cell Line	Uninjured	GSE48251	Arnold et al., 2014
RNA-seq	<i>Danio Rerio</i>	Heart	Injured and Uninjured	GSE81865	Goldman et al., 2017
RNA-seq	<i>Mus Musculus</i>	Liver	Injured and Uninjured	GSE76926	Sun et al., 2016
Histone Profiling	<i>Danio Rerio</i>	Heart	Injured and Uninjured	GSE81893	Goldman et al., 2017
ATAC-seq	<i>Danio Rerio</i>	24h embryo	Uninjured	GSE61065	Gehrke et al., 2015

**Table 4 - Genome-wide data used.** Detailed list of the genome-wide data used. The table shows the type of experiment from where the data was obtained, as well as the species, the tissues and the condition (injured or uninjured sample) of each sample. The GEO accession number and the reference also shown.

## Softwares

Softwares used and its resource information are depicted in Table 5.

Software	Reference
ImageJ	Schindelin et al. 2012
grape-nf	<a href="https://github.com/guigolab/grape-nf">https://github.com/guigolab/grape-nf</a>
STAR 2.4.0j	Dobin et al. 2013
RSEM	Li and Dewey 2011
d3js	<a href="https://d3js.org/">https://d3js.org/</a>
ggplot2	Wickham 2009
DAVID	Huang et al. 2008, 2009
reviGO	Supek et al. 2011
Cytoscape	Shannon et al. 2003
iREGULON	Janki et al. 2014
KEGGmapper	Kanehisa and Goto 2000; Kanehisa et al. 2016, 2017
FlyFactorSurvey	<a href="http://mccb.umassmed.edu/ffs">http://mccb.umassmed.edu/ffs</a>
CROC	Pignatelli et al. 2009
MACS2	Zhang et al. 2008
BEDOPS v. 2.4.14	Neph et al. 2012
bwtool summary v. 1.0	Pohl and Beato 2014
chip-nf	<a href="https://github.com/guigolab/chip-nf">https://github.com/guigolab/chip-nf</a>
GEM mapper	Marco-Sola et al. 2012
Picard	<a href="http://broadinstitute.github.io/picard/">http://broadinstitute.github.io/picard/</a>
SPP	Kharchenko et al. 2008; Landt et al. 2012
preprocessCore	Bolstad et al. 2003
UCSC - liftOver	Tyner et al. 2017
BEDTools intersectBed v2.17.0	Quinlan and Hall 2010
iCIS target	Herrmann et al. 2012
Ensembl79	Yates et al. 2016

**Table 5 - Softwares used.** List showing all the softwares used and its reference.



## Methods

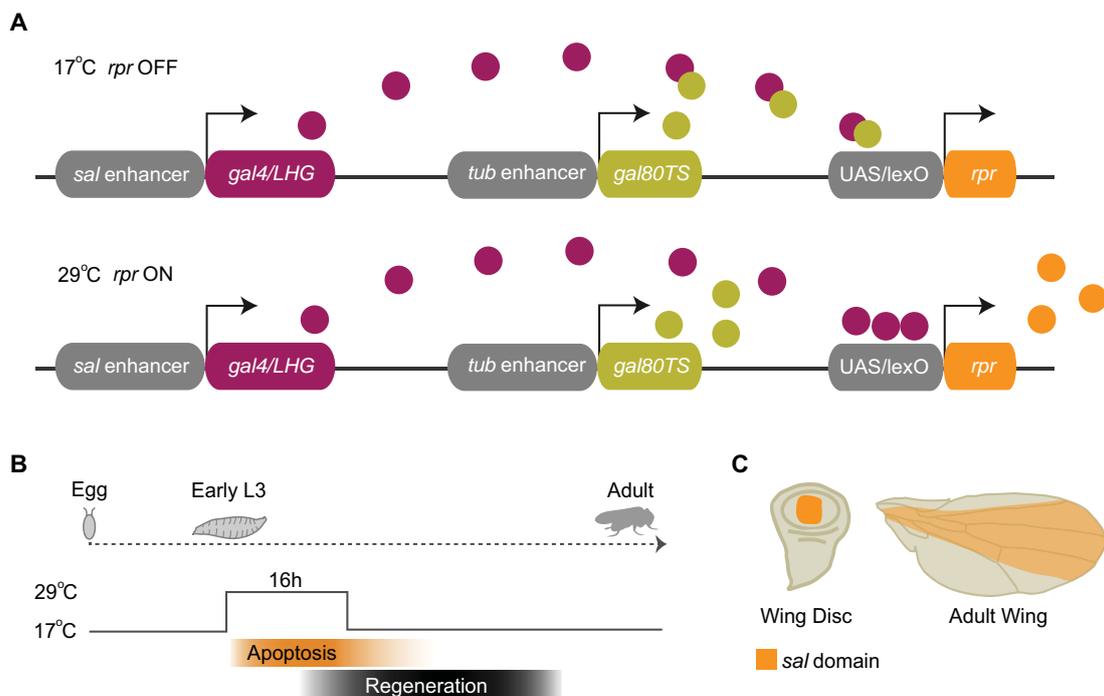


## METHODS TO STUDY REGENERATION

### *In vivo*: CELL DEATH INDUCED REGENERATION

To induce cell death in a particular developmental time and in a specific domain of the wing disc, expression of the pro-apoptotic gene *rpr* was driven using the Gal4/UAS binary system (Brand and Perrimon 1993) in combination with a thermo-sensitive Gal80 (*Gal80<sup>ts</sup>*) that blocks Gal4 protein at 17°C and became inactive at 29°C. Inductions in all the experiments were performed for 16h at 96h after egg laying (AEL) (early Larvae 3, L3) in the *sal*m domain (using the *sal*m-GAL4). Control samples without *rpr* expression were always treated in parallel (Fig. 15).

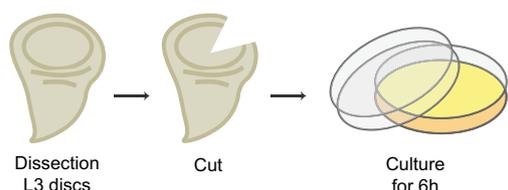
When we needed to combine cell death induction with the expression of other transgenes, we used the ILHG-lexO system following the same logic stated above to drive *rpr* expression (*sal*-LHG, *lexO*-*rpr*).



**Figure 15 - *In vivo* genetic ablation using binary systems.** (A) Outline of the method. At 17°C Gal80 is active and blocks the activity of the Gal4/LHG. At 29°C, Gal80 is inactive and relieves the Gal4-LHG which can bind UAS-*lexO* sequences to promote *rpr* expression. (B) Scheme showing the timing of cell death induction. (C) Drawing showing the *sal*m domain.

## Ex vivo: IMAGINAL DISC CULTURE AFTER PHYSICAL INJURY

We dissected wing discs at L3 stage in Schneider's insect medium and cut them with tungsten needles. Discs were then cultured for 6h at 25°C in culture medium (Schneider's medium supplemented with 2% heat activated foetal calf serum, 2.5% fly extract and 5ug/ml insulin) (Fig. 16).

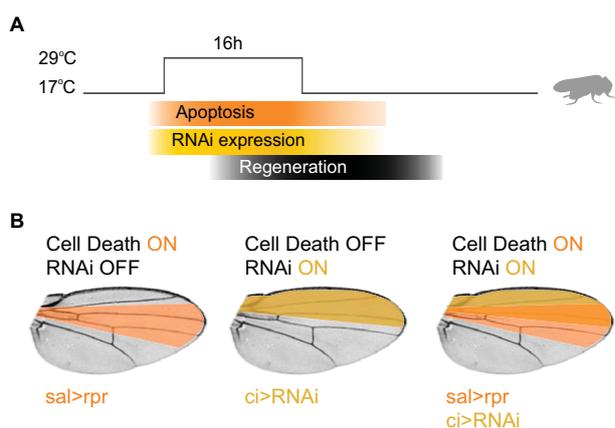


**Figure 16 - Ex vivo culture of imaginal discs.** Schematic showing the design for physical injury experiments.

## TEST FOR REGENERATED ADULT WINGS

To test the capacity to regenerate we analyzed adult wings emerged from flies where cell death was induced using the *LHG/lexO* system and genes were depleted by RNAi using the *UAS/Gal4* system. Cell death was induced in the *salm* domain and the RNAi in the *cubitus interruptus (ci)* domain, and we activated both systems for 16h at 8th day AEL (Fig. 17). Adult flies were fixed in glycerol:ethanol (1:2) for 24h. Wings were mounted on 6:5 lactic acid:ethanol and analyzed and imaged under a microscope.

Wings with defects in patterning (at least one vein or one intervein missing) were considered as aberrant. We also measured the area of each wing. We used as controls wings where only cell death or only the RNAi was expressed (Fig. 17B).



**Figure 17 - Test for regenerated adult wings.** (A) schematic showing the timeline of apoptosis and RNAi induction. (B) Control and experimental wings showing the domain used for cell death and RNAi expression.

# METHODS TO STUDY THE TRANSCRIPTOME

---

## RNA-SEQ

### RNA-seq: library preparation and sequencing

We used 40 wing discs of each genotype (regeneration and control) and time point (0h, 15h and 25h after *rpr* induction). Two biological replicates of each sample were performed. RNA was extracted with ZR RNA microprep and RNA clean and concentrator Kit from (Zymo Research). Five µg of total RNA were used for reverse transcription and cDNAs were subjected to Illumina TruSeq library preparation. All libraries were sequenced on Illumina NextSeq500 according to manufacturer's instruction. Sequencing was done by Sandor Life Sciences Pvt. Ltd. in Hyderabad, India.

### RNA-seq: data processing and analysis

Data was processed using grape-nf (available at <https://github.com/guigolab/grape-nf>). RNA-seq reads were aligned to the fly genome (dm6) using STAR 2.4.0j software (Dobin et al. 2013) with up to 4 mismatches per paired alignment using the FlyBase genome annotation r6.05. Only alignments for reads mapping to ten or fewer loci were reported. Gene and transcripts FPKMs were quantified using RSEM (Li and Dewey 2011). Genes showing at least 1.7 fold change difference in expression levels between control and regeneration at each time point were considered differentially expressed. Plots were performed using d3js (available at <https://d3js.org/>) and ggplot2 (Wickham 2009) and R scripts (available at <https://github.com/abreschi/Rscripts>). *(These analyses were carried out in collaboration with Cecilia Klein, from Roderic Guigó's Lab at the CRG).*

## GENE ENRICHMENT ANALYSIS

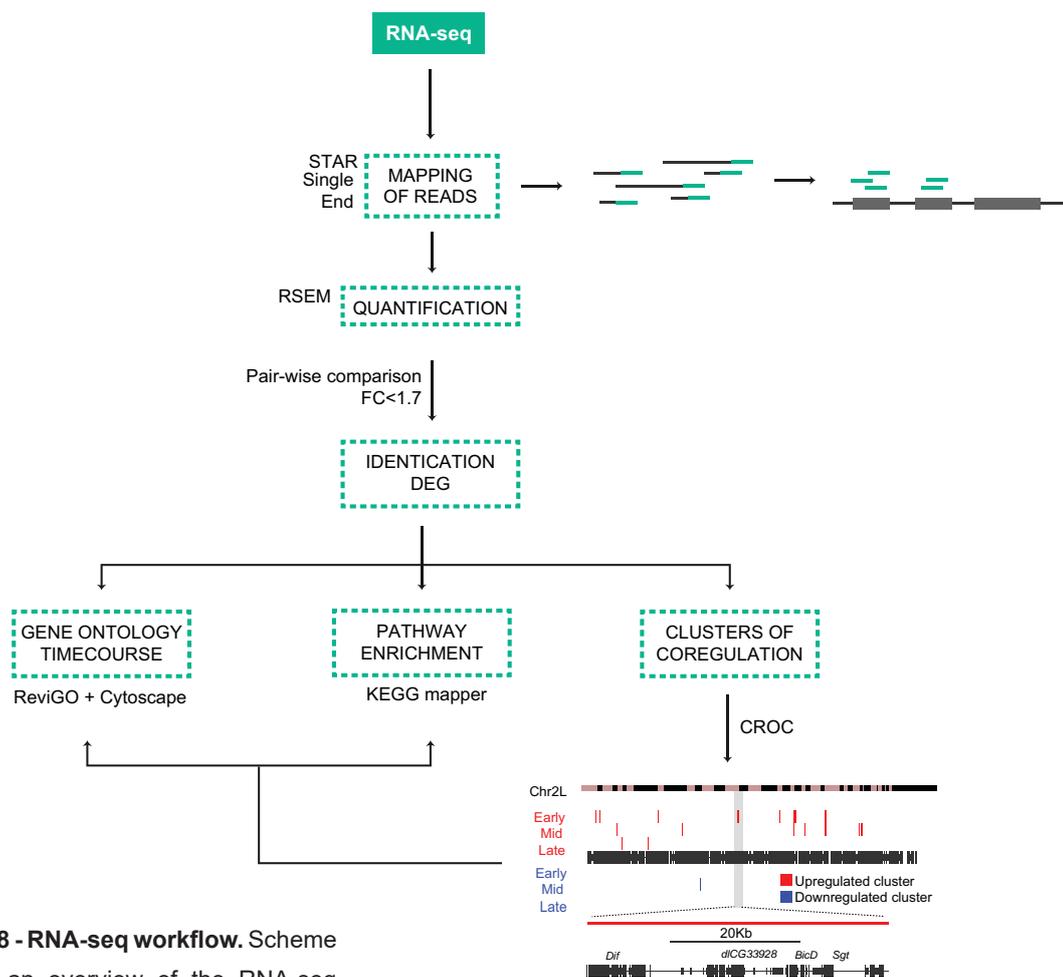
We used DAVID (Huang et al. 2008, 2009) web tool to identify Gene Ontology terms. For time-course analysis of Molecular Function terms, we used revigo (Supek et al. 2011) to compute a network based on semantic terms, term enrichment and gene number for each time. We used Cytoscape (Shannon et al. 2003) for merging and visualizing all time points.

We used KEGGmapper (Kanehisa and Goto 2000; Kanehisa et al. 2016, 2017) to map upregulated genes in fly pathways.

TF annotation was obtained from FlyFactorSurvey (<http://mccb.umassmed.edu/ffs>). We used iREGULON (Jankins et al. 2016) to compute the proportion of genes upregulated at the early stage that can be regulated by the TFs within the same set of genes.

## IDENTIFICATION OF CLUSTERS OF CO-REGULATION

Chromosomal clusters were identified for early, mid and late up and downregulated protein-coding genes using CROC (Pignatelli et al. 2009) with default parameters. To assess co-regulation of genes in the same cluster, we computed the pearson coefficient of correlation for every protein-coding gene pair through time using the R script `gene.pair.correlation.R` (available at <https://github.com/abreschi/Rscripts>) with parameters `--log --pseudocounts 0.01`. The expression profile of genes inside clusters in regeneration samples through the three time points was analyzed as follows: genes for which maximal expression divided by minimal expression is greater than two FPKM were considered variable; the others were classified based on the average expression in the three time points (highly expressed for average expression greater than 30 FPKM; moderately expressed for average expression greater than five FPKM and smaller or equal to 30; lowly expressed for average expression greater than one FPKM and smaller or equal to 5; and silenced for average expression smaller or equal to one FPKM. Cluster hotspots were also identified using CROC on chromosomal clusters. For that, window of 1000000 was defined and no p-value or multiple test correction were required. *(These analyses were carried out in collaboration with Cecilia Klein, from Roderic Guigó's Lab at the CRG).*



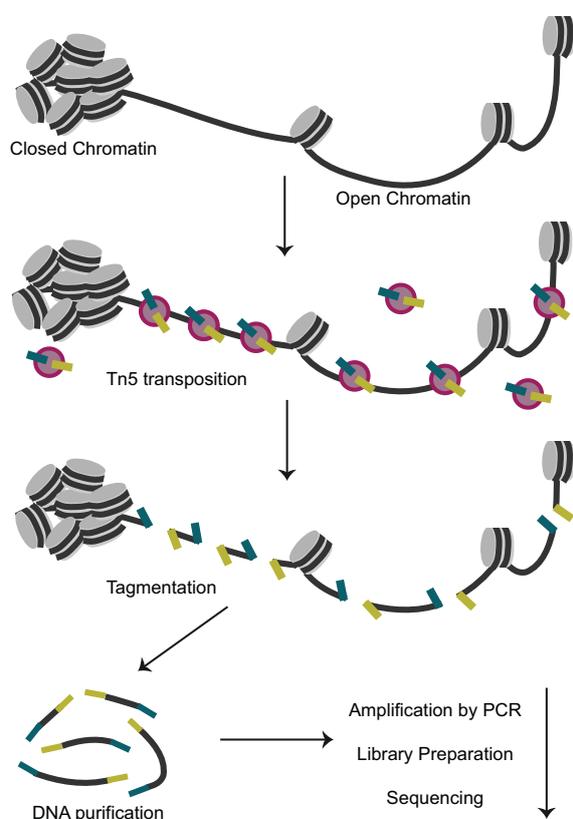
**Figure 18 - RNA-seq workflow.** Scheme showing an overview of the RNA-seq data analysis.

## METHODS TO STUDY THE CHROMATIN LANDSCAPE

### ATAC-SEQ

#### ATAC-seq: the method

ATAC-seq (Assay for Transposase-Accessible Chromatin using sequencing) is a genome-wide technique that allows the study of chromatin accessibility based on the action of the transposase Tn5 (Buenrostro et al. 2013). The Tn5 efficiently recognizes accessible DNA, and cuts and ligates it to specific sequences used as adaptors. The adaptor-ligated DNA fragments are isolated and amplified by PCR and then used for next generation sequencing (Fig. 19).



**Figure 19 - ATAC-seq technique.** Scheme showing an overview of the ATAC-seq technique. The Tn5 transposase recognizes the open chromatin, cuts it and incorporates sequencing adaptors in the tagmentation reaction. Chromatin is then purified and sequenced. (Adapted from Buenrostro et al. 2015)

#### ATAC-seq: library preparation and sequencing.

We used 10 wing discs of each genotype (regeneration and control) and time point (0h, 15h and 25h after *rpr* induction) as well as third instar larva (L3). Two biological replicates of each sample were performed, as previously described (Gehrke et al. 2015; Davie et al. 2015) with some modifications. Briefly, samples were lysed in Lysis Buffer (10 mM Tris-HCl, pH 7.4, 10 mM NaCl, 3 mM MgCl<sub>2</sub>, 0.1% NP40) by gently pipetting. Lysates were centrifuged for 10min at 500g to isolate the nuclei. Nuclei were resuspended and incubated for 30 min at 37°C in transposition reaction mix (Illumina). Right after the transposition reaction, samples were

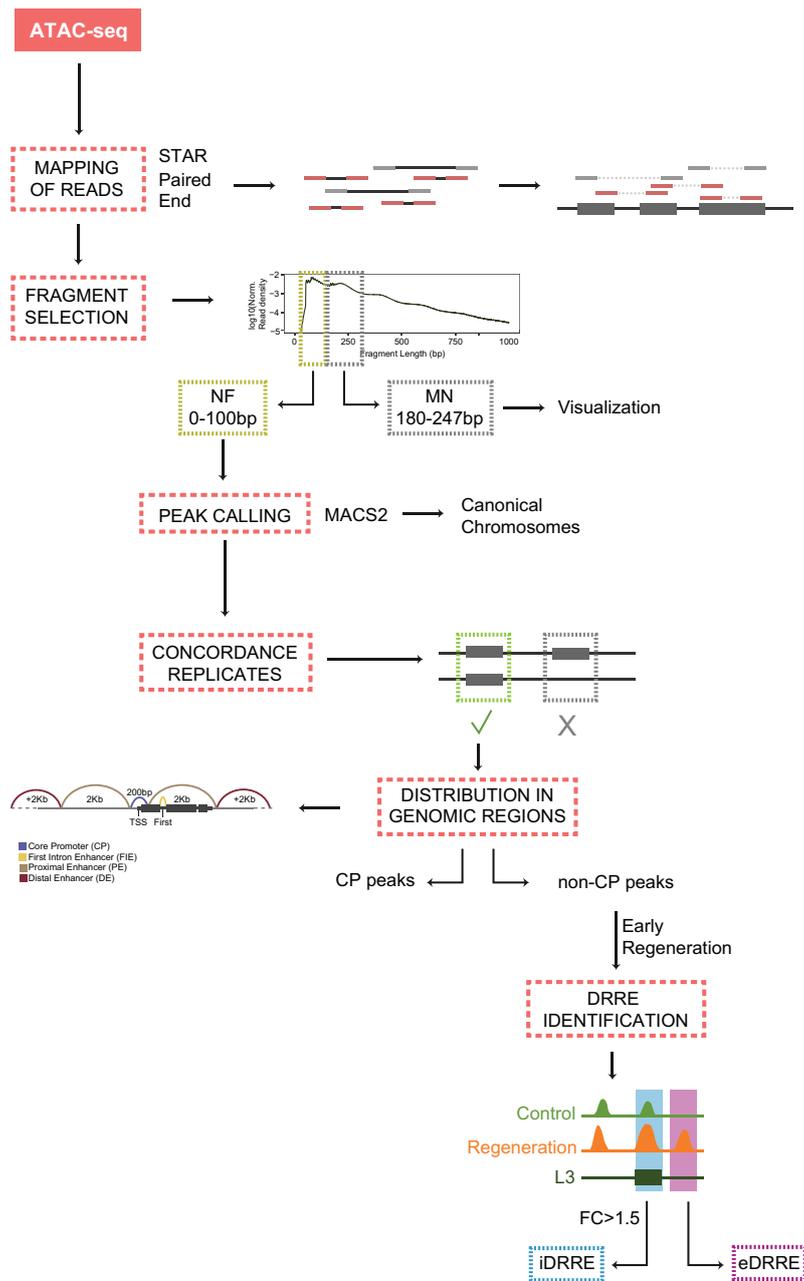
purified using Qiagen MinElute Kit and eluted in Elution Buffer (10mM Tris buffer, pH8).

For library preparation we amplified the transposed DNA fragments by running a conventional PCR (5 min at 72°C, 2.5 min at 95 °C, thermocycling 13 cycles 20 sec at 98°C, 15 sec at 63°C and 1 min at 72°C) with Nextera barcoded primers. Libraries were purified using Qiagen PCR CleanUP Kit and eluted in Elution Buffer. All libraries were sequenced on Illumina HiSeq2500 according to manufacturer's instruction. Sequencing was performed at the Centre Nacional Anàlisi Genòmica- Centre de Regulació Genòmica (CNAG-CRG) sequencing facilities in Barcelona, Spain.

### **ATAC-seq: data processing and analysis**

Reads were continuously mapped to the fly genome (dm6) using STAR 2.4.0j software (Dobin et al. 2013). Only uniquely aligned reads to canonical chromosomes were selected. To generate the nucleosome position data, reads shorter than 100 bps were considered nucleosome-free and reads between 180 and 247 bps were considered to be mononucleosomes (Buenrostro et al. 2015). Peaks were called using paired-end mode of MACS2 software (Zhang et al. 2008) and signal profiles were normalized by the total number of sequenced reads. Concordant peaks (i.e. called in both replicates) of all samples were merged to define a set of consensus regions using BEDOPS v. 2.4.14 (Neph et al. 2012). To identify differentially accessible regions we did pairwise comparison between peaks called in control and regeneration at each time point. We analyzed presence and absence of peaks or peak summits showing at least 1.5 fold change difference in height when called in both conditions. Peak height of each sample was defined using bwtool summary v. 1.0 (Pohl and Beato 2014). *(These analyses were carried out in collaboration with Cecilia Klein, from Roderic Guigó's Lab at the CRG).*

We assigned a unique genomic annotation for each peak by using the following order: core promoter ( $\pm 100$ bp from the transcription start site, TSS); first intron (region between the first and second projected exons, i.e. merged exons of all annotated transcripts of a gene); proximal ( $\pm 2$ kb from TSS); distal (more than  $\pm 2$ kb from TSS). To identify DRRE (Damage Responsive Regulatory Elements) we used more accessible regions at early stage and filtered based on the presence in L3. eDRRE (*emerging* DRRE) were defined as peaks exclusively called in regeneration samples (excluding L3 and control peaks), while iDRRE (*increasing* DRRE) were called in L3, early control and regeneration samples (Fig. 20 for a detailed workflow of all the analysis performed).

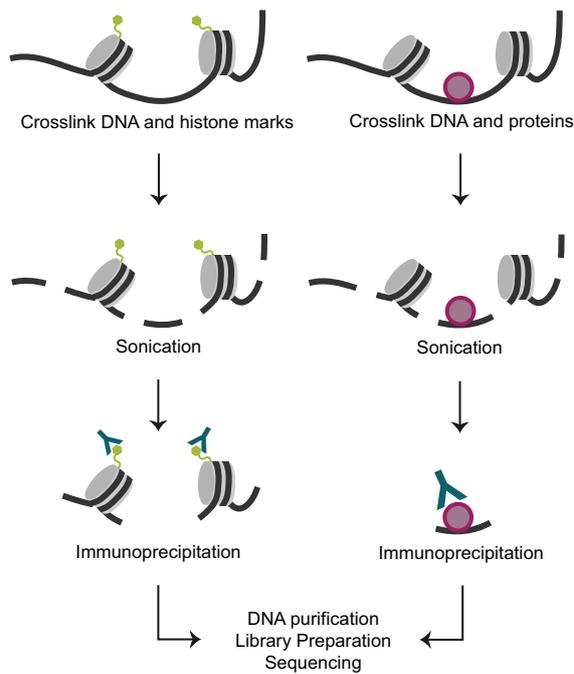


**Figure 20 - ATAC-seq workflow.** Scheme showing an overview of the ATAC-seq data analysis and the identification of DRREs.

## ChIP-seq

### ChIP-seq: the method

ChIP-seq (Chromatin ImmunoPrecipitation followed by sequencing) is a genome-wide technique that allows to recognize protein interactions with DNA as well as histone-tail modifications (Johnson et al. 2007). Chromatin is crosslinked, sonicated and precipitated by an antibody against the protein or the modification in study. The immunoprecipitated chromatin is purified and used for next generation sequencing (Fig. 21).



**Figure 21 - ChIP-seq technique.** Scheme showing an overview of the ChIP-seq technique. Chromatin is crosslinked, sonicated and the feature of interest is recognized by a specific antibody. The complex feature-antibody is precipitated and decrosslinked. Then the chromatin is purified and sequenced.

### ChIP-seq: library preparation and sequencing

We isolated 100 wing discs per sample (early control and regeneration). Discs were fixed, pooled in 700  $\mu$ l of sonication buffer (10 mM Tris-HCl, pH 8.0, 2 mM EDTA and 1 mM EGTA) and processed as described (Pérez-Lluch et al. 2011). Immunoprecipitations were performed in RIPA buffer and using 1  $\mu$ g of the corresponding antibody. Immunocomplexes were recovered by incubation with Invitrogen Protein A magnetic beads for 2 h. The beads were washed three times in RIPA, once in lithium chloride buffer and twice in TE buffer. After, RNAase treatment was done and samples were decrosslinked at 65°C overnight by adding Proteinase K. Samples were purified with Qiagen MinElute Kit and eluted in Gibco water. Library preparation and sequencing using HiSeq 2000 were carried out at CRG Genomic Unit (Barcelona, Spain).

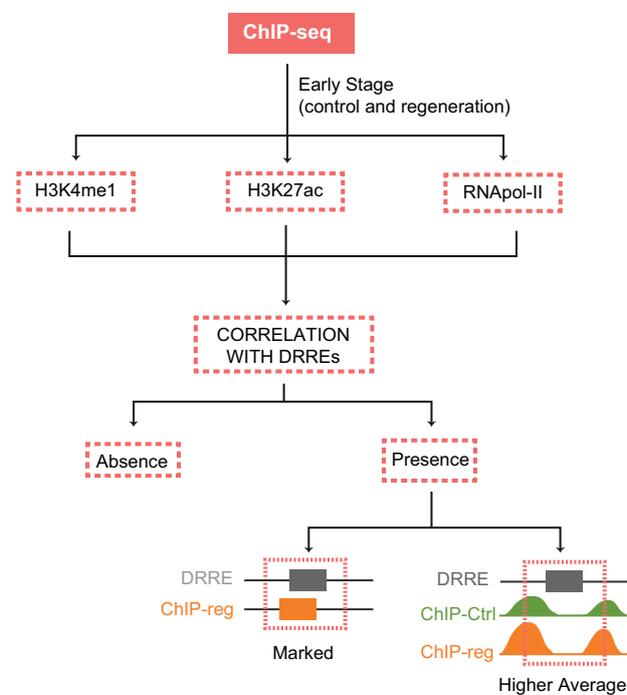
ChIP-qPCR analyses were done following the same protocol. ChIP eluates and input (10%) were assayed by real-time PCR with SYBR Master Mix (Roche).  $\Delta\Delta$ Ct method was used to normalize the data. Both samples are normalized against the input. Average Standard Error of the Mean (SEM) of two biological replicates was computed for each one based on three technical replicates by the  $\Delta\Delta$ Ct method. ChIP enrichment is shown as FC between regeneration and control.

### ChIP-seq: data processing and analysis

Data was processed using chip-nf pipeline (available at <https://github.com/guigolab/chip-nf>). Reads were continuously mapped to the fly genome (dm6) with up to 2 mismatches using GEM mapper (Marco-Sola et al. 2012). Only alignments for reads mapping to ten or fewer loci were reported. Duplicated reads were removed using Picard (<http://broadinstitute.github>).

io/picard/). Fragment length was estimated using SPP (Kharchenko et al. 2008; Landt et al. 2012). Peak calling was performed using MACS2 (Zhang et al. 2008). Signal profiles were quantile normalized using R package preprocessCore (Bolstad et al. 2003). Quality check was based on the signal level of H3K27ac, H3K4me1 and Pol-II at the TSS of modEncode stable and silent genes (Graveley et al. 2011). We computed the coefficient of variation of gene expression for 12 developmental time points and selected 1000 stable genes (lowest values of the coefficient of variation) and 1000 silent genes in this same dataset. (These analyses were carried out in collaboration with Cecilia Klein, from Roderic Guigó's Lab at the CRG).

To characterize chromatin along the predicted DRREs, we intersected both datasets using BEDTools intersectBed v2.17.0 (Quinlan and Hall 2010). We considered as active DRREs the ones showing higher ChIP average signal in regeneration than in control (FC>1.5 in a window of +250bp) and the ones intersected with of ChIP peak in regeneration. (Fig. 22)



**Figure 22 - ChIP-seq workflow.** Scheme showing an overview of the ChIP-seq data analysis and the intersection with DRREs.

## TEST FOR ENHACER ACTIVITY IN REGENERATION

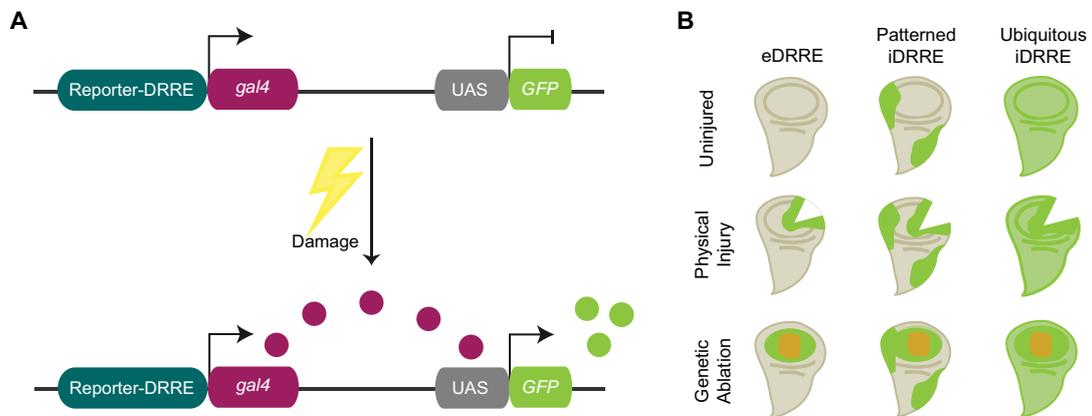
We tested enhancer activity upon damage using reporter lines obtained from the *Janelia-FlyLight* and *VDRC-VT* collections (Table 6).

In physical injury analysis we crossed reporter lines with a *UAS-mCD8GFP*, then we dissected discs, cut them and cultured them for 6 hours. After, we checked for an enhancement or *de novo* GFP expression in cells nearby the wound (Fig. 23). In genetic ablation analysis we combined the *UAS-Gal4* and the *LHG-lexO* system. We used the *UAS-Gal4* to test the enhancer activity in the same way we did in physical injury analysis, and the *LHG-lexO* to drive genetic

ablation. We induced cell death for 16h in the *salm* domain at 96h of development and checked for enhancer activity in early regeneration (Fig. 23).

Reporter Line	DRRE Coordinates	DRRE type	Genomic Location	Associated Gene
GMR21F09-Gal4	chrX:19697549-19697888	iDRRE	First Intron	<i>cdc42</i>
GMR25D02-Gal4	chr3L:6216670-6217317	iDRRE	First Intron	<i>lanA</i>
GMR35A10-Gal4	chr3R:30765485-30765822	iDRRE	Proximal	<i>zfh1</i>
VT39456-Gal4	chr3R:10437127-10437322	iDRRE	Proximal	<i>Cyp12e2</i>
GMR17D09-Gal4	chr2L:7286945-7287225	iDRRE	Distal	NA
GMR26G03-Gal4	chrX:16254669-16254835	iDRRE	Distal	<i>Cyp</i>
GMR32B11-Gal4	chr2R:10680648-10680812	eDRRE	First Intron	<i>stan</i>
GMR24G07-Gal4	chr2L:7266527-7266631	eDRRE	First Intron	<i>wnt4</i>
GMR42G10-Gal4	chr3R:11351859-11352162	eDRRE	First Intron	<i>pros</i>
GMR85E02-Gal4	chr2L:11472301-11472450	eDRRE	Distal	NA
GMR69F06-Gal4	chr2R:8640900-8641040	eDRRE	Distal	NA
GMR41E03-Gal4	chr2R:5741496-5741591	eDRRE	Distal	NA
GMR36C06-Gal4	chr3R:13938246-13938465	Neg Ctrl	First Intron	<i>AI</i>
GMR47D05-Gal4	chr2R:14807712-14808131	Neg Ctrl	Proximal	<i>hui</i>
GMR88H01-Gal4	chr3R:13939143-13939310	Neg Ctrl	Distal	NA

**Table 6 - Reporter lines features.** The table shows the name of the reporter line used for each tested DRRE. Also are depicted the coordinates, the type and the genomic location of each DRRE. If the DRRE is associated to the TSS of any gene, the gene name is also shown.



**Figure 23 - Validation of DRREs activity.** (A) Scheme of the genetic strategy used to validate enhancer activity. (B) Drawing showing the expected GFP expression driven by the different DRREs types upon physical injury and genetic ablation.

## CHROMOSOME CONFORMATION CAPTURE (3C)

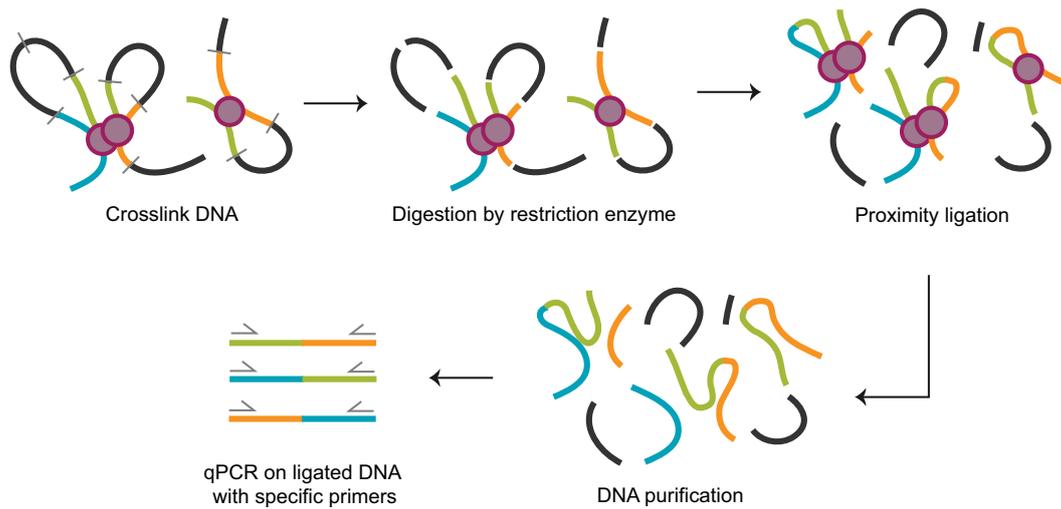
### 3C: the method

Chromosome Conformation Capture based techniques are used to study the spatial organization of chromatin. They rely on the ability of the chromatin to create loops that bring in contact regions that are far away in the linear genome such as, for instance, promoter-enhancer interactions (Reviewed in de Wit and de Laat, 2012). In 3C-qPCR the chromatin is crosslinked and digested by a restriction enzyme and then ligated again. Primers for ligated regions are design and quantified by qPCR. As the chromatin is crosslinked, the frequency of ligate two regions that are in close contact than two random regions is much higher, therefore, the amount of qPCR product is also higher (Fig. 24).

### 3C-qPCR: sample preparation and analysis

We developed a 3C protocol for wing disc following previously described 3C procedures in *Drosophila* (Li 2016). We used 300 wing discs for each condition (0h after *rpr* induction control and early regeneration) and we did 2 replicates of each experiment. Rounds of 50 larvae (100 discs each) were turned and fixed in PBS1x-37%formaldehyde for 15 min at 25°C. Fixation was quenched with glycine (0.125M) and cold down in ice for 5 min. Larvae was resuspended in PBS1x and discs dissected. All discs were pooled together and spun down and lysated in Lysis Buffer (10 mM Tris-HCl, pH 7.4, 10 mM NaCl, 3 mM MgCl<sub>2</sub>, 0.1% NP40) by gently pipetting for 10 min. Nuclei were pelleted by centrifugation at 600g 4°C for 10 min and washed 5 min in 1XRestriction Enzyme Buffer (Pst1). Nuclei were centrifuged again, resuspended and incubated for 1h at 37°C in Nuclei Lysis Buffer (1X Restriction Enzyme Buffer, SDS). TritonX-100 was added (final concentration 2%) for 1h more. Pst1 (Promega) was added to the sample and incubated overnight 37°C; after, SDS was added for 1h. Following the addition of ligation Buffer and TritonX-100, samples were incubated for 1h at 37°C. Temperature was lowered by incubation on ice for 5 min and then ATP and T4 ligase (Roche) were added. Ligation reaction lasted for 4h at 16°C and for 1h at 25°C. After ligation, samples were decrosskinked at 65°C overnight by adding Proteinase K. Right after, RNAase treatment was done. Samples were purified with Qiagen MinElute Kit and eluted in Gibco water.

3C eluates were assayed by qPCR with SYBR Master Mix (Roche).  $\Delta\Delta C_t$  method was used to normalize the data. 3C interaction enrichment is shown as FC between regeneration and control. Both samples were normalized against a known interaction in *Drosophila* (Bieli et al. 2015). Average SEM of two biological replicates was computed for each one based on three technical replicates by the  $\Delta\Delta C_t$  method.



**Figure 24 - 3C-qPCR technique.** Scheme showing an overview of the 3C-qPCR technique. Chromatin is crosslinked, digested by a restriction enzyme and ligated. Then the chromatin is purified and the ligated DNA is quantified by qPCR using pairs of specific primers.

## MOTIF ENRICHMENT ANALYSIS

We used iCIS target (Herrmann et al. 2012) with default parameters to get the enriched motifs in each enhancer type. Only TFs expressed in the RNAseq (>1FPKM) were considered as a hit. Redundant hits were manually removed.

## METHODS TO STUDY PROTEINS

---

### Immunostaining of imaginal discs

Discs were dissected in Schneider's medium and fixed at room temperature for 30' in PBS1x-4%paraformaldehyde. Then they were washed twice in PBS, twice in PBS1x-0.5% Triton X-100) and 30' in blocking solution (PBS1x- 2%BSA, 0.3% Triton X-100). Samples were incubated overnight with the antibody in its appropriate concentration at 4°C. After, discs were washed twice (PBS1x-0.5% Triton X-100) and incubated with the secondary antibody for 2h at 25°C. Discs were twice washed (PBS1x-0.5% Triton X-100) and mounted in Slowfade (Life Technologies). Nuclei were stained using NucRed (Life Technologies) in *in-vivo* imaging and TO-PRO-3 (Life Technologies) in fixed tissues.

### TUNEL assay

For apoptotic cell detection we used TUNEL assay. After fixation, apoptotic cells were detected using labelled dUTP ChromaTide BODIPY-FL-14 or Alexa Fluor® 647-aha-dUTP (Life Technologies) and the Terminal deoxynucleotidyl transferase (TdT, Roche). The sample was incubated in for 1h 30' at 37°C. Then, EDTA was added to stop the reaction and discs were washed and mounted as stated for immunostaining.

### Western Blot

We used 15 wing discs for each condition (early control and early regeneration) and we performed 2 replicates of each experiment. Discs were dissected in Schneider's Medium and transferred to Lysis buffer (50mM Tris, 150mM NaCl, 1%Triton X-100, 1mM EDTA) supplemented with protease and phosphatase inhibitors, frozen at -20°C for at least 30' and thawed. Then, we added the  $\beta$ -mercaptoethanol and boiled the sample at 95°C. Samples were run in a SDS-Polyacrylamide gel and transferred to a Polyvinylidene difluoride (PVDF) membrane. The membrane was incubated for one hour in Blocking solution (PBS1x, 2% Tween-20, 5% BSA) and then with the primary antibody (dissolved in Blocking solution) overnight at 4°C. The membrane was then washed twice in (PBS1x- 2% Tween-20) and incubated with the secondary antibody (conjugated to a peroxidase) for 2 hours at room temperature. The membrane was washed twice and then develop with ECL in LICOR-cDigit. We quantified bands using ImageStudio program.

### Imaging

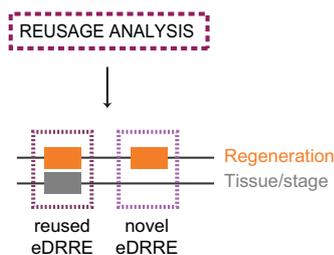
Confocal images were obtained with Leica SPE confocal microscopes. All pictures were processed with Fiji and Adobe Photoshop software.

## METHODS BASED ON GENOME-WIDE COMPARISONS

### REUSAGE ANALYSIS

To check whether DRREs were used in other developmental stages and tissues we obtained Formaldehyde-Assisted Isolation of Regulatory Elements followed by sequencing (FAIRE) data for embryo (2-4h, 6-8h, 16-18h), L3 (Central Nervous System and eye-antenna, leg and haltere discs) and pharate (haltere, leg and wing) stages (GSE38727, McKay and Lieb 2013).

Peak coordinates were converted from dm3 to dm6 using liftOver tool from UCSC genome browser (Tyner et al. 2017). eDRREs overlapping FAIRE open regions in developmental stages different from L3 or in tissues other than the wing imaginal disc were considered reused. eDRREs not overlapping any open region were considered novel-eDRREs. Such overlap was computed using BEDTools intersectBed v2.17.0 (Quinlan and Hall 2010). (These analyses were carried out in collaboration with Cecilia Klein, from Roderic Guigó's Lab at the CRG).



**Figure 25 - Reusage workflow.** Scheme showing an overview of the reusage analysis. eDRREs overlapping FAIRE data are considered reused-eDRRES and no overlapping eDRREs are considered novel-eDRRES.

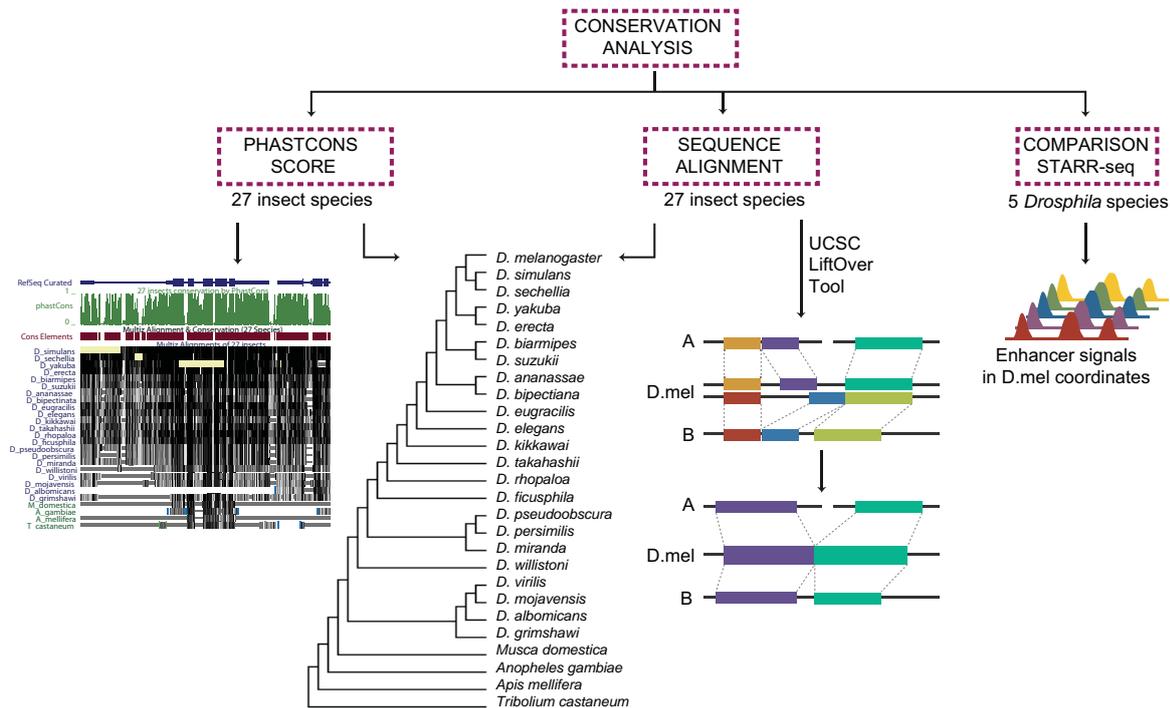
To *in vivo* validate the usage of eDRREs in other tissues (Central Nervous System and eye-antenna, leg and haltere discs) we crossed eDRRE reporter lines with a *UAS-mCD8GFP* and checked for GFP expression.

### DRREs CONSERVATION

To investigate DRRE conservation we used the dm6 27-way multiple alignment (23 *Drosophila* sequences, house fly, *Anopheles* mosquito, honey bee and red flour beetle) and the phast-Cons measurement of evolutionary conservation from the UCSC genome browser (Tyner et al. 2017). Bwtool was used to intersect peaks with the conservation track (Pohl and Beato 2014). (These analyses were carried out in collaboration with Cecilia Klein, from Roderic Guigó's Lab at the CRG).

We also compared DRREs with Self-Transcribing activatie Regulatory Region followed by sequencing (STARR-seq) data, genome-wide enhancer activity profiles for five *Drosophila* species, namely *D. ananassae*, *D. melanogaster*, *D. pseudoobscura*, *D. yakuba*, *D. willistoni* from NCBI GEO database (GSE48251, GSE40739 (Arnold et al. 2014)). (These analyses were

carried out in collaboration with Cecilia Klein, from Roderic Guigó's Lab at the CRG).



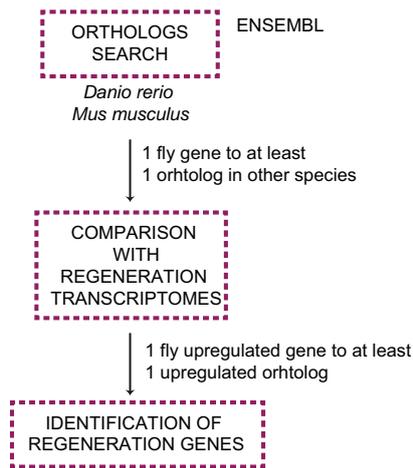
**Figure 26 - DRREs conservation workflow.** Scheme showing an overview of the conservation analysis. Phastons score and sequence alignment in 27 insect species as well as STARR-seq comparison are computed for each DRRE

## CONSERVATION OF GENES IMPLIED IN REGENERATION ACROSS METAZOANS

To identify regenerative genes shared between fly, zebrafish and mouse we first mapped fly gene identifiers to zebrafish and mouse orthologous genes using Ensembl79 (<http://mar2015.archive.ensembl.org>) (Yates et al. 2016). Genes mapping to one or more orthologous genes in zebrafish or mouse were analyzed in a fly-oriented manner.

Besides, we obtained RNA-seq data of regeneration in zebrafish heart and in mouse liver were (NCBI GEO database GSE81865 (Goldman et al. 2017) and GSE76926 (Sun et al. 2016)). We identified higher expressed genes in regeneration for each species (at least 1.5 fold change difference between injured and uninjured expression levels). We selected fly upregulated genes (early and/or mid) mapping to at least one ortholog in mouse and/or zebrafish.

The set of genes upregulated in fly, in zebrafish and in mouse regeneration data were used to identify the regenerative core genes.



**Figure 27 - Regenerative genes discovery workflow.** Scheme showing an overview of the identification of the core regenerative genes.

## CONSERVATION OF THE REGULATORY LOGIC

To identify if the regulatory logic in fly was conserved in higher organisms we first obtained the genome-wide map of histone variant H3.3 occupancy in zebrafish cardiomyocytes undergoing regeneration (same experimental conditions as the RNA-seq zebrafish heart data) and compared to an uninjured sample (NCBI GEO database GSE81893 (Goldman et al. 2017)). (These analyses were carried out in collaboration with Cecilia Klein, from Roderic Guigó's Lab at the CRG).

Concordant peaks (i.e. peaks called in both replicates) were classified as *emerging* (eDRRE: exclusively called in regeneration) or *increasing* (iDRRE: called both in uninjured and injured samples and at least 1.5 fold higher in samples undergoing regeneration). Peaks were classified based on non-overlapping regions of genomic location: core-promoter (0.5kb up/downstream the TSS), first intron enhancer (region between the first and second projected exons, i.e. merged exons of all annotated transcripts of a gene); proximal enhancer ( $\pm 2$ kb from TSS); distal enhancer ( $\pm$ more than 2kb away from TSS) based on Ensembl release 89 of zebrafish (GRCz10) (Yates et al. 2016).

DRREs were compared to ATAC-seq data from 24 hours postfertilization (hpf) zebrafish embryo (NCBI GEO database GSE61065 (Gehrke et al. 2015)). Raw zebrafish ATAC-seq data was mapped to GRCz10 assembly and processed as described for the herein presented fly ATAC-seq NF fraction. Re-usage analysis as for fly was based on the overlap between zebrafish DRRE with open regions in embryo. (These analyses were carried out in collaboration with Cecilia Klein, from Roderic Guigó's Lab at the CRG).

# RESULTS

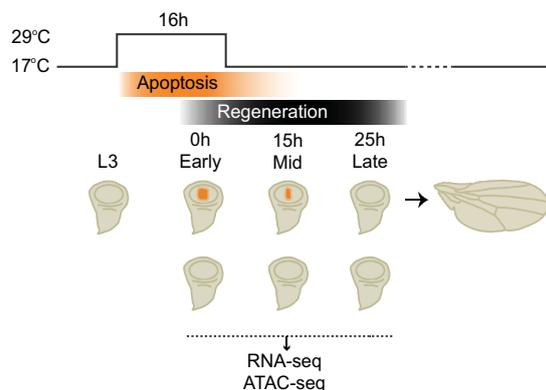




## Experimental design

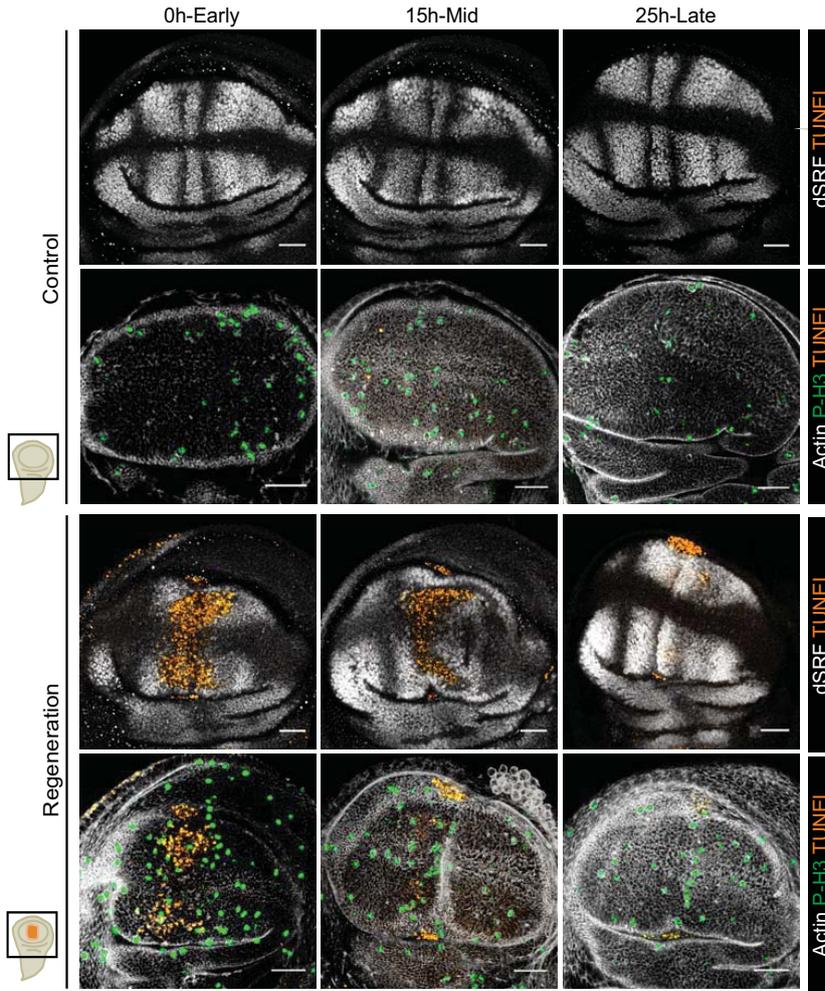
To elucidate the transcriptional regulatory network controlling tissue regeneration in *Drosophila* wing imaginal discs, we characterized the gene expression profiles (by RNA-seq) and the map of accessible regions (by ATAC-seq) associated with the response to damage.

We induced cell death in the *salm* domain for 16h at the L3 stage (96h AEL) and then collected two replicate samples of RNA-seq and ATAC-seq at three different stages. We collected samples of both regenerating and control discs (discs lacking *rpr* expression but kept at the same temperatures) (Fig. 28). (See Annex II for mapping statistics and replicate analysis of RNA-seq and ATAC-seq).



**Figure 28 - Experimental design.** Flies were raised at 17°C until the 8th day after egg-laying (equivalent to third instar larva or L3). After they were moved to 29°C for 16 hours to induce apoptosis triggered by *rpr* specifically in the *salm* domain of the wing pouch (orange region), and then back to 17°C to switch off *rpr* and allow the tissue to regenerate. RNA-seq and ATAC-seq samples were collected at early, mid and late stages.

The three selected time points rank from the initial response to apoptosis up to complete re-patterning. Since cell death is stochastic, and not synchronized, there is no clear separation between the end of apoptosis and the initiation of regeneration. Hence, the first selected time point, which we named **early regeneration stage** (0h: early), corresponds to immediately after genetic ablation, when some of the early signals are known to act (Smith-Bolton et al. 2009, Bergantiños et al. 2010; Repiso et al. 2013; Santabárbara-Ruiz et al. 2015). At this stage, apoptotic cells are extruded from the epithelium, the patterning is completely disrupted and mitotic cells are localized mostly at the edges of the wound (Fig. 29). The second selected time point corresponds to an intermediate step, or 15 hours after apoptosis, which we have named **mid regeneration stage** (15h: mid). At this stage, patterning has not yet been recovered, although living cells have almost closed the wound. Moreover, a localized mitotic zone is also found at the edges of the wound closure (Fig. 29). Finally, the third time point, named **late regeneration stage** (25h: late), corresponds to 25 hours after apoptosis. At this stage, the wound is completely closed and both size and patterning are mostly reconstructed (Fig. 29).



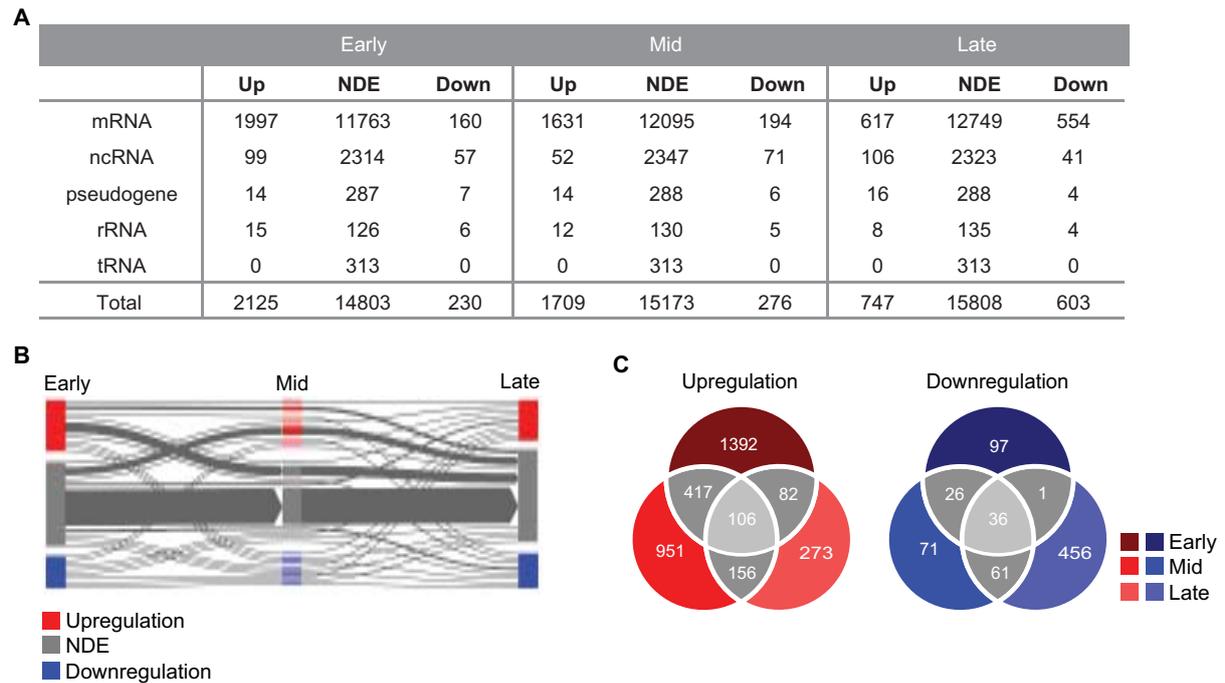
**Figure 29 - Regeneration stages description.** Confocal images of wing discs stained with dSRF (*Drosophila* serum response factor) antibody and actin to visualize the patterning, TUNEL assay to detect cell death and P-H3 antibody to detect mitosis.

**Chapter I:**  
**The transcriptome of regeneration**



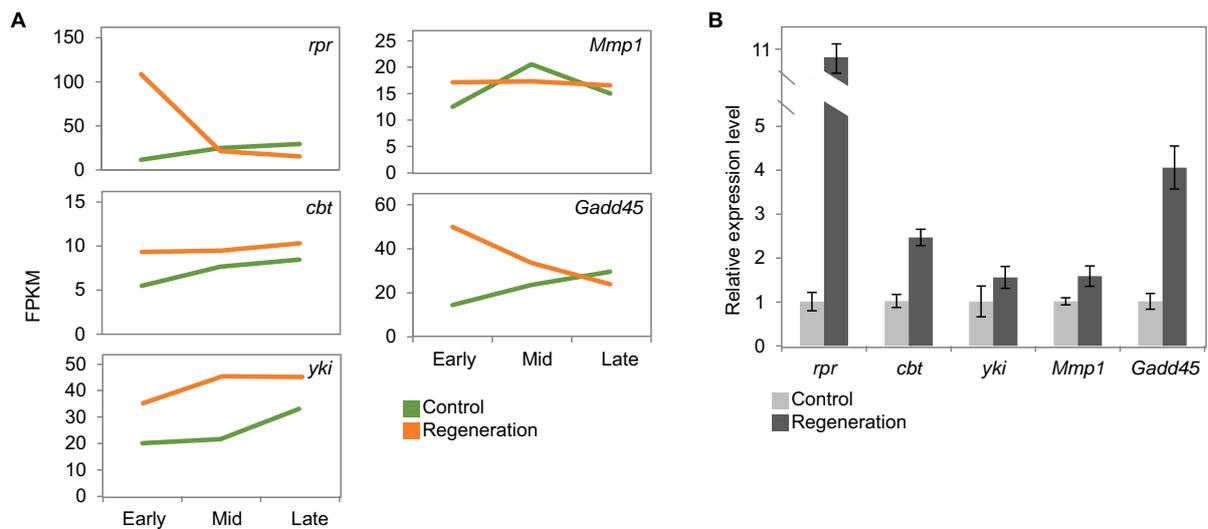
## Regeneration: a burst of active transcription

RNA-seq analysis showed that the great majority of changes were due to an increase of gene expression that mainly occur at the early stage: 92% of differentially expressed (DE) genes were upregulated at that time point. As expected, this number decreased with time, correlating with the recovery of tissue morphology (Fig. 30A,B). The most predominant type of differentially expressed genes were protein coding genes, the great majority of which was indeed stage specific (Fig. 30C). For instance, only a 28% of upregulated genes at the early stage was also upregulated in other stages (Fig. 30B,C).



**Figure 30 - Differentially expressed genes after induction of cell death.** (A) Number of differentially expressed genes at all time points. Gene types are specified. (B) Flux-plot showing RNA-seq dynamics throughout the different time points. Each line represents a set of gene equally behaving through time. The line width denotes the number of genes. (C) Venn diagram showing the intersection of DE genes in the three timepoints.

We realized further validations of the data obtained by doing qPCR at the early stage of protein coding genes. We observed the same behaviour in the genes tested by qPCR than in the RNA-seq data set, which validated the results obtained (Fig. 31). Moreover, among the upregulated genes, we found *unpaired 3 (upd3)*, *Jun- related antigen (Jra)*, *cabut (cbt)* and *p38a MAP kinase (p38a)*, which are known to be required only in a few cells around the wound both after cell death and physical injury (Blanco et al. 2010; Katsuyama et al. 2015; Santabárbara-Ruiz et al. 2015). Also, we found upregulation of genes, such as *activating transcription factor 3 (Atf-3)*, *moladietz (mol)*, *fruitless (fru)*, *LaminC (LamC)* and *pickled eggs (pigs)*, which were identified and validated in a previous transcriptomic study of wing disc regeneration (Khan et al. 2017). These demonstrates that our approach is sensitive enough to identify changes affecting small numbers of cells within the monitored population.



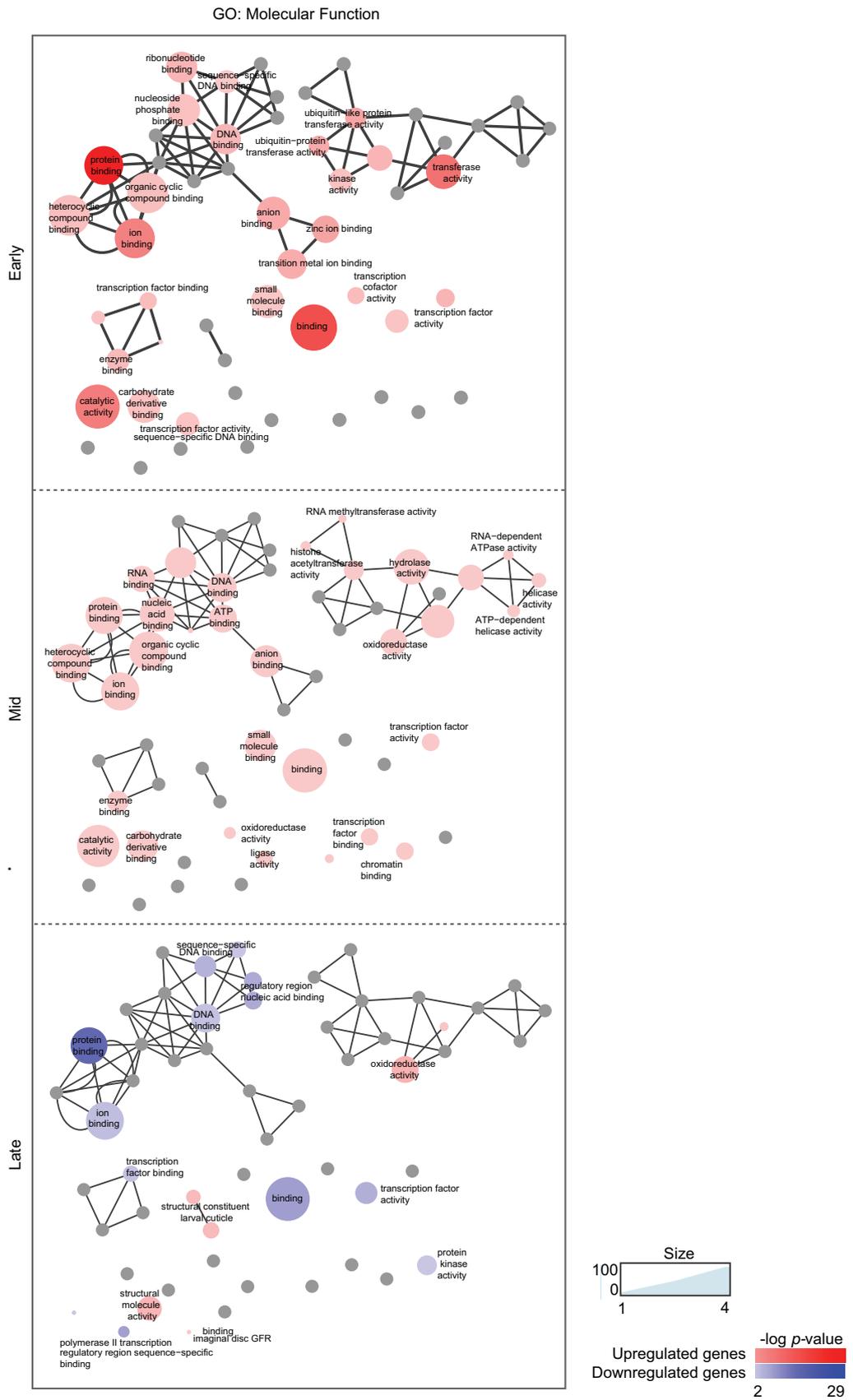
**Figure 31 - Validation of DE genes.** (A) RNA-seq expression profile (FPKM) across time in control and regeneration of the genes selected for qPCR validation. (B) qPCR analysis at the early stage of the selected protein coding genes. This results are presented as fold change enrichment between control and regeneration. Error bars represent standard error of the mean from three biological replicates.

Although we cannot discard that ncRNA or other type of genes could play an important role in the process of regeneration, we have focused our work into protein coding genes, as they account for the great majority of changes in expression.

### Regeneration genes are enriched in transcription related terms

In order to know which kind of molecular role could have DE genes at each timepoint, we did a time-course Gene Ontology (GO) analysis of functional categories (Fig. 32). We found that genes upregulated at the early stage were specifically related to functions as transcription factor activity, kinase activity or DNA binding. However, there were no functional categories enriched in downregulated genes in the early stage neither in the mid. Moreover, we observed that some of the categories enriched in upregulated genes of the early stage were also enriched in the late downregulated ones. Indeed, genes falling in such categories are the same in the early and the late stages, meaning that even if they are upregulated at the beginning, they become silenced towards the end of the process.

Because we detected an enrichment in transcription factor activity related terms, we looked for TFs upregulated throughout the process. We found a set of 195 TFs, a 68% of which is induced at the early stage (Fig. 33). Interestingly, within the set there were several TFs related to signalling pathways involved in the early activation of regeneration.

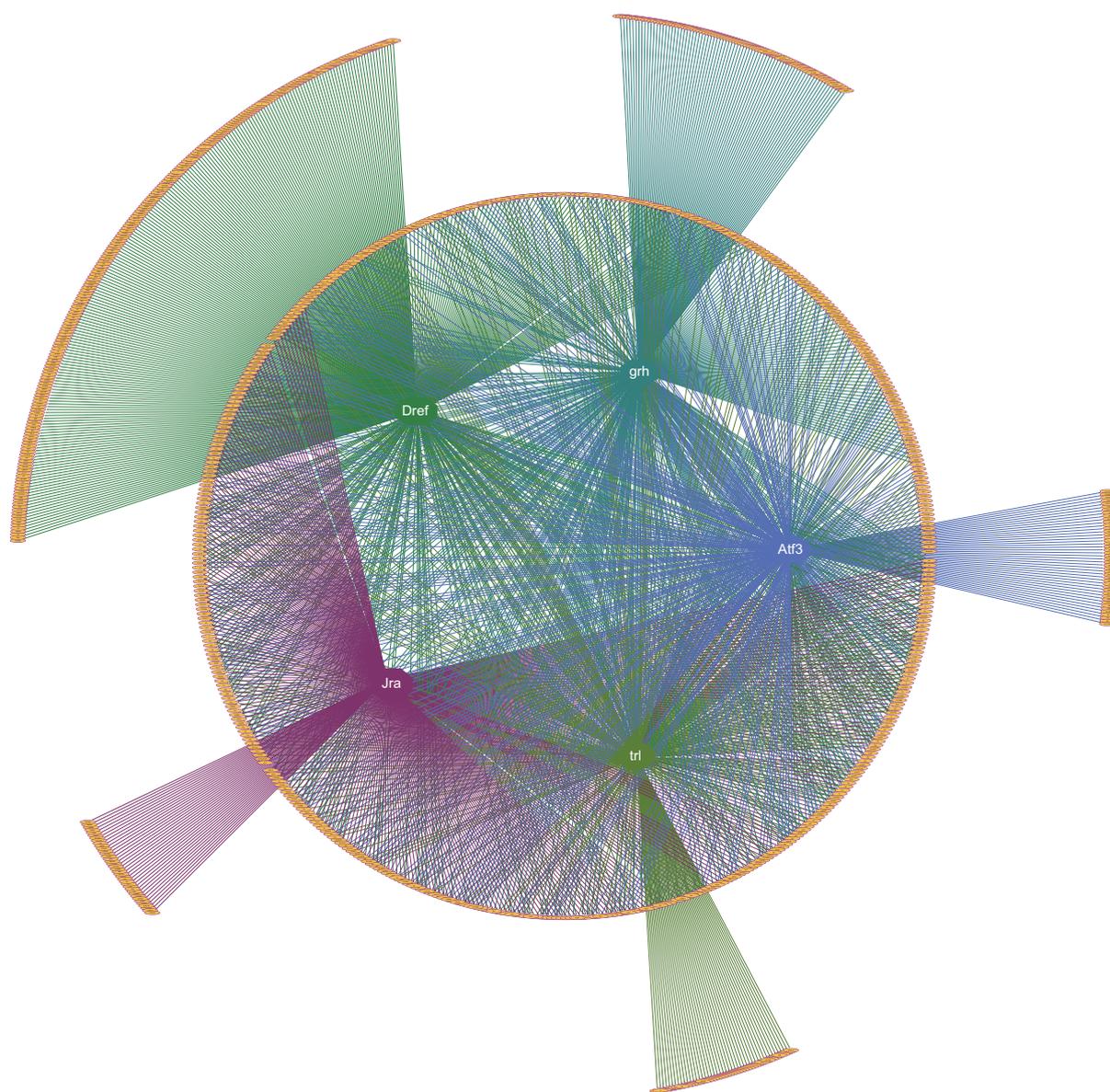


**Figure 32 - Time-course GO.** Gene Ontology term enrichment of differentially expressed genes at successive time points visualized by ReviGO. Size of circles denotes number of genes. Circle color indicates the  $p$ -value of each term. Highly similar GO terms are linked by edges in the graph.



**Figure 33 - Expression profiles of upregulated transcription factors.** Heat map showing the expression fold change of genes encoding transcription factors upregulated in at least one time point throughout the recovery process. Gene names are shown.

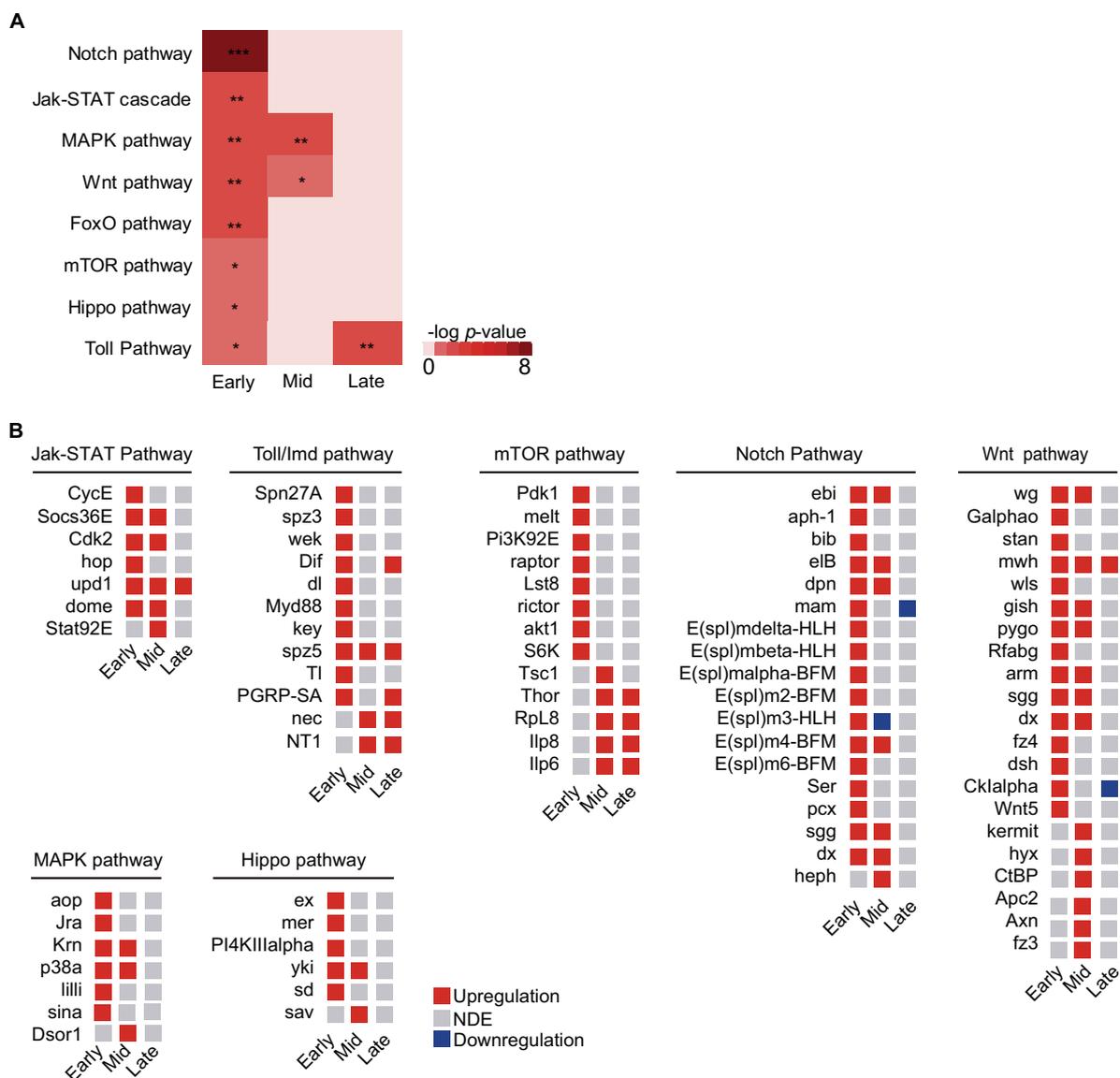
Then we applied iRegulon (Janki et al. 2014) to our set of DE genes at the early stage and computed the proportion of genes that can be regulated by the TFs within the same set of genes. We found that a 50% of the upregulated genes can be explained by just five of the TFs upregulated at that same stage: *Jra*, *Atf-3*, *grainy head (grh)*, *thritoxax-related (trl)* and *DNA replication-related element factor (Dref)*. (Fig. 34).



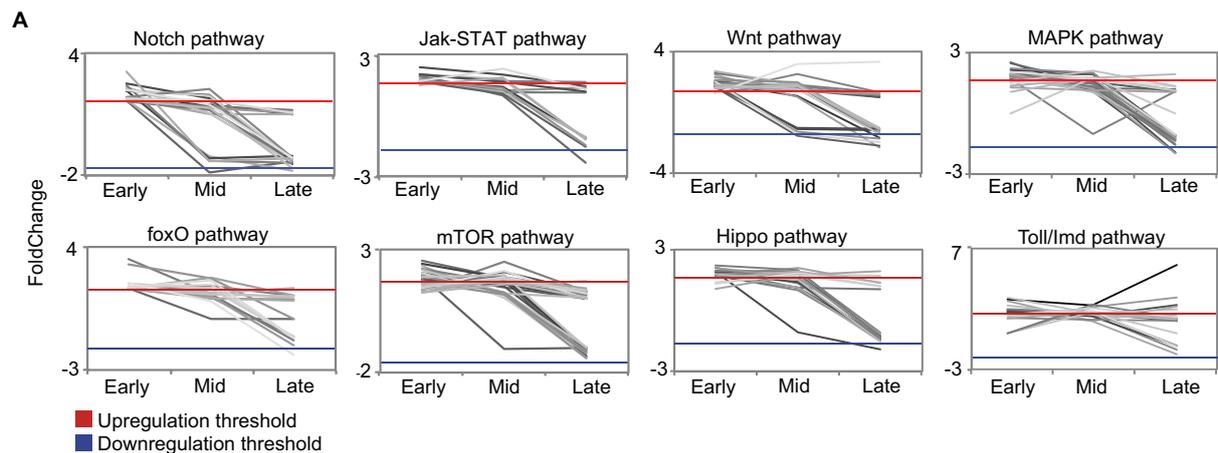
**Figure 34 - iREGULON.** Network showing TFs (centred in the network) connected to all their putative targets. Each yellow dot is a single gene upregulated at the early stage. Each line connects the TF with its associated gene.

## Co-expression of genes involved in signalling throughout regeneration

Next we analyzed which signalling pathways could be involved in the process. We used KEGG mapper and found 8 pathways enriched at the early stage of regeneration, the most enriched being Notch, Jak-STAT, MAPK and Wnt (Fig. 35A,B). In agreement with the main burst of active transcription, the enrichment occurred mainly at the early stage and was recovered though time. Moreover, such pathways are already known to be active at the onset of regeneration. Hence, upregulation of such pathways means that a more sustained activation throughout the process is required. Finally, when we analyzed expression of DE genes in each pathway over time, we found similar transcription patterns for several members of the pathway, indicating they are mainly co-expressed (Fig. 36).



**Figure 35 - Signaling pathways upregulated in regeneration.** (A) Heat map of pathway enrichment in the set of upregulated genes at each time point. The level of significance is denoted with stars (\*  $p < 0.05$ , \*\*  $p < 10^{-2}$ , \*\*\*  $p < 10^{-3}$ ). (B) Expression profile of the upregulated members of the enriched pathways at the different time points.



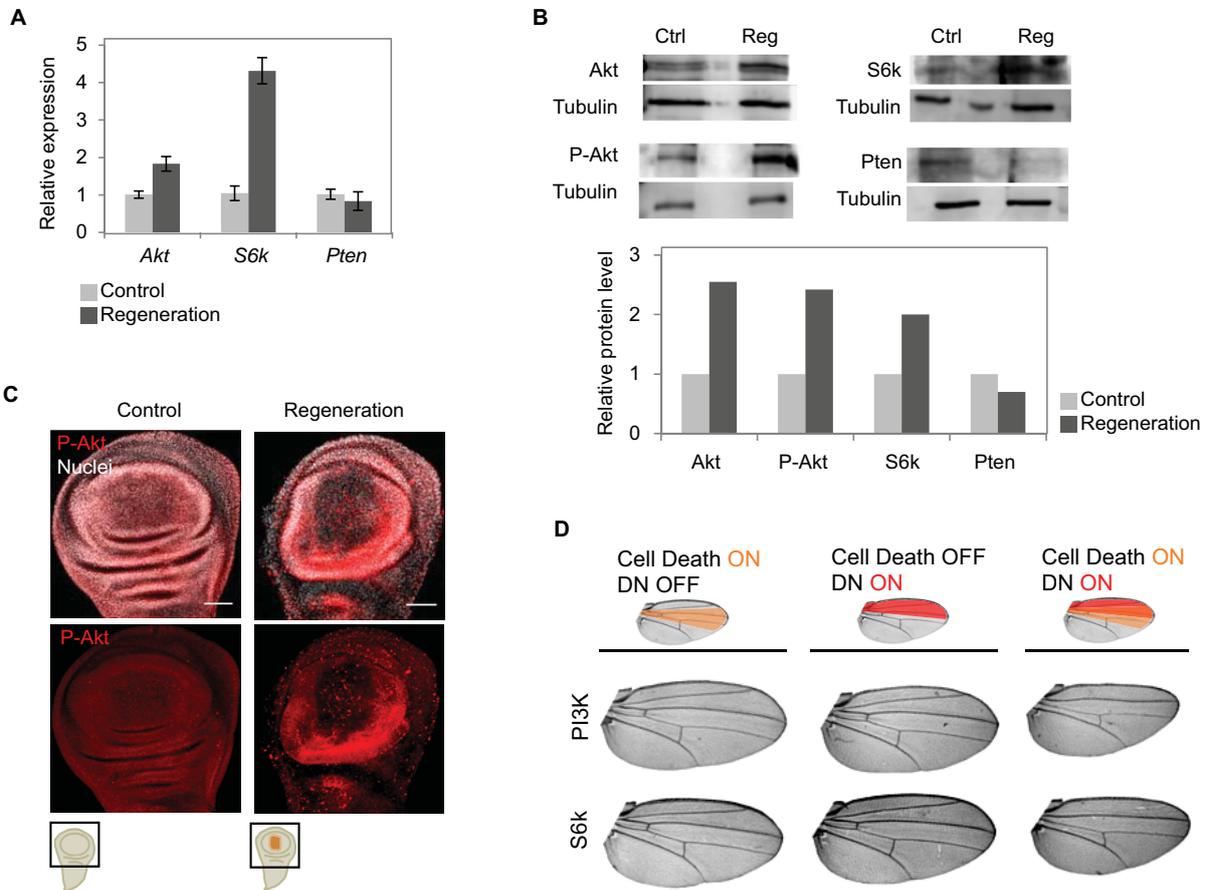
**Figure 36 - Expression of DE genes from signaling pathways .** (A) Line plots showing expression changes through time of DE genes that belong to signaling pathways significantly enriched in regeneration. Expression is shown as fold change between control and regeneration at each time point. Each gene is plotted as a single line.

### mTOR pathway is required for regeneration

Many of the aforementioned pathways have already been demonstrated as essential for wing disc regeneration (Reviewed in Hariharan and Serras, 2017), which gives robustness to our study. Nonetheless, to further validate the results obtained we looked for the requirement in the response to cell death of the mTOR pathway, which, despite its role in development has not been analyzed in the regeneration context in the wing disc.

First, we performed qPCR on DE genes of the pathway and observed that the elements that activate the pathway are upregulated meanwhile the inhibitors are downregulated (Fig. 37A), in agreement with the RNA-seq results. Some mechanisms as the non-mediated decay are known to inhibit the production of RNA, therefore even if there is more transcript does not necessarily mean that there is more protein (Reviewed in Schmid and Heick-Jensen 2018). By checking the same elements as before but using western blot, we observed a perfect correlation between the increase of transcripts and the increase of protein (Fig. 37B). Then, we assayed by immunofluorescence the activity of Akt, the core activator of the pathway, by analyzing its active form, which is the phosphorylated one (P-Akt). We observed an increase of the signal in the cells surrounding the wound, meaning that the pathway is precisely activated in the regenerating tissue (Fig. 37C). Finally, we assessed the requirement of Ribosomal protein S6 kinase (S6k) and PI3K in regeneration by testing the capability to regenerate upon their depletion. We used the double transcriptional transactivator system consisting of the *sa<sup>E/Pv</sup>-LHG lexO-rpr* to induce apoptosis combined with the *UAS/Gal4* and *Gal80<sup>ts</sup>* systems to drive the expression of a Dominant Negative (DN) construct of the two proteins in the anterior domain (*cubitus interruptus*, *ci-Gal4*) of the wing disc. The induction of cell death alone or the expression of each DN construct without the induction of cell death had no effect on the adult

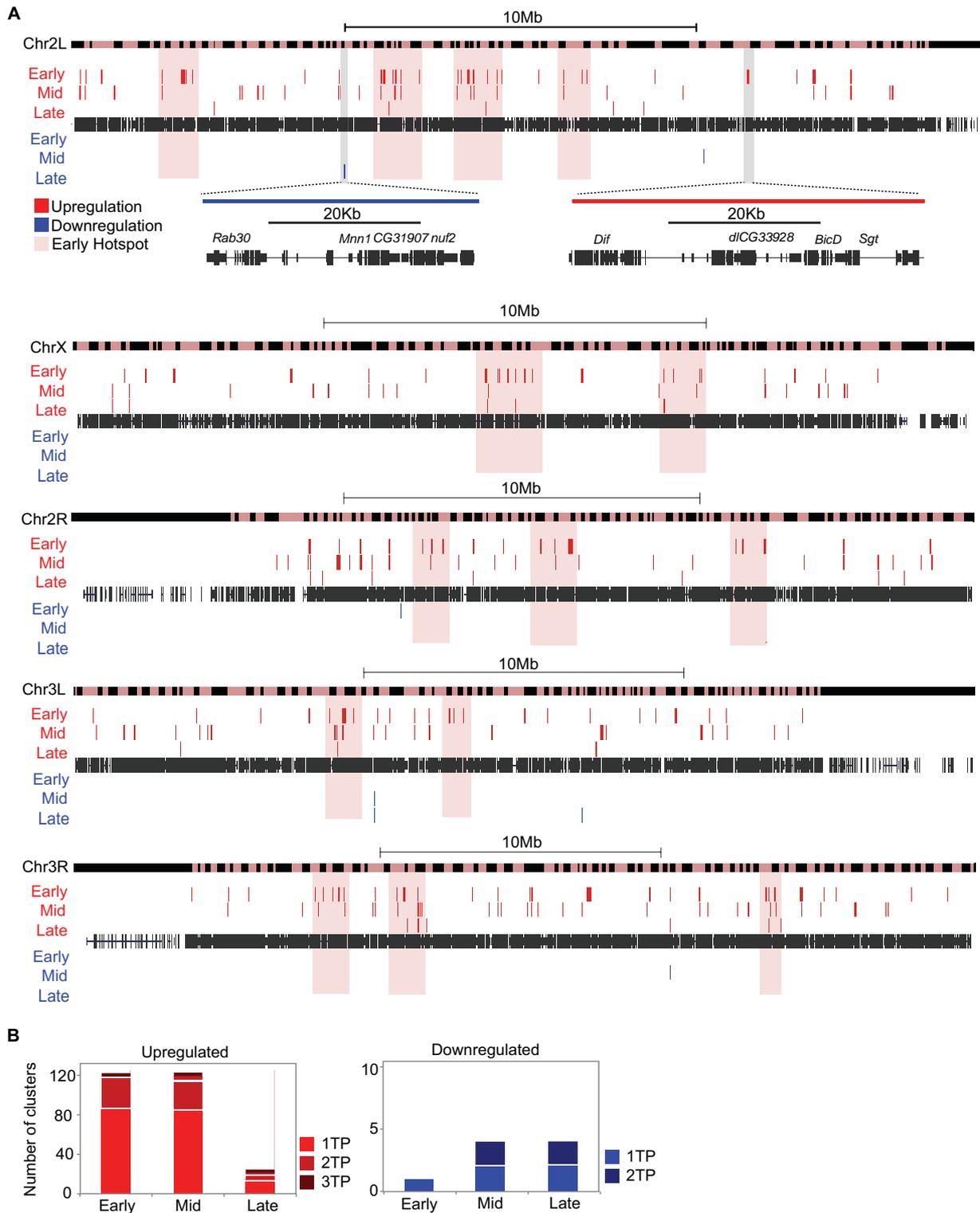
wings. However, the combination of each DN with genetic ablation resulted in non-regenerated wings, with a reduction of the wing size and an aberrant vein patterning (Fig. 37D). Altogether, these results indicate that the mTOR pathway is not only upregulated but also activated and required for proper regeneration.



**Figure 37 - Requirement of mTOR pathway in regeneration.** (A) qPCR analysis at the early stage of DE genes of the mTOR pathway. The results are presented as fold change enrichment between control and regeneration. Error bars represent standard error of the mean from three biological replicates. (B) Western Blot analysis of the early stage of DE genes of the mTOR pathway (top). Bar plot showing the relative protein levels, which are calculated as pixel intensity of the band. (C) Confocal images of wing discs (control and regeneration at the early stage) stained with P-Akt antibody. (D) Adult wings showing the predominant phenotype observed upon the depletion of S6k an PI3K in control and regeneration conditions (bottom). The region where the DN was expressed is highlighted in red and the apoptotic region in orange (top).

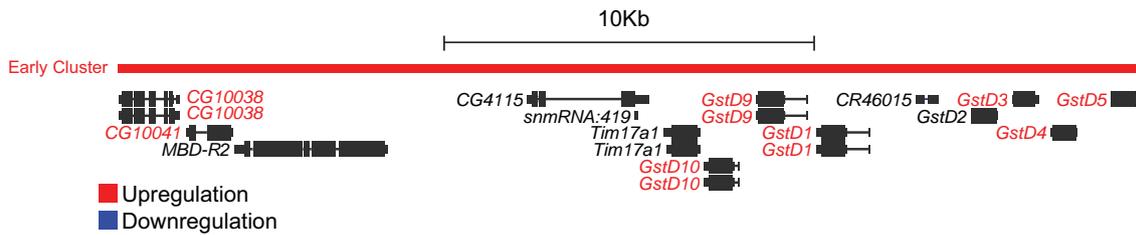
### Genomic clustering of differentially expressed genes involved in regeneration

A number of transcriptomic studies have demonstrated that genes with similar patterns of gene expression are frequently located close to one another in linear genomes (Boutanaev et al. 2002; Sproul et al. 2005; Michalak 2008; Corrales et al. 2017). Hence, we examined the chromosomal distribution of DE genes in all the time points using CROC (Pignatelli et al. 2006), a program that takes into account the genomic distance between genes in a given set to find which ones are in more proximity than expected (Fig. 38A).



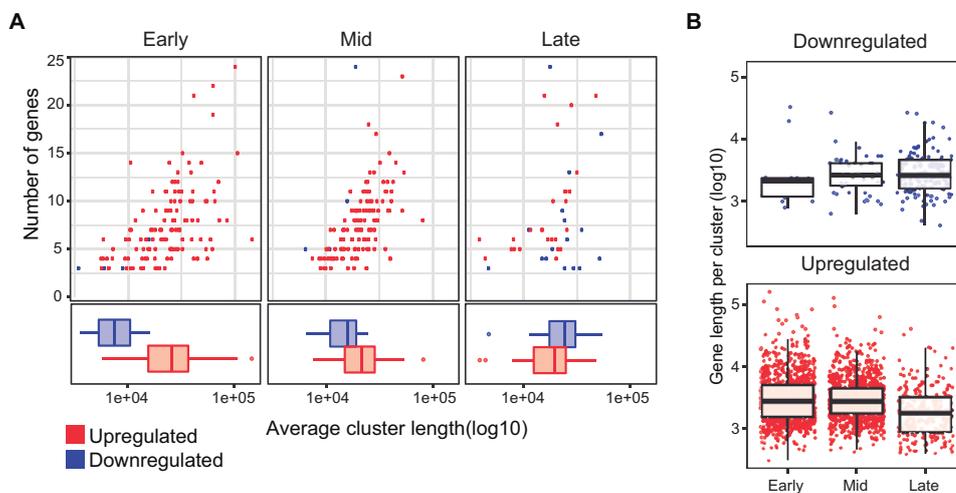
**Figure 38 - Genomic distribution of clusters.** (A) Genomic map of clusters of differentially expressed genes on *Drosophila* chromosome 2L. Each red or blue box represents one single cluster. The size of each box denotes the length of each cluster. Zoomed regions show one down regulated cluster (blue, left) and one up regulated cluster (red, right). Hotspots at the early stage are highlighted in pink. (B) Bar plot showing the number of clusters identified only in one time point, in two time points and in all three time points.

We identified several clusters of upregulated genes mostly at early and mid regeneration (126 and 124 respectively), meanwhile we only found few clusters for downregulated genes at the same time points (2 and 4 respectively) (Fig. 38B). Among upregulated genes we found members of the same family that are already known to belong to clusters, which confirms our approach. One example is the Glutathione S transferase (GstD) cluster formed by *GstD9*, *GstD10*, *CG10038*, *GstD1*, *GstD5*, *GstD3*, *CG10041* and *GstD4* (Fig. 39). (See Annex III for a complete list of clusters). When we analyzed the distribution of the clusters across the genome we observed that in some cases they were also located close to one another. Thus, we applied CROC to the clusters itself and found that, indeed, some of them create hotspots across the genome (Fig. 38A) (See Annex III for a complete list of clusters hotspots).



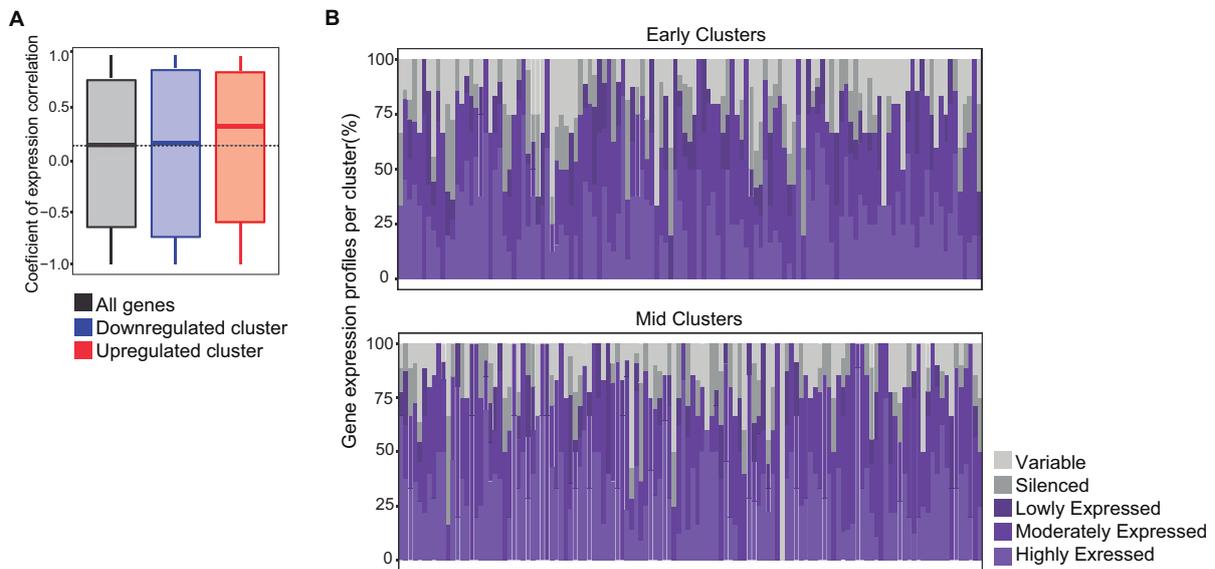
**Figure 39 - GstD cluster.** Screenshot showing an early upregulated cluster containing the members of the GstD family.

In order to better characterize the clusters we performed deeper analyses on different features. First we studied cluster size, which varies greatly depending on the number and the size of genes (Fig. 39). We found that average cluster length is ~25.5Kb and the average number of genes is ~7.7, being the shortest cluster in genomic length ~3.5Kb with 3 protein coding genes, while the longest ~150Kb with 6 protein coding genes. Also, there are three clusters with 24 genes, the largest numbers, with genomic sizes ranging from 17 to 100Kb. Besides, mean gene length in DE genes (whether up or downregulated) is almost the same independently of the time point.



**Figure 40 - Cluster size.** (A) Scatter plot showing the number of protein coding genes and the length of the cluster. Each dot represents a cluster (top). Box plot showing the average cluster length (bottom). (B) Box plot showing the average gene length per cluster. Each dot represents a cluster.

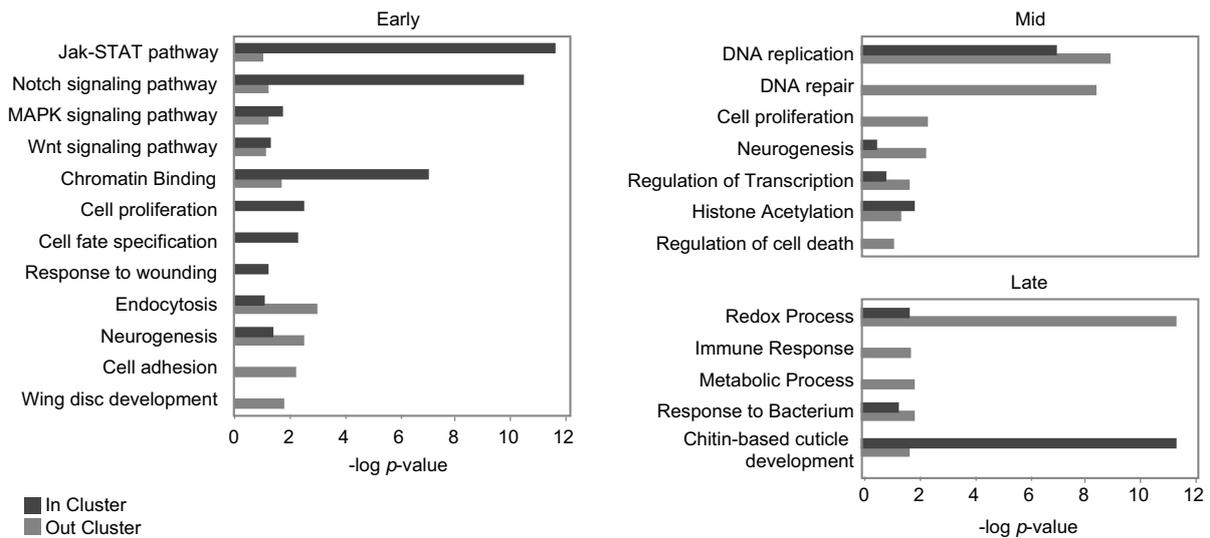
To determine whether there was co-regulation of genes inside the same cluster we computed the Pearson coefficient of expression correlation for every pair of genes within a cluster through time. We found that genes within upregulated clusters have higher median pairwise correlation than genes in downregulated clusters and genes overall (Fig. 40A). Then, we calculated the average gene expression profile through time per upregulated cluster at early and mid regeneration stages. We found that the proportion of variable genes is small compared to genes with moderate or high expression, meaning that genes inside upregulated clusters show similar expression profiles through time (Fig. 40B). Altogether, upregulated clusters are composed by genes that are not only in close proximity but which expression follows a similar behaviour through time.



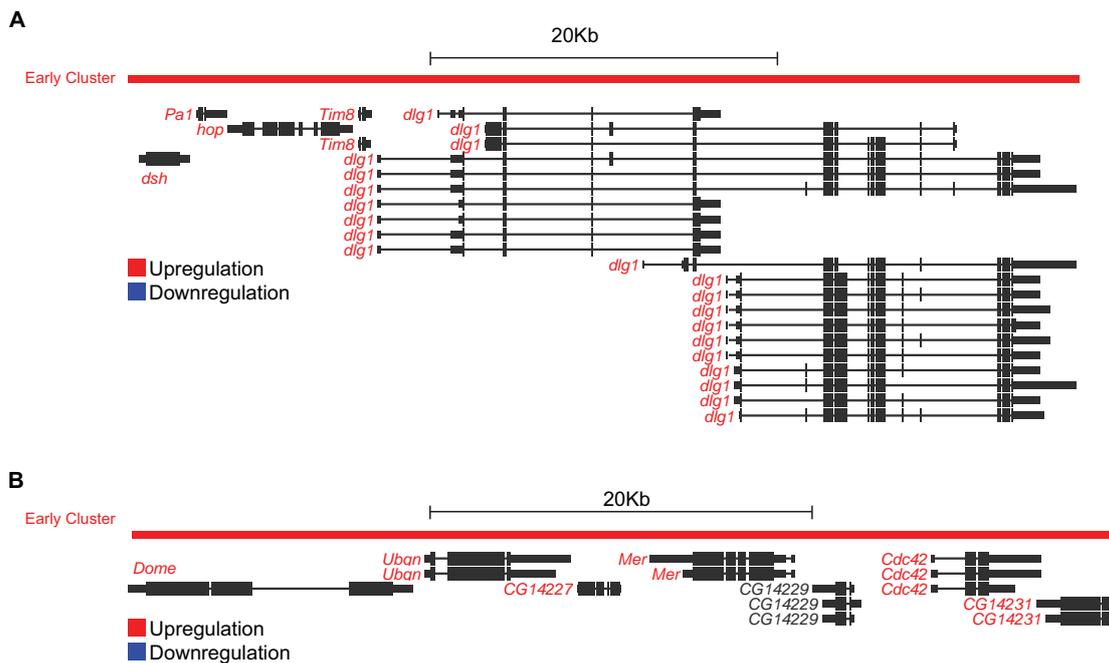
**Figure 41 - Gene expression in clusters.** (A) Box plot showing the coefficient of correlation between pairs of protein-coding genes inside the same cluster through time. (B) Bar plot showing in percentage the average gene expression profile through time per cluster. Each bar represents one individual cluster.

Finally, we wondered if upregulated genes inside or outside clusters could be acting in different biological processes. We performed GO analysis and found that genes within genomic clusters at the early time point are significantly linked to signaling pathways, proliferation and response to wound, while upregulated genes outside clusters are more associated with development, cell adhesion or neurogenesis (Fig. 42).

Next we investigated the composition of clusters at the early stage containing signalling pathway members. We found that even if some clusters can be composed by members of the same pathway, the great majority are composed by members of different pathways. For instance, one cluster contains the *hopscotch* (*hop*), *discs large 1* (*dlg1*) and *dishevelled* (*dsh*) genes, which belong to Jak-STAT, Notch and WNT pathways respectively (Fig. 43A). Another cluster contains *domeless* (*dome*), *Merlin* (*Mer*) and *Cdc42*, which belong to Jak-STAT, Hippo and MAPK pathways respectively (Fig. 44B).



**Figure 42 - Biological processes related to clustered genes.** Gene Ontology term enrichment for the set of upregulated genes located inside or outside the clusters at early, mid and late regeneration time points. All the categories plotted are significant in at least one group of genes (absence of bar denotes no enrichment in that group).



**Figure 43 - Clusters containing signalling pathway members.** (A, B) Screenshot showing early upregulated clusters containing members of different signaling pathways.

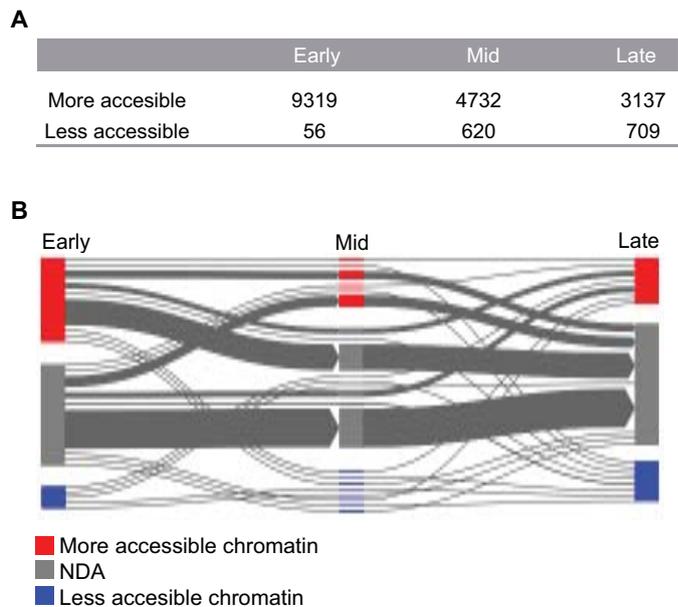
Taken together, all these results suggest that large regions, rather than individual genes, may be controlled by the same regulatory elements to be turned on at once, in bulk, by cluster co-regulation.

## **Chapter II: The regulome of regeneration**



## The chromatin landscape of regeneration

To shed light in how chromatin dynamics could trigger the regenerative transcriptional profile we performed in depth-analysis of chromatin accessibility data. Consistently with RNA-seq, the analysis of ATAC-seq data showed that the number of regions more accessible in injured than controls samples was the highest at early regeneration and decreased with time, correlating with the activation of transcription at the initial steps (Fig. 44).

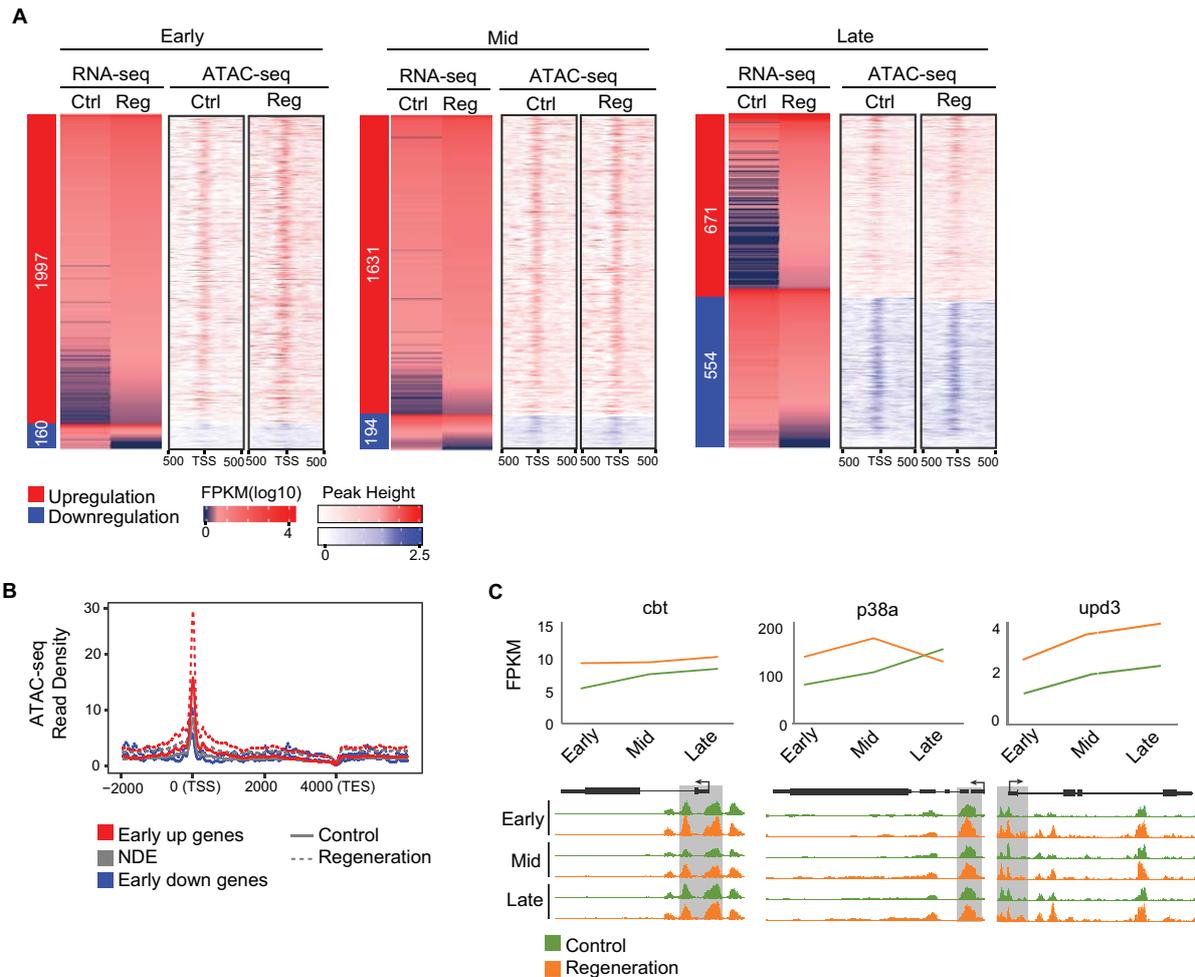


**Figure 44 - ATAC-seq.** (A) Number of differentially accessible regions at all time points. (B) Flux-plot showing ATAC-seq dynamics (differential accessible regions) throughout the different time points. Each line represents a set of accessible regions equally behaving through time. The line width denotes the number of accessible regions. (absence of bar denotes no enrichment in that group).

We combined both data-sets (RNA-seq and ATAC-seq) to examine the CP region ( $\pm 100$ bp of TSS) of DE genes. All expressed genes at the different time points presented and accessible CP in the same stage. (Fig. 45A). Besides, we examined how the CP of DE genes at the early stage behaved over the regeneration process. We observed that in the early stage, the TSS of upregulated genes was clearly more accessible in regeneration compared to the TSS overall and it was recovered to control levels towards the end of the process (Fig. 45B).

Among the upregulated genes with a more accessible CP, we found genes such as *upd3*, *Jra*, *cbt* and *p38a* (Fig. 45C), which are known to be required only in a few cells around the wound both after cell death and physical injury (Blanco et al. 2010; Katsuyama et al. 2015; Santabárbara-Ruiz et al. 2015). This demonstrates that our approach is sensitive enough to identify changes affecting small numbers of cells within the monitored population.

As differences in gene expression correspond mainly to a burst of active transcription after damage, we focused on regions that presented higher accessibility in regeneration compared to control.

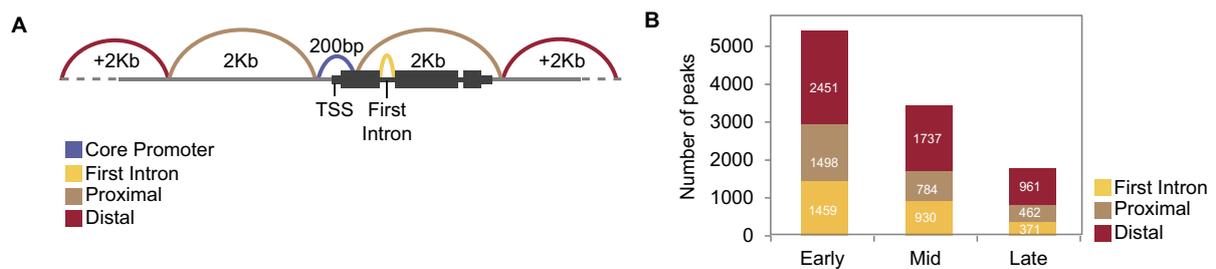


**Figure 45 - Correlation of RNA-seq and ATAC-seq data.** (A) Heatmaps showing RNA-seq signal of DE genes at each timepoint and and ATAC-seq signal around  $\pm 500$ bp of the TSS of protein coding DE genes at each timepoint. Sites are ordered by up and downregulation (shown in the left) and by gene expression based on regeneration samples. (B) Aggregation plot showing ATAC-seq read density at early stage (control and regeneration) for each set of DE genes (upregulated, non differentially expressed or NDE and downregulated) at early regeneration. The TSS of upregulated genes show the highest number of ATAC-seq reads in regeneration. (C) Expression profile (FPKM) of *cbt*, *p38a* and *upd3* across time in control and regeneration (top). Genome Browser screenshots depicting ATAC-seq peaks at the core promoter of *cbt*, *p38a* and *upd3* in control and regeneration through time (bottom)

## Damage Responsive Regulatory Elements

In order to differentiate the CP from regions representing putative enhancers we first distributed ATAC-seq peaks based on their genomic location according to their position relative to the TSS of the closest gene. Thus, we classified regions that become more accessible under damage conditions as being in: a) the CP ( $\pm 100$ bp of the TSS); b) first intron (FI) regions, the ones falling in the first intron; c) proximal regions, the ones falling  $\pm 2$ Kb of the TSS; d) distal regions, the ones falling more than  $\pm 2$ Kb away of TSS (Fig. 46A). We also obtained ATAC-seq data for untreated L3 wing imaginal discs, which represents the basal developmental stage of our tissue of study, to better characterize specific traits of putative damage enhancers (See Annex II for mapping statistics and replicate analysis of L3 ATAC-seq).

We consider as putative regeneration enhancers the more accessible regions belonging to first intron, distal and proximal locations that are concordant in all the replicates of both controls (early/mid/late controls and L3) and regeneration (see *Material and Methods for a detailed explanation*). We have named these regions **Damage Responsive Regulatory Elements (DRREs)**. Despite a decreasing number of accessible chromatin regions with time, we did not observe differences in the genomic distribution, suggesting that proportions of each type of enhancer are maintained over time (Fig. 46B).



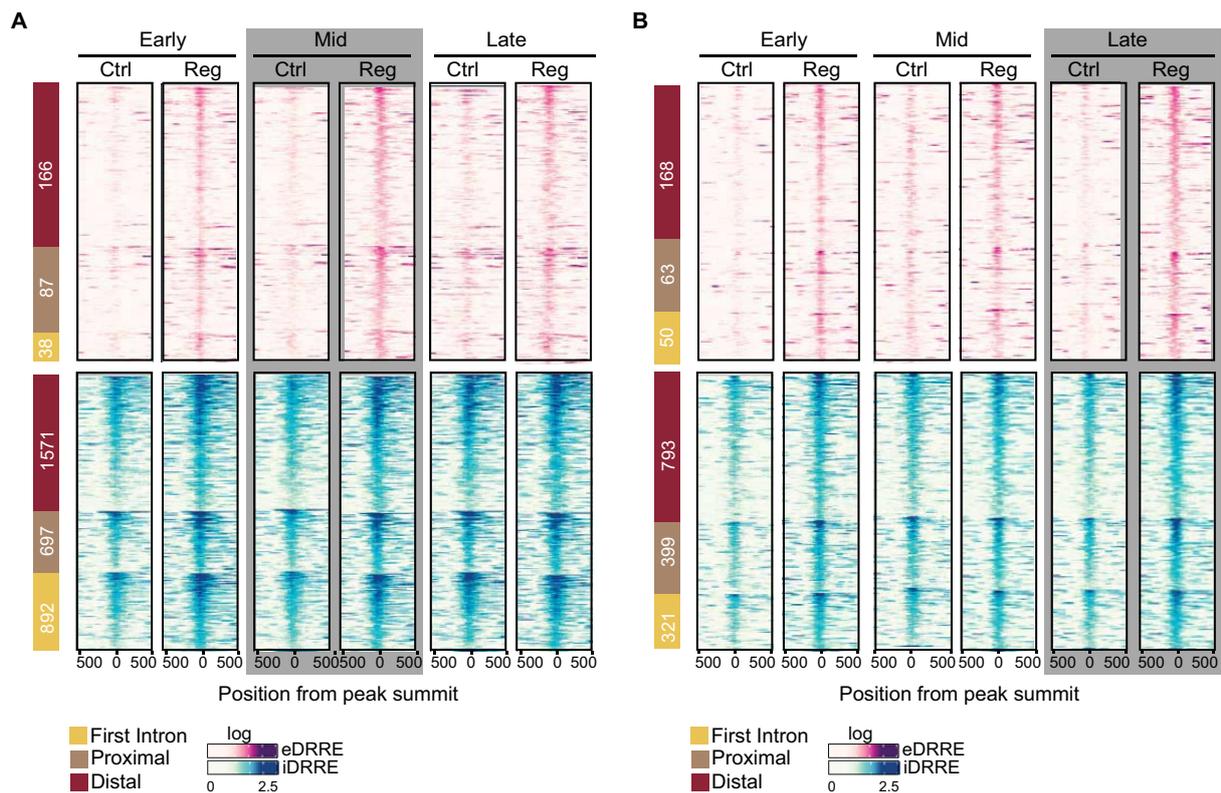
**Figure 46 - Genomic distribution of DRREs.** (A) Schematic overview of peak distribution in the genome. (B) Bar plot showing the number of DRREs at each time point falling in each genomic region: first intron, proximal and distal.

### ***emerging vs increasing DRREs***

By comparing peak height of accessible regions between regeneration and controls (L3 and control), we distinguished two types of DRREs, which following previous studies in zebrafish heart regeneration (Goldman et al. 2017) we named as: ***emerging* (eDRRE)**, open regions only found after damage or ***increasing* (iDRRE)**, regions already open in controls but displaying increased accessibility after damage (Fig. 47A). To ensure the robustness of our system we proceed as follows: a DRRE can only be considered as *emerging* if it is called in both replicates of regeneration and in none of control and L3 at each stage. In the same direction, an *increasing* DRRE had to be called in all the replicates, whether regeneration or control ones. Among the iDRREs we found the damage-activated WNT enhancer that has already proven to be crucial in imaginal disc regeneration (Fig. 47B) (Harris et al. 2016).

First we characterized the early stage as it is the one presenting the highest number of DRREs. We observed that *increasing* enhancers are the most frequent DRRE type as they represent a 93.7% of all DRREs compared to a 6.3% of *emerging*. We also analyzed their genomic distribution and found that compared to the iDRREs, eDRRE tend to occur more often in distal locations (Fig. 47C). Finally, since regions with or without nucleosomes may present different features (Jung et al. 2017), we also compared the nucleosome-free (NF) and mononucleosome (MN) fractions from the ATAC-Seq experiments. In iDRREs, we detected ATAC-Seq reads in the NF region, which was flanked by well-positioned nucleosomes both in control and regeneration samples, whereas for eDRRE we observed reads in the NF regions





**Figure 48 - Accessible chromatin landscape after cell death induction through time.** (A, B) Heatmaps showing nucleosome-free (NF) enrichment around  $\pm 500$ bp of the peak summit of DRREs trough time. Sites are ordered by genomic distribution (shown in the left) and by peak height based on ATAC-seq regeneration sample highlighted in gray.

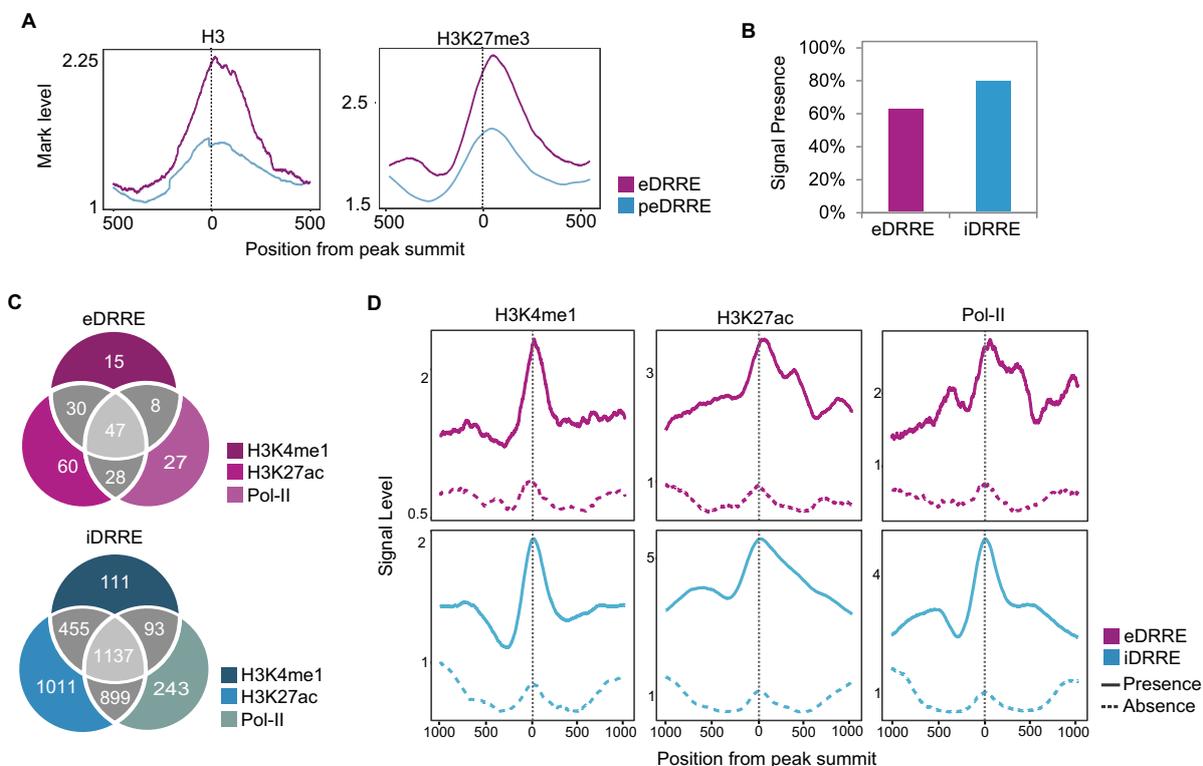
### The chromatin state of DRREs

Certain post-translational modifications of histone residues, such as H3K4me1 and H3K27ac, are predictive of active enhancers, whereas inactive ones are associated to H3K27me3 (Supek et al. 2011; Calo and Wysocka 2013; Shlyueva et al. 2014; Koenecke et al. 2016; Long et al. 2016). Pol-II occupancy and transcription are also predictive of active enhancers (Mikhaylichenko et al. 2018).

To further characterize DRRE states we first took advantage of the available ChIP-seq data at L3 wing discs on histone modifications (Pérez-Lluch et al. 2015; Loubière et al. 2016) and found that eDRREs but not iDRREs showed a positioned nucleosome (histone 3, H3) modified with the repressive mark H3K27me3 (Fig. 49A), which reinforces that eDRREs but not iDRRES are in closed chromatin in a non-regenerative context.

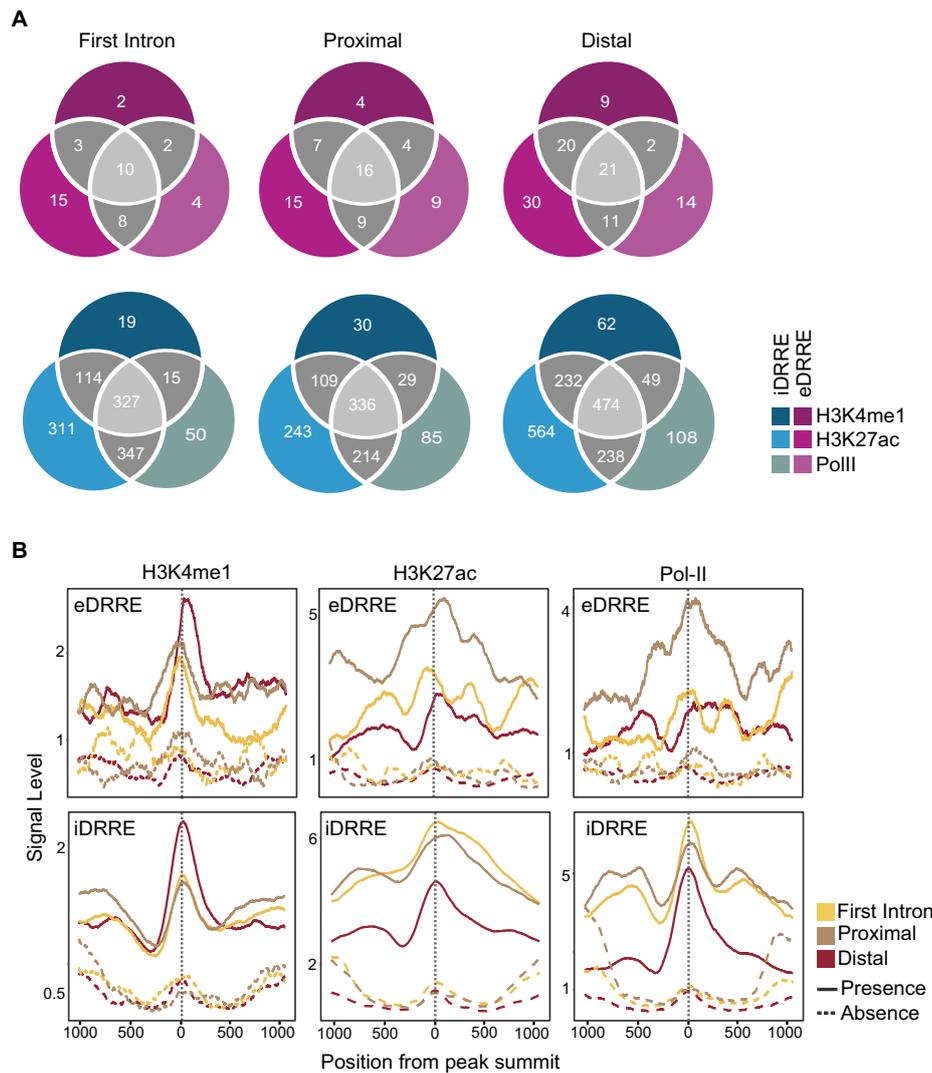
Afterwards, we performed ChIP-seq on H3K27ac, H3K4me1 and Pol-II at the early stage (See Annex II for mapping statistics). We found that 80% of iDRREs and 63% of eDRREs displayed features of active enhancers in regeneration as they presented at least one of the analyzed marks (Fig. 49B). Around 30% of eDRREs were marked only by one of them whereas

14% contained all of them. In the case of iDRREs, 8.1% presented only one feature in contrast to 20% containing all. In both types of DRREs the most representative mark is H3K27ac, being the one present in more active-like enhancers (74% and 70% of marked eDRRE and iDRREs respectively) and the one with a higher signal overall (Fig. 49C,D).



**Figure 49 - Chromatin features of DRREs.** (A) Average profile of H3 and H3K27me3 around  $\pm 500$ bp of the peak summit of DRREs at L3 wing discs. (B) Bar plot showing the percentage of presence or absence of signal at DRREs. (C) Venn Diagrams showing the intersection of H3K4me1, H3K27ac and PolII at DRREs in regeneration. (D) Average profile of H3K4me1, H3K27ac and Pol II at DRREs. A straight line denotes DRREs with the presence of at least one CHIP-seq signal; and a dotted line denotes the absence of any CHIP-seq signal

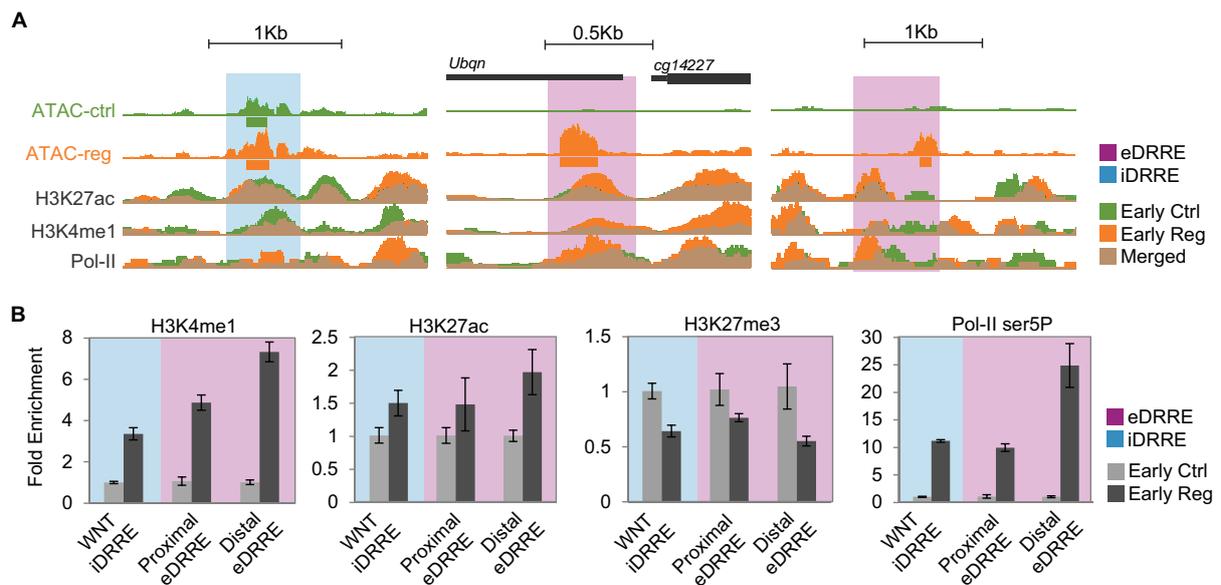
It is known that based on their genomic location, enhancers can be more associated to one mark or another (Reviewed in Catarino and Stark, 2018). Accordingly, we observed that H3K4me1 is preferentially located in distal positions meanwhile H3K27ac and Pol-II are more likely to be in proximal regions and in the first intron (Fig. 50A,B).



**Figure 50 - Chromatin features of DRREs by genomic distribution.** (A) Venn diagrams showing the intersection of ChIP signal in DRREs per genomic distribution. (B) Average profile of ChIP-seq signal around  $\pm 1000$ bp of the peak summit of DRREs in regeneration. A straight line denotes DRREs with the presence of at least one ChIP-seq signal; and a dotted line denotes the absence of any ChIP-seq signal.

Finally, we further confirmed ChIP-seq results by doing ChIP-qPCR analysis in individual DRREs. We selected the following genomic regions: the WNT enhancer (Harris et al. 2016) as an iDRRE; a proximal eDRRE located inside a cluster of upregulated genes and 1,5 Kb away of the TSS of *CG14227*; and a distal eDRRE located more than 48 Kb away of the nearest upregulated protein-coding gene *leucine-rich-repeats and calponin homology domain protein (Irch)*. We observed a decrease of H3K27me3 as well as an increase of H3K4me1, H3K27ac and the active form of Pol-II phosphorylated in Serine 5 (Pol-II ser5P) in the WNT enhancer (iDRRE) and in the proximal and distal eDRREs (Fig. 51A,B).

Altogether these analyses indicate that eDRREs are indeed in close chromatin at L3 wing discs, becoming accessible and acting as enhancers only after damage. Meanwhile, even if iDRREs are already active, they gain active marks as they become accessible in more cells.

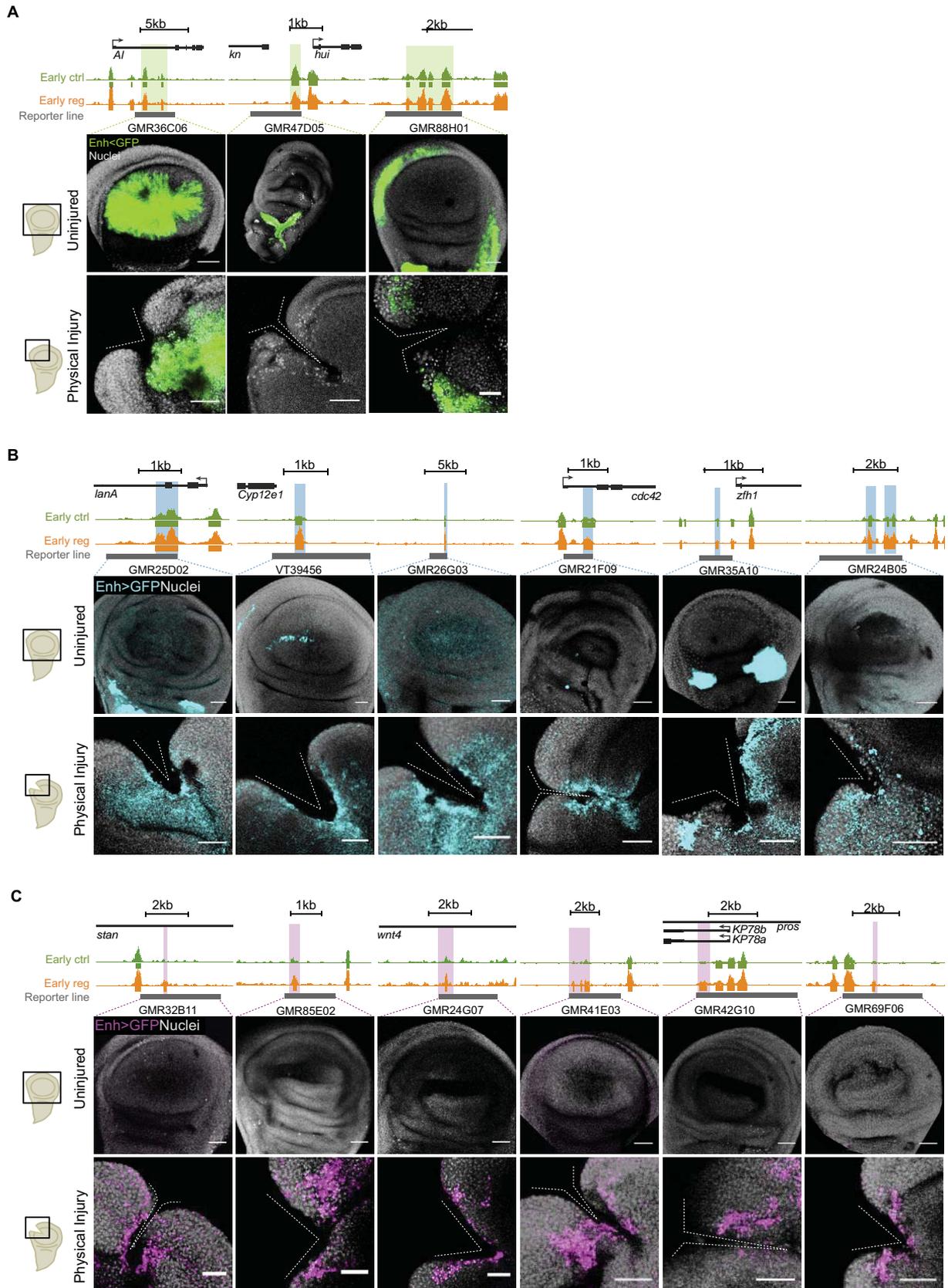


**Figure 51 - ChIP-seq validation of DRREs.** (A) Genome browser screenshot showing ATAC-seq and ChIP-seq profiles (control and regeneration) of the DRREs tested by ChIP-qPCR. (B) ChIP-qPCR analysis of H3K4me1, H3K27ac, H3K27me3 and Pol II-ser5P on individual DRREs at the early stage. ChIP results are presented as fold change enrichment between control and regeneration. Error bars represent standard error of the mean from two biological replicates.

## DRREs are active upon different types of damage

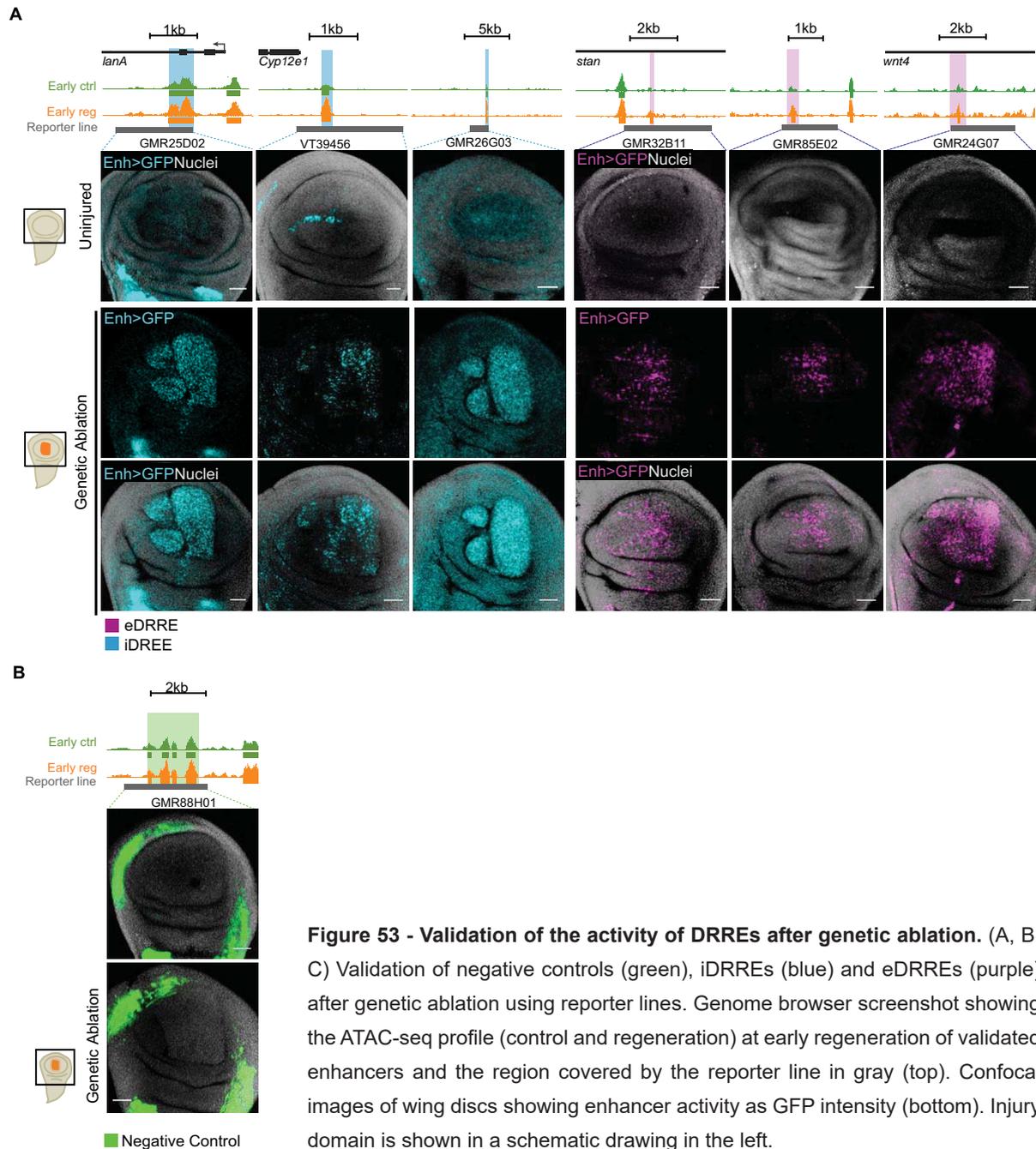
One property of enhancers is their ability to retain transcription-activating functions outside their endogenous contexts. Thus, we used reporter lines and tested them upon different types of injury to confirm the damage-induced activity of DRREs *in vivo*.

First we tested 15 different lines upon physical injury which include: 3 lines for negative controls (open regions in control and regeneration but not differentially accessible), 6 lines for eDRREs and 6 lines for iDRREs. We combined reporter lines containing DRREs cloned upstream a Gal protein with a *UAS-GFP* line and checked GFP expression after cutting and culturing the disc (See *materials and methods*). Negative control lines only drove GFP expression in their wild-type activation pattern and not in the damaged area (Fig. 52A). However, we observed that in iDRREs with a defined endogenous activation pattern there was ectopic GFP expression in the wound, meanwhile in iDRREs ubiquitously active we observed an increase of the GFP expression in the wounded zone (Fig. 52B). Regarding, eDRREs we confirmed once more that there is no endogenous activity in the disc although they were able to drive GFP expression in the injured region (Fig. 52C).



**Figure 52 - Validation of the activity of DRREs after physical injury.** (A, B, C) Validation of negative controls (green), iDRREs (blue) and eDRREs (purple) after physical injury using reporter lines. Genome Browser screenshot showing the ATAC-seq profile (control and regeneration) at early regeneration of validated enhancers and the region covered by the reporter line in gray (top). Confocal images of wing discs showing enhancer activity as GFP intensity (bottom). Injury domain is shown in a schematic drawing in the left.

Then we selected 7 of the previously validated lines (1 line for a negative control, 3 lines for eDRREs and 3 lines for iDRREs) to test them upon genetic ablation. We combined the *UAS-Gal4* system to assess the enhancer activity with the *LHG-lexO* system to induce cell death (See *Materials and Methods*). In accordance with the results obtained after physical injury, we observed ectopic or increased GFP expression around the wound in iDRRE and eDRRE (Fig. 53A) and not in the negative controls (Fig. 53B). Altogether, these results confirm the occurrence of bona-fide enhancers responding to injury.

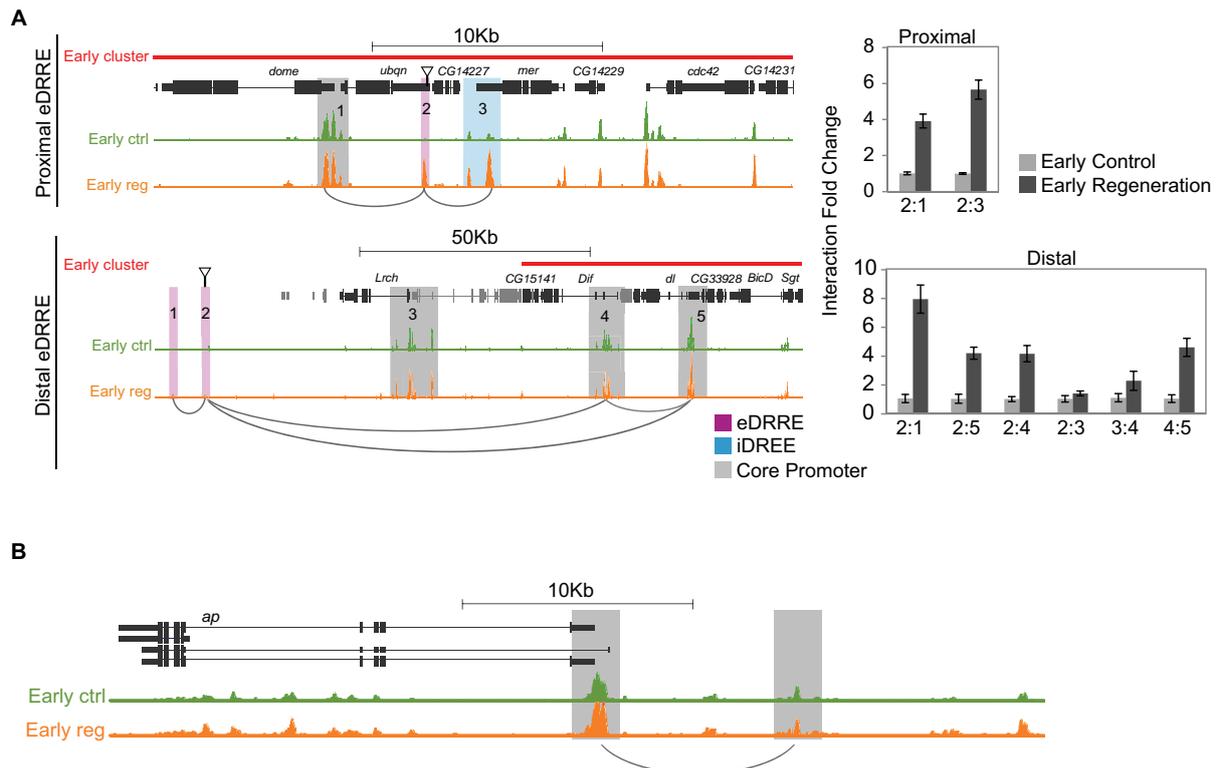


## Chromatin 3D-rearrangements in regeneration

The analysis of the transcriptome of regeneration suggested that one well positioned enhancer could turn on, in bulk, all genes inside a cluster of upregulated genes. As already mentioned, it is known that spatial chromatin organization connects active enhancers to target promoters in *cis* to regulate gene expression (Dekker et al. 2013; Rowley and Corces 2016; Schwartz and Cavalli 2017; Cubebñas-Potts et al. 2017). Consequently, we studied whether eDRREs could contact other accessible regions, iDRREs or CPs, associated to clustered genes to putatively drive their expression.

We selected the same proximal and distal eDRREs tested for enhancer features by ChIP-qPCR analysis (Fig. 51). The proximal eDRRE is located inside a cluster containing upregulated genes related to growth, belonging to the most enriched pathways at early regeneration: *dome*, *mer* and *cdc42*, from Jak-STAT, hippo and MAPK pathways respectively (Fig. 54A). The distal eDRRE is located 50Kb away of a cluster which contains *Dorsal-related immunity factor (Dif)* and *dorsal (dl)*, the effectors of the Toll pathway (Fig. 54A).

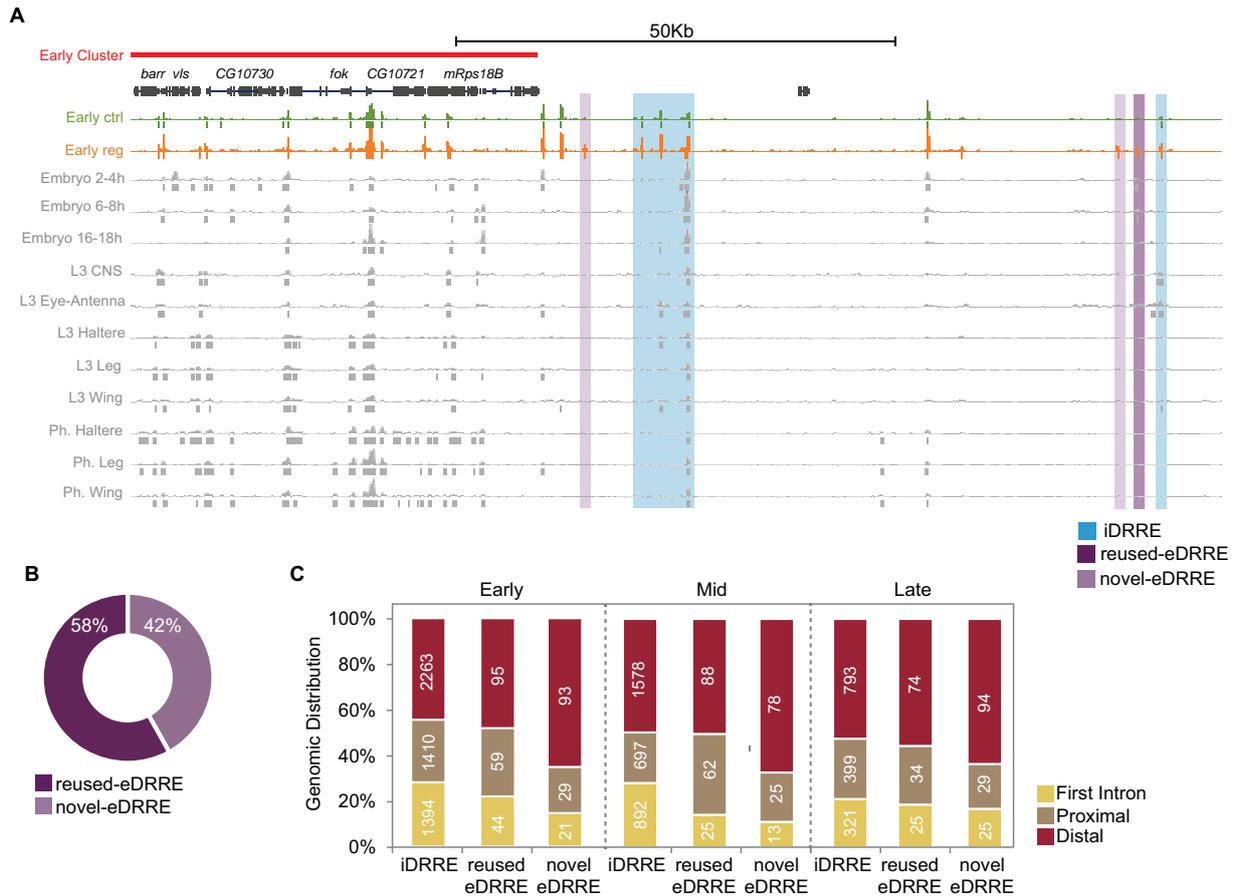
Next, we performed 3C-qPCR analyses at the early stage to test if the selected eDRREs could be interacting with other accessible regions (CPs and iDRREs). We used regions already proven to establish contacts in L3 wing discs (Bieli et al. 2015) as control (Fig. 54B). We detected interactions between the proximal eDRRE, and both the CP of *dome* and a proximal iDRRE located at the transcription-ending site of *mer* (Fig. 54A), within the same cluster of co-regulated genes. We also detected physical contact between the distal eDRRE, and the CP of the *Dif* and of *CG33928* genes, which are located within the cluster of upregulated genes. In contrast, we did not detect any interaction between the same distal eDRRE and the CP of the *Irch* gene, which, in spite of being also upregulated and closer in the genome, is outside the cluster. This suggests that the distal eDRRE could specifically regulate the entire cluster (Fig. 54A).



**Figure 54 - Chromatin rearrangements in regeneration.** (A) Genome Browser screenshots at the early stage highlighting regions used for 3C analyses (left). Arrows indicate the eDRRE peaks used as bait. 3C-qPCR analysis showing interaction levels between eDRREs and both CP and iDRRE (right). 3C results are presented as the fold change of the interaction between control and regeneration. Error bars represent the standard error of the mean from two biological replicates. Positive interactions after damage are marked as connectors in the genome browser screenshot. (B) Genome Browser screenshot depicting the interaction of a known enhancer on the Apterous (*ap*) gene (Bieli et al. 2015) used as a control on 3C experiments.

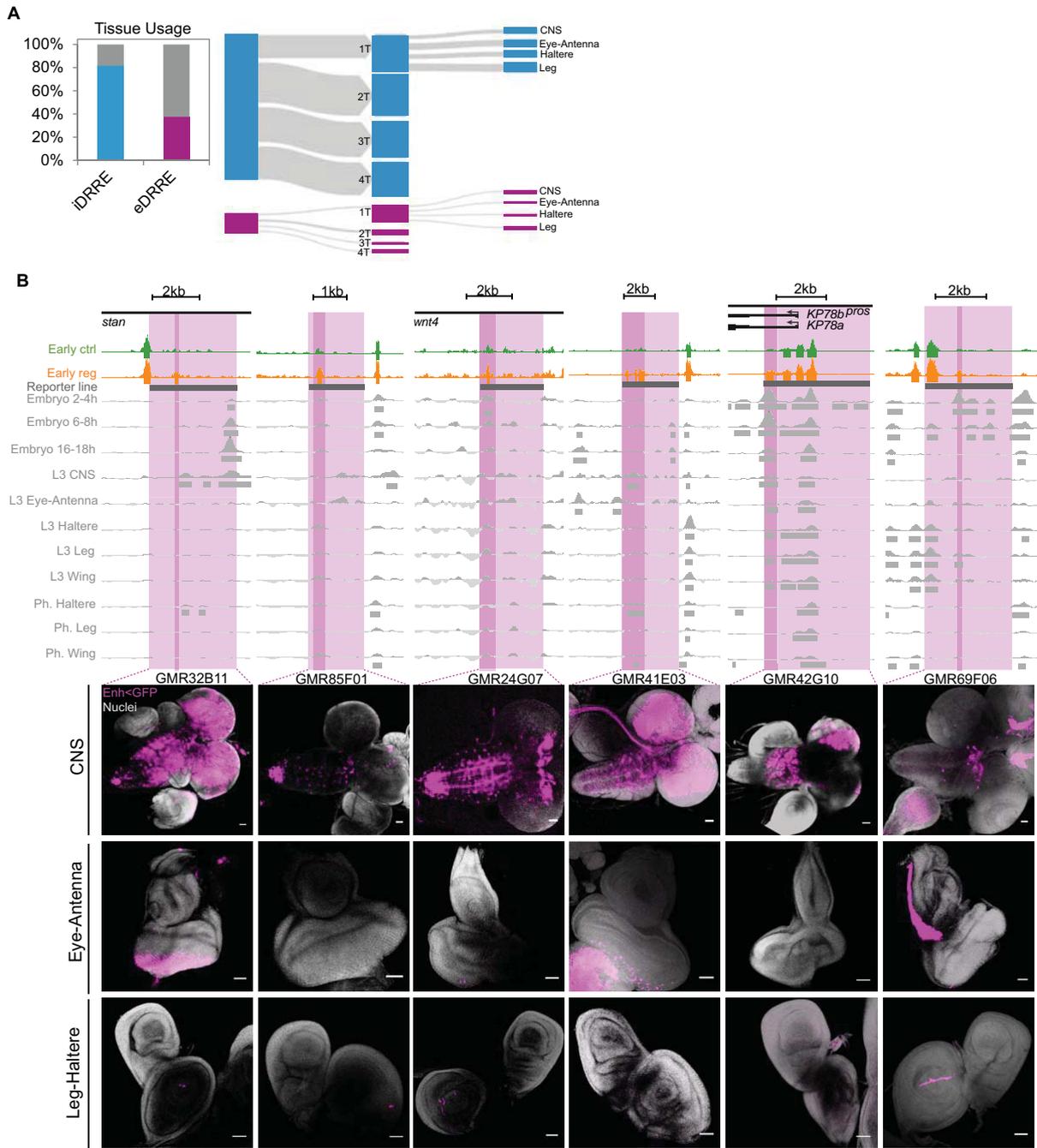
## A specific regeneration regulome

Since enhancers can be used in a context-dependent manner (Nègre et al. 2011; McKay and Lieb 2013; Wei et al. 2016; Erceg et al. 2017), we assessed whether DRREs were involved in other developmental events, regardless of their specific role in wing disc regeneration. We took advantage of chromatin accessibility data from different tissues (Central Nervous System and eye-antenna, leg and haltere discs) and stages of fly development (2-4h, 6-8h, 16-18h of embryonic development and pharate) (McKay and Lieb 2013) and checked for accessibility of DRREs. We first focused on the early stage and found that, 58% (198) of eDRREs were already used in other tissues or across different developmental stages; we renamed these as reused-eDRREs. The remaining 42% (143) therefore represented novel-eDRREs (Fig. 55A,B), a class of enhancers likely to be, thus, regeneration specific. Regarding their genomic distribution we observed that novel-eDRREs tend to be more distal than iDRREs and reused-eDRREs in all the regeneration time points (Fig. 55C). We observed that both the proportion of DRRE types and the genomic distribution was conserved through time (Fig. 55C).



**Figure 55 - DRREs used in other tissues and at other developmental stages.** (A) Genome Browser screenshot depicting the alignment of iDRRE, reused-eDRRE and novel-eDRRE with previously accessible regions identified by FAIRE in different developmental stages and tissues. (B) Classification of eDRREs usage. FAIRE data for embryo (2-4h, 6-8h, 16-18h), L3 (Central Nervous System, Eye Antenna, haltere, leg) and pharate (haltere, leg, wing) is used. eDRREs falling in at least one of the data sets are considered as reused-eDRREs. (C) Genomic distribution of DRREs: first intron, proximal and distal.

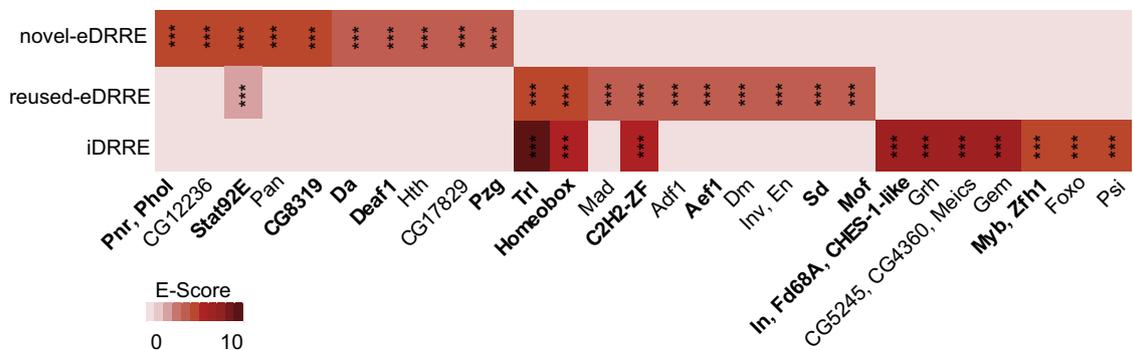
To further confirm the usage of the identified reused-eDRREs, we looked for the endogenous activity of the previously tested eDRREs reporter lines in other tissues. We found that in accordance with the comparative analysis, reused-eDRREs are indeed also active in some other tissues (Fig. 56A,B).



**Figure 56 - Tissue usage of DRREs.** (A) Flux plot of enhancers activated after damage showing their usage in other tissues at L3 stage. DRRE peaks are classified in base of their presence in tissues (Central Nervous system – CNS, eye- antenna disc, haltere disc and leg disc) and in the amount of occasions each enhancer is employed (up to 4 tissues -T). (B) Validation of tissue usage using reporter lines. Genome browser screenshot showing the ATAC-seq profile (control and regeneration) at early regeneration of validated eDRREs (highlighted in dark purple), the FAIRE profiles (in gray) and the region covered by the reporter line (highlighted in light purple) (top). Confocal images of CNS, eye-antenna discs, leg discs and haltere discs at L3 stage showing enhancer activity as GFP intensity (bottom).

## Novel-eDRREs possess a specific motif grammar

Next, we analyzed if each type of DRRE contained specific sequence features that could be recognized by different TFs for their further activation. We used iCIS target (Herrmann et al. 2012) and searched for motif enrichment in DRREs. First, selected only the motifs for which the corresponding TF was expressed in the wing disc (See Annex IV for a detailed list of the motifs found at each stage). Of these, we found that at the early regeneration stage 52% of TFs putatively binding eDRREs (whether novel or reused) and 43% putatively binding iDRREs were upregulated (Fig. 58). Moreover, we observed that, with the exception of Stat92E, the motifs found in novel-eDRREs were not enriched in the other DRREs (Fig. 58). Altogether, novel-eDRREs appear to be under a regulatory program triggered by TFs different from that controlling reused-eDRREs and iDRREs.



**Figure 58 - Motif enrichment in DRREs.** (A) Heatmap showing top 10 enriched TF motifs in each DRRE type. Upregulated TFs are marked in bold. The level of significance is denoted with stars (\*\*\*)  $p < 10^{-3}$ .



**Chapter III:**  
**Conservation of the regeneration  
regulatory logic across metazoans**



Because regeneration is widespread in nature (Tanaka and Reddien 2011; Vriza et al. 2014; Chen and Poss 2017), we investigated the conservation of genes and regulatory regions triggering regeneration to determine whether there is a core molecular toolkit underlying organ regeneration in metazoans. We performed a comparative study with zebrafish heart and mammalian liver, two organs with capacity to regenerate in adult organisms, and for which transcriptomics data similar to the one produced in this thesis are available.

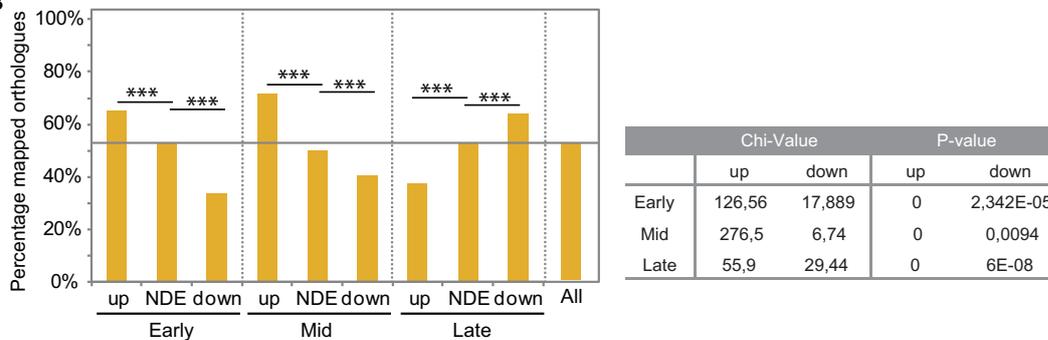
### High homology in the genes involved in fly regeneration

First we identified the fly protein-coding genes that have orthologs in at least one of these two species (7,458 genes, 54% of all fly genes) (Fig. 59A). We found that fly genes upregulated at early stages after injury have more vertebrate orthologues than downregulated genes and genes overall. For instance, at the early stage, 65% of upregulated genes have vertebrate orthologs, while only 35% of downregulated genes have them (Fig. 59A,B).

**A**

	Early			Mid			Late		
	up	NDE	down	up	NDE	down	up	NDE	down
Mapped in both species	1269	5880	54	1154	5972	77	227	6625	351
Only mapped in zebrafish	19	133	2	23	128	3	5	141	8
Only mapped in mouse	18	83	0	14	86	1	5	92	4
non-mapped	691	5667	104	440	5909	113	380	5891	191
total	1997	11763	160	1631	12095	194	617	12749	554
% mapped	65,3980971	51,8235144	35	73,022685	51,1451013	41,7525773	38,411669	53,7924543	65,523466

**B**

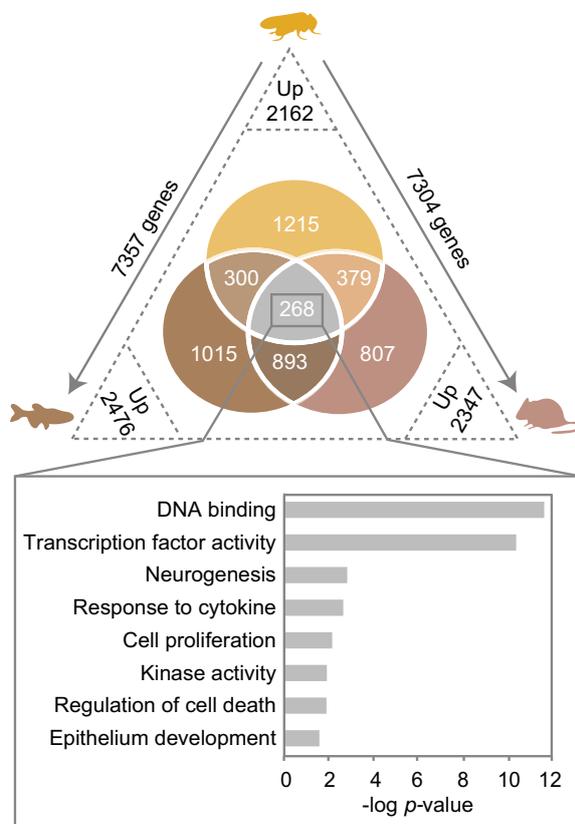


**Figure 59 - Homology of fly regenerative genes.** (A) Table showing the number and percentage of DE fly genes that map to an ortholog in zebrafish, mouse or in both. (B) Bar plot showing the percentage of fly genes with an ortholog in zebrafish, mouse or both at each time point and for each set of genes (upregulated, non differentially expressed or NDE and downregulated). Such percentage for all fly protein-coding genes is also shown and highlighted as a horizontal line. Also it is shown the Chi-value and the *p-value* of the comparison between NDE and upregulated or downregulated genes in each time point.

## A conserved core of regeneration genes highly related to transcription

Then, we analyzed RNA-seq data recently produced during heart regeneration in zebrafish (Goldman et al. 2017) and liver regeneration in mouse (Sun et al. 2016).

We identified the genes that were upregulated after injury in zebrafish and in mouse using the same bioinformatics protocol that we employed in fly. Then, we performed a fly-oriented way comparison between the three transcriptomes, which means that to identify regeneration core genes, we selected fly upregulated genes (early and/or mid) mapping to at least one upregulated ortholog in each dataset (See Annex II for mapping statistic of shared regenerative genes). We obtained 2,476 fly genes with at least one ortholog upregulated in zebrafish regeneration, and 2,347 fly genes with at least one ortholog upregulated in mouse regeneration.



**Figure 60 - Conservation of regenerative genes.**

Venn diagram showing the intersection of upregulated genes that have an orthologous gene in a fly oriented way. Numbers in the axis outside the triangle indicate the amount of fly genes that map an orthologous gene in zebrafish and mouse. Numbers in vertices show the amount of upregulated orthologous genes in regeneration and in each species. The set of upregulated shared genes between the three species is highlighted and the bar plot shows GO term enrichment in the same.

When we compared the three gene sets we found 268 genes shared by all species. We performed GO analysis in these genes and found enrichment of GO terms similar to the ones obtained when considering fly regeneration genes alone (Fig. 60). The highest enrichment occurs in terms related to DNA binding and transcription factor activity (Fig. 60), and, indeed, there is a set of 21 TFs shared genes by all the species (Table. 7A, B). Remarkably, among the 21 TFs shared by the three species, 43% of them had an enriched motif in DRREs (Table. 7C). Furthermore, within the set there are the effectors of the signalling pathways found to be enriched in the transcriptomic analysis: *Stat92E*, *lilli*, *Dif*, *sd* and *Jra* from Jak-STAT, MAPK, Toll, Hippo and JNK pathways respectively.

A

			
<i>lilli</i> CG13775 pdm2 elB <b>Dif</b> Jra cyc Su(z)12 Rbf2 <b>sr</b> <b>Stat92E</b> <b>E2f1</b> <b>zfh1</b> Sox100B	Med Rbf bi <b>CHES-1-like</b> <b>sd</b> <b>Myb</b> <b>disco-r</b>	<b>drm</b> vri E2f2 Optix sug CG6701 twi <b>pita</b> CG7839 <b>caup</b> Eip75B kin17 ato hon	C15 dys CG2120 <b>Her</b>
aop CG17612 CG7099 Clamp d4 Br140 <b>vis</b> Psc mip120 row Dek <b>CG9890</b> CG15011 <b>fd68A</b>	ash1 Gnf1 <b>MTA1-like</b> <b>Kdm2</b> <b>svp</b> MBD-R2 <b>pnr</b> woc Ets98B <b>wdn</b> yem <b>CG12054</b>	<b>cbt</b> fu2 YL-1 abo esc CG6686 <b>crp</b> az2 CG1603 <b>CG1602</b> <b>dpn</b> <b>CG1663</b> Sox15 Blos1	<b>CG8089</b> <b>CG10543</b> Hmg-2 <b>CG10321</b> NC2alpha <b>ken</b> <b>CG11414</b> CG2790 hng3 ecd ERR Gua ac sc crm Mnt
Tip60 CG3815 Nf-YC CG6769 <b>CG2116</b> pad CG3407			
21	18	26	47

B

	Shared all	Fly Zebrafish	Fly Mouse	Fly
All genes	268	300	379	1168
TFs	21	18	26	47
%TFs	7,84	6	6,86	4,02

D

Shared effector TFs		
Dif	→	Toll pathway
Stat92-E	→	Jak-STAT pathway
Sd	→	Hippo Pathway
Lilli	→	MAPK pathway
Jra	→	JNK pathway

C

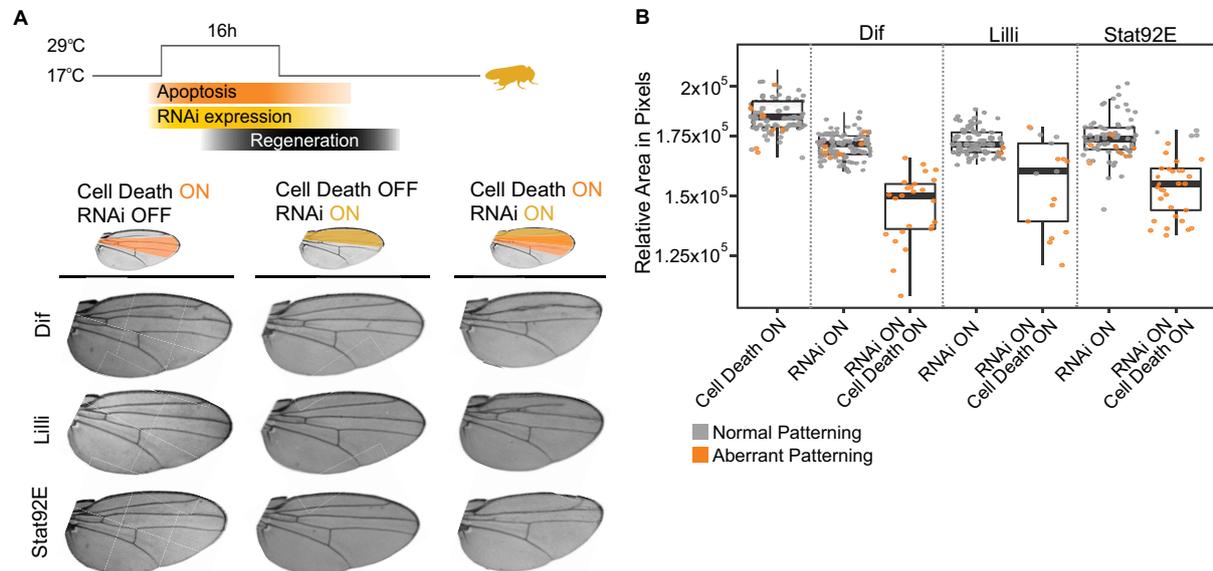
	Shared all	Fly Zebrafish	Fly Mouse	Fly
TFs	21	18	26	47
TFs with an enriched motif	9	4	9	16
% TFs with an enriched motif	42,86	22,22	34,61	34,04

**Table 7 - List of conserved TFs upregulated in regeneration.** (A) Table showing the list of mapped and upregulated TFs in fly. TFs are classified depending if they are upregulated in all the species, in two or uniquely in fly. (B) Table showing the number of TFs over the total number of genes in each of the classes described in panel A. (C) Table showing the number of TFs containing an enriched motif in fly DRREs over all the TFs in each of the classes described in panel A.

We selected *Dif*, upregulated at the early and late time points, *lilli*, upregulated at the early time point and *Stat92E*, upregulated at the mid time point to study their requirement by testing the capability to regenerate upon depletion of such TFs. For that, we used the double transcriptional transactivator system consisting of the *sa<sup>EPV</sup>-LHG lexO-rpr* to induce apoptosis combined with the *UAS/Gal4* and *Gal80<sup>ts</sup>* systems to drive the expression of the RNAi against the TFs in the anterior domain (*cubitus interruptus*, *ci-Gal4*) of the wing disc.

As previously described (Bergantiños et al. 2010; Repiso et al. 2013; Santabárbara-Ruiz et al. 2015), we found that the induction of cell death resulted in properly regenerated adult wings (Fig. 61A,B). The expression of each RNAi without cell death induction had a minor effect in the adult wings, characterized by a slight decrease in the area and by a 10 % of wings with a minor patterning defect. However, the combination of each individual RNAi with genetic ablation resulted in impaired regeneration, characterized by a major reduction of the wing area

as well as in an increase in the number of wings with an aberrant vein patterning (100%, 70% and 90% for *Dif*, *Lilli* and *Stat92E* respectively). These results demonstrate that these TFs are not only upregulated but also required for proper regeneration regeneration.

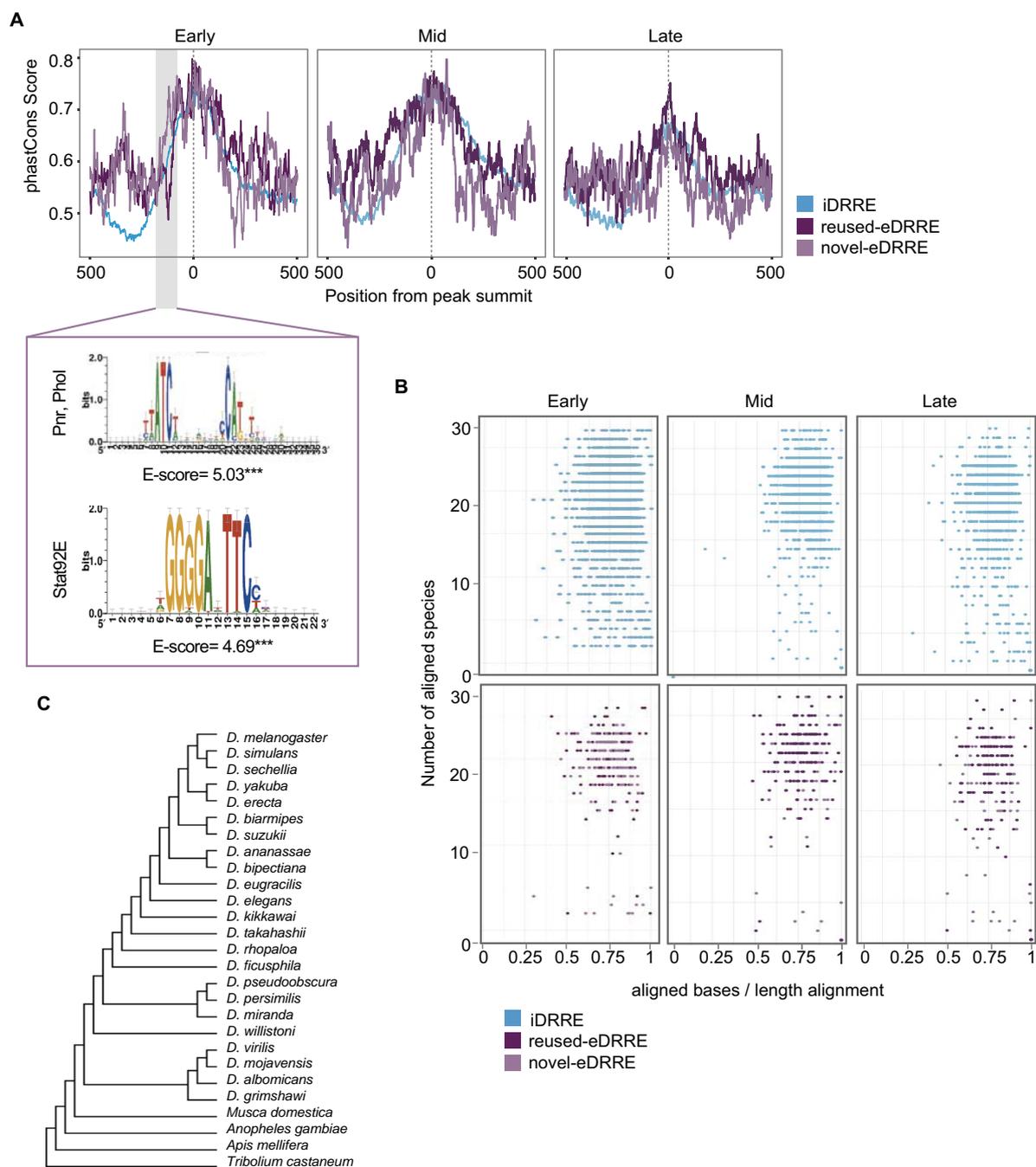


**Figure 61 - Requirement of shared TFs in fly regeneration.** (A) Schematic of the experimental design used to score the ability to regenerate upon the depletion of a gene by RNAi and the induction of cell death (top). Adult wings showing the predominant phenotype observed upon each condition (bottom). The region where the RNAi was expressed is highlighted in yellow and the apoptotic region in orange. (B) Boxplot showing the average area of adult wings obtained after the expression of cell death, the RNAi or the combination of both. Each dot represents one wing. Wild type pattern (gray) and aberrant pattern (orange).

## The regulome of regeneration is evolutionary conserved

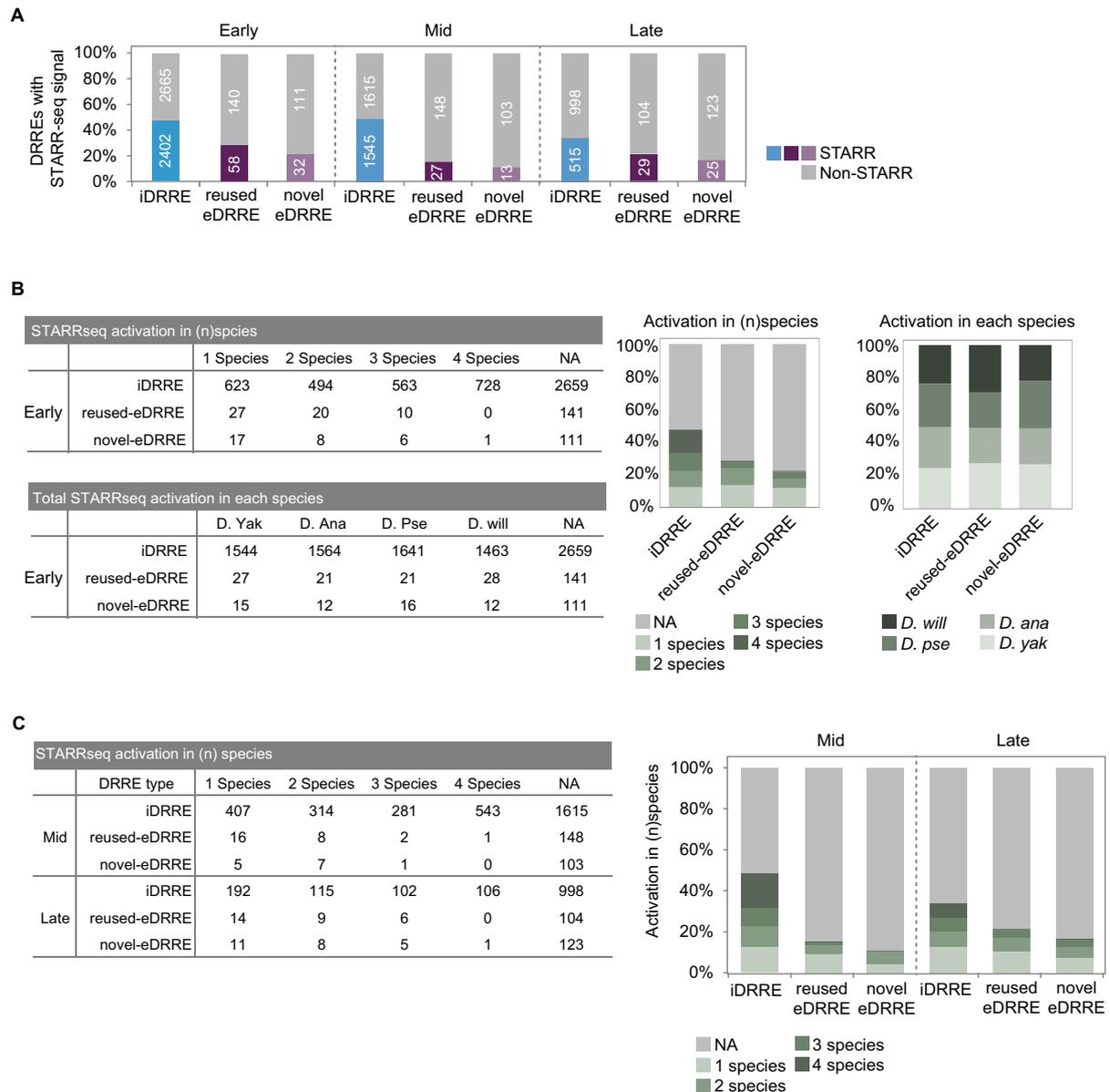
As enhancer activity is often deeply conserved in insects (Stark et al. 2007; He et al. 2011; Arnold et al. 2014), we explored whether DRREs are present in other insect species. We first calculated the average phastCons score for each DRRE using 27 different insect species, and found a similar pattern of conservation among DRRE types, with the exception being that iDRRE are less conserved in the nucleosome region upstream of the peak summit of the NF region (Fig. 62A). We also observed that novel-eDRREs at the early time point present higher conservation around 100bp upstream of the peak summit. When we applied motif discovery to these regions, we found enrichment for *Stat92E*, *pannier (pnr)* and *pleiohomeotic like (phol)*, which correspond to the most enriched motifs for novel-eDRREs (Fig. 62A).

Next, we calculated the number of species containing DRREs and found that most eDRREs are present in a large number of species, while iDRREs tend to be more species specific (Fig. 62B,C), suggesting that eDRREs might be involved in the core regulation of regeneration pathways common to all insects.



**Figure 62 - Conservation of DRREs.** (A) Average distribution of PhastCons scores derived from 27 insect species in the DRRE sequences (defined as 500 bp upstream and downstream of the NF peak summit). Highlighted a more conserved region at -100bp from the peak summit in novel-eDRREs with its motif enrichment. The level of significance is denoted with stars (\*\*\*)  $p < 10^{-3}$ . (B) Conservation of DRREs across 27 insect species. Each dot corresponds to one independent enhancer. Y-axis denotes the number of species that present the enhancer conserved. X-axis represents the percentage of aligned bases per sequence length. (C) Phylogenetic tree showing the 27 insect species used for conservation analysis.

Previous work has already demonstrated the activity of several of these conserved enhancers among different *Drosophila* species using the STARR-seq technique (Arnold et al. 2014). Taking advantage of these data, we found that around 40% of iDRREs are active in other *Drosophila* species, whereas enhancer activity has been proven for only 20% of novel-eDRREs (Fig. 63A, B), which is consistent with these enhancers being activated only after damage.

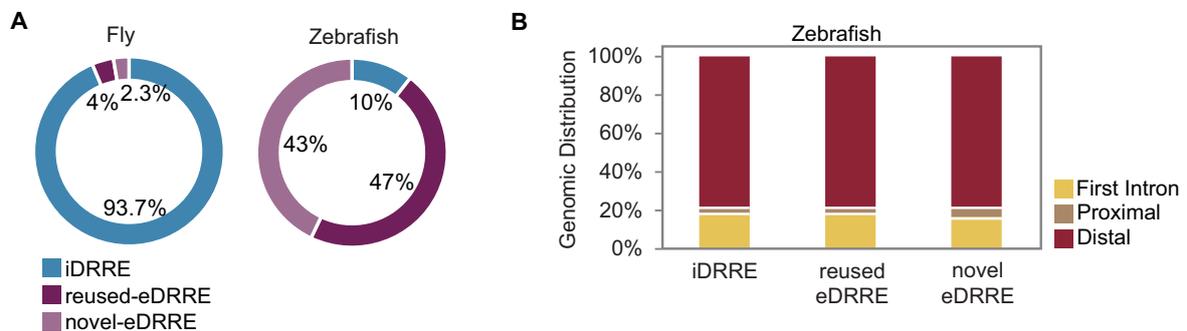


**Figure 63 - Activation of DRREs in other species.** (A) Percentage of conserved DRREs that are active according to the STARR-seq technique. (B) Tables showing the number of species containing the same active enhancer (top) and the number of active enhancers present in each species (bottom). Bar plots showing the same numbers in percentage (left). Numbers refer to early stage. (C) Tables showing the number of species containing the same active enhancer (right). Bar plots showing the same numbers in percentage (left). Numbers refer to mid and late stages.

## The fly regulatory logic is conserved in zebrafish regeneration

We next investigated to what extent the pattern of emergence and reuse of DRREs that we found in fly is also present in other systems. We took advantage of the only available data on open chromatin in regeneration, which is the histone profiling of H3.3 from the zebrafish heart regeneration experiment used above in the transcriptomic analysis (Goldman et al. 2017). We followed the same bioinformatics procedure that we used in fly to obtain DRREs in the zebrafish regenerating heart. In agreement with published results (Goldman et al. 2017) we found that the majority of DRRE in zebrafish are emerging, in contrast to the fly, where the great majority of DRRE are increasing (Fig. 64A).

To investigate the possible reuse of eDRREs we analyzed open chromatin data obtained during zebrafish embryonic development (Gehrke et al. 2015). As observed in the fly, we found that a fraction of open regions classified as *emerging* had already been identified as enhancers in embryos: reused-eDRRE (47% of all DRREs, Fig. 64A). In contrast to the fly, where we observed that novel-DRREs tend to be located in more distal regions than all DRREs, we did not observe any difference in the genomic distribution of DRREs in the zebrafish (Fig. 64B). Nonetheless, considering that the great majority of DRREs in zebrafish are *emerging* and that we only take into account one embryonic stage, the percentage of reused-eDRRE in zebrafish is considerably higher than in the fly.



**Figure 64 - Reusage in zebrafish.** (A) Classification of DRREs identified in fly and zebrafish in novel-eDRRE, reused-eDRRE and iDRRE based on co-option from embryo stages. (B) Bar plot showing genomic distribution of the three types of DRREs identified in zebrafish.



# DISCUSSION



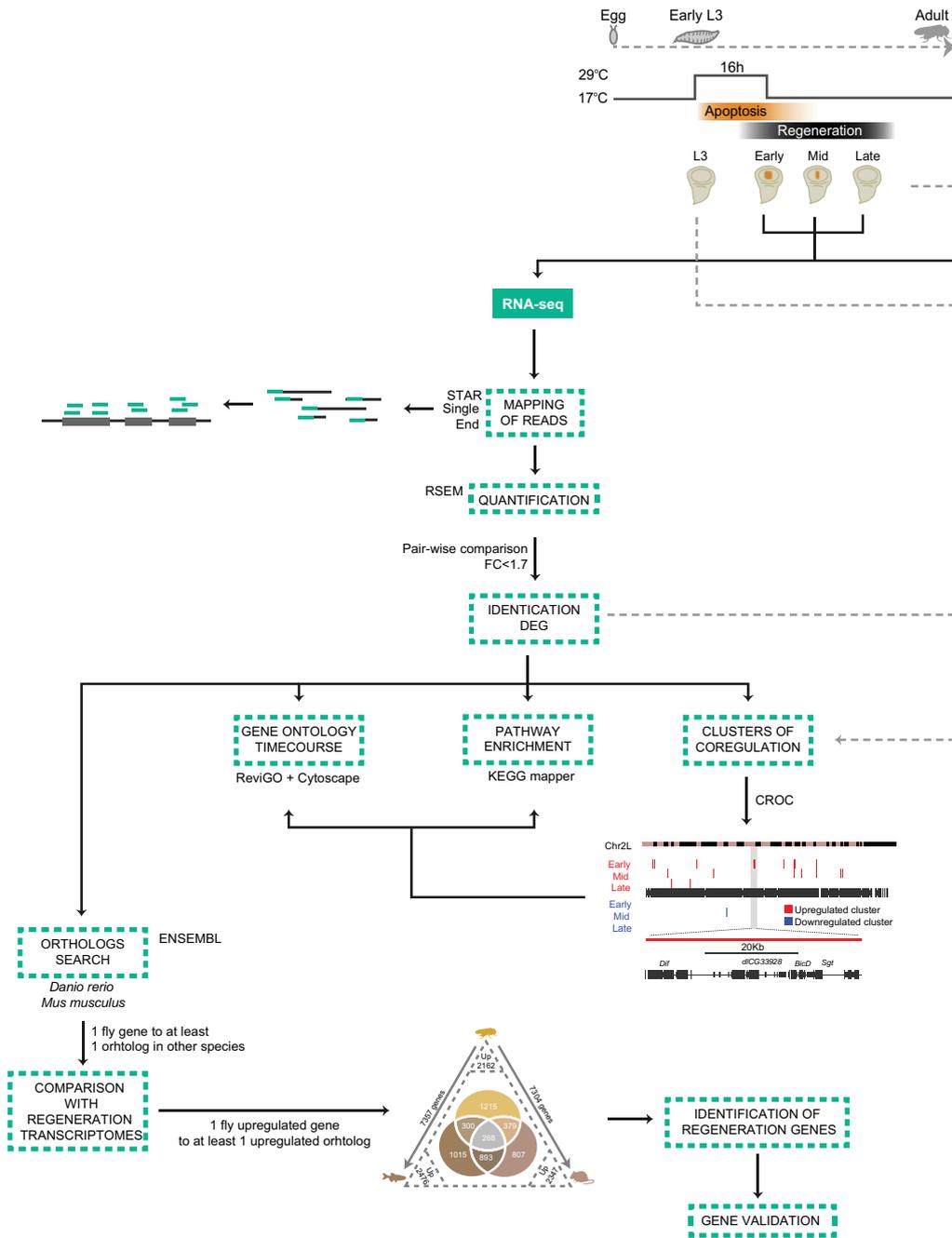


Resetting of gene expression patterns during response to injury is essential for regeneration. During tissue and organ regeneration, certain cells detect damage and switch their transcriptional programs to reconstruct lost structures. This process involves spatial and temporal regulation of gene expression (Maurange et al. 2006; Katsuyama et al. 2015; Kang et al. 2016; Rodius et al. 2016). In this work, we have identified the genes and regulatory elements involved in regeneration of fly imaginal discs. Our results point to a dynamic pattern of gene expression, which is controlled by different types of regulatory elements, including novel enhancers activated only after damage acting in combination with enhancers re-used from other tissues and stages.

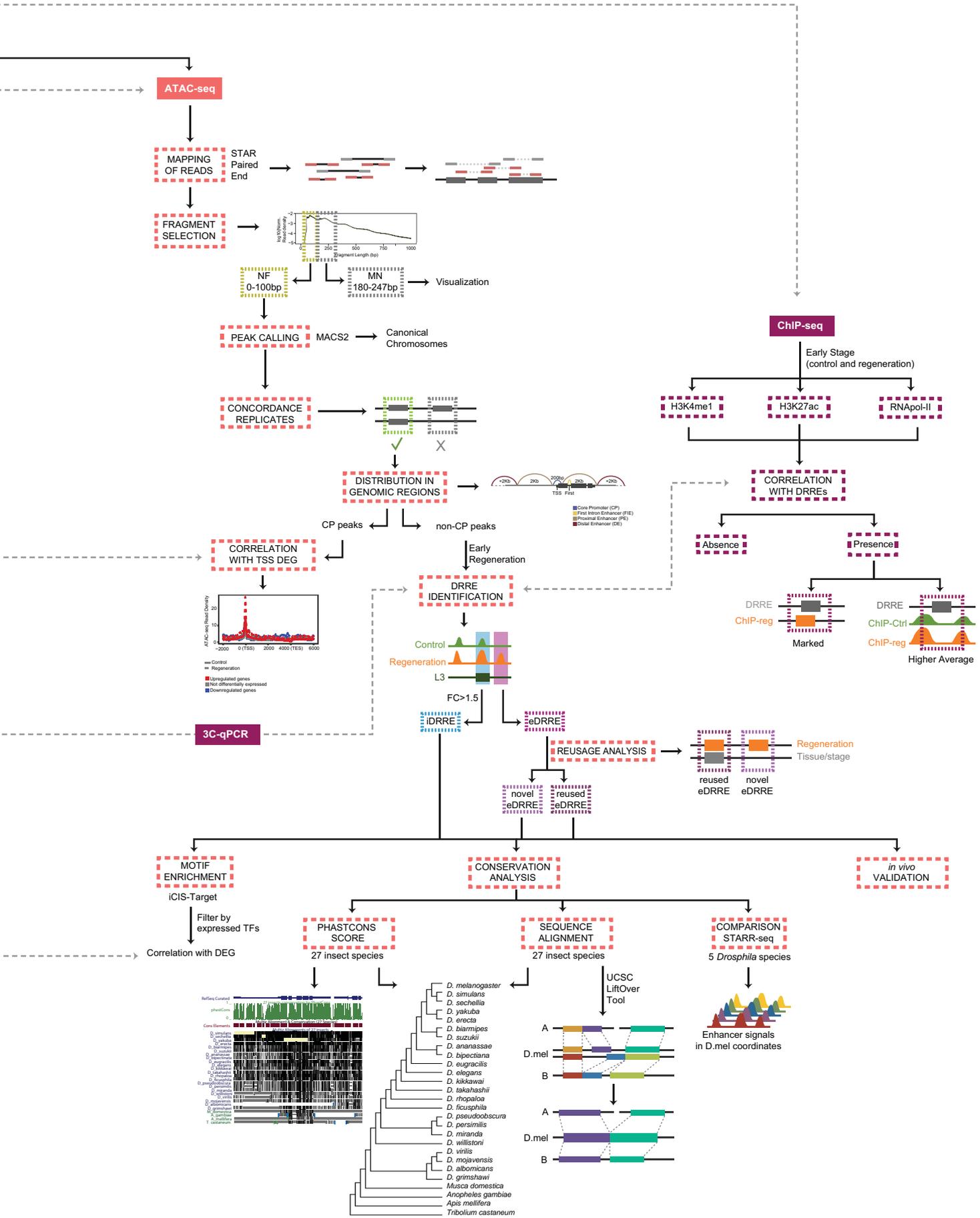
## **A GENOME-WIDE WORKFLOW TO STUDY REGENERATION**

A complete understanding of the regulatory landscape of regeneration requires, among others, the discovery of DE genes and DRRE regions across metazoans. Nowadays, this is possible thanks to genome-wide techniques. Although such techniques allow us to acquire countless datasets, they have to be properly interpreted to extract as many results as possible. Indeed, one of the most challenging works in this thesis has been the design of a workflow that could take into account the highest number of variables, integrated into an easy to follow pipeline. Thus, we have created a complete overview of the used workflow. Its logic can now be applied to any other model used in regeneration (Fig. 64). Besides this, and to make all the generated data available and easy to access, we have created a UCSC track data hub: [https://public-docs.crg.es/rguigo/Papers/2018\\_vizcaya-klein\\_regeneration/hub.txt](https://public-docs.crg.es/rguigo/Papers/2018_vizcaya-klein_regeneration/hub.txt)

Moreover, the vast amount of data generated can also be used as a resource for other researchers performing developmental or comparative genomics studies.



**Figure 64 - Genome-wide workflow used to study regeneration.** Workflow depicting all acquired data, bioinformatics processing, and the steps followed to analyze and integrate other datasets.



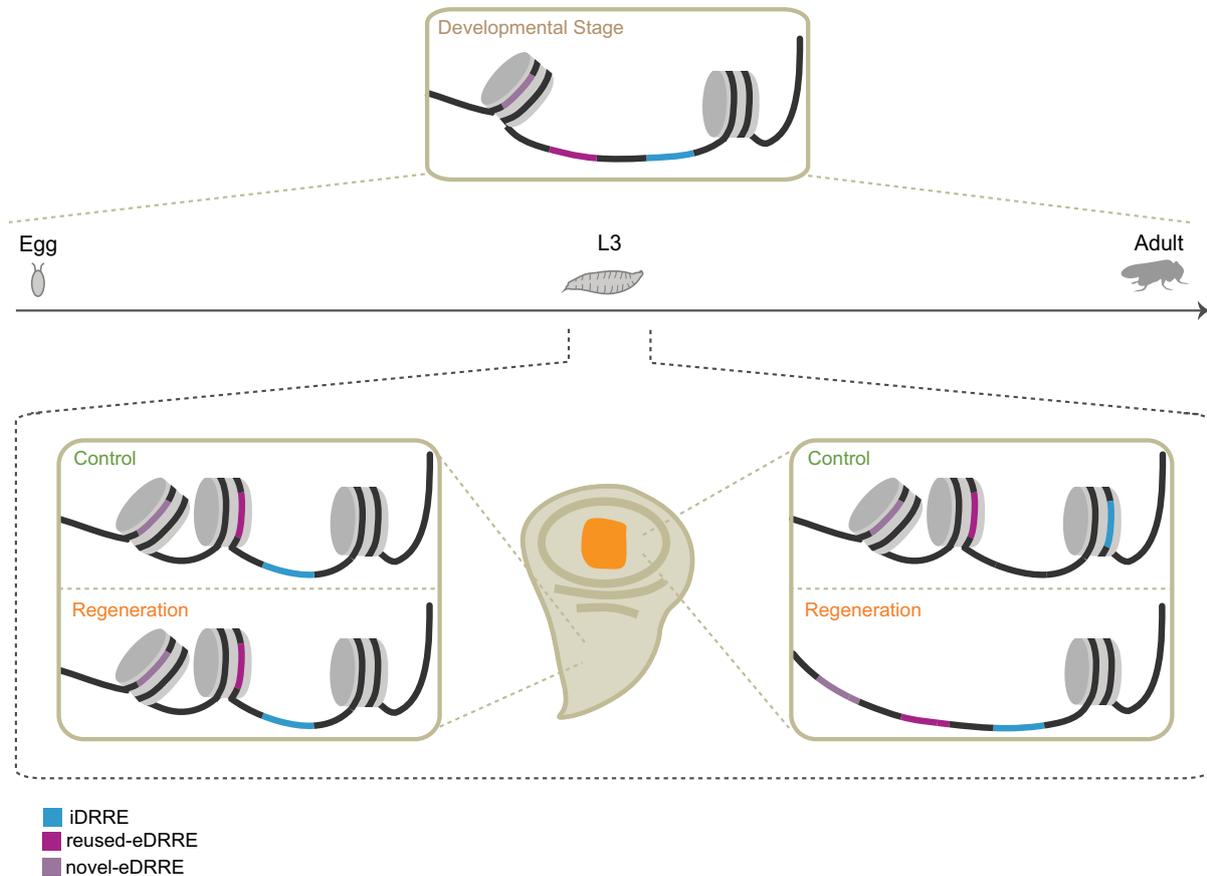
## INTEGRATING THE ONSET OF REGENERATION SIGNALS INTO THE GENOME: DRRES

Early signals that initiate regeneration in different species involve calcium waves and the production of ROS (Razzell et al. 2013; Santabárbara-Ruiz et al. 2015; Niethammer 2016; Hariharan and Serras 2017). The early burst of ROS activates JNK and p38 MAPK pathways, which in turn activate the Jak-STAT pathway (Gauron et al. 2013; Santabárbara-Ruiz et al. 2015; Fogarty et al. 2016). The final outcome of these signals is their integration into the genome, through specific regions, that become accessible after sensing damage, as **Damage Responsive Regulatory Elements (DRREs)**. We have classified these elements in three different types, according to their accessibility: iDRRE, reused-eDRRE and novel-eDRRE (Fig. 65).

Among the different DRREs, **increasing DRREs** are in regions that are already open in the wing imaginal disc, but that become **more accessible** during regeneration, indicating a fine-tuning mechanism of gene expression, as it occurs with the WNT damage enhancer (Harris et al. 2016). Being already opened in some cells of the wing disc, does not necessarily mean that such regions are already acting as enhancers in the cells responding to damage. In fact, analysis using reporter lines showed that compared to their basal activation pattern, some iDRREs are ectopically activated in the wound. **Reused-eDRREs** are **co-opted** from other developmental stages or tissues and are reused in regeneration. The existence of elements used during development and reused in injury was previously proposed for zebrafish regeneration (Reviewed in Chen and Poss 2017). During regeneration, cells have to reespecificate and proliferate to replace lost tissue, thereby, the tissue needs to somehow recapitulate some developmental traits. Our work confirms that the same regions can participate in development as well as in regeneration, both in flies and zebrafish. The last fraction of enhancers corresponds to **novel-eDRREs** that act exclusively in the damaged tissue. This last category could, in theory, represent **genuine regeneration enhancers**. The leptin B enhancer, found in zebrafish regeneration, seems to be such an enhancers type: it has been proven that it plays a crucial role in regeneration, but it does not seem to be required at all neither during development nor in basal heart homeostasis (Kang et al. 2016). Further comparative analysis with more tissues and stages will allow to refine the occurrence of the different enhancer types.

Moreover, we have seen, that DRREs are **highly conserved** through 27 different **insect species**, regardless of the DRRE type. Even if it remains to be determined which species are able to regenerate, many live wild in nature and suffer insults that may alter cell homeostasis and activate damage responses. Therefore, DRREs could have been **positively selected** to allow **fly survival** upon environmental aggressions. In this regard, the regeneration capability and the degree of enhancer conservation could be directly linked. Genome-wide comparison of close species with different regeneration capabilities would help to highlight the difference between good and bad regenerators.

Altogether, it seems, that a complete regenerative response requires the combinatorial effect of all enhancer classes. Our results suggest a gene regulatory program triggered by different types of DRREs, acting either on individual genes or on clusters of co-regulated genes.



**Figure 65 - DRRE types.** Illustration showing the three DRRE types and their accessibility patterns in control, regeneration, and other tissues and stages. The model does not reflect a real situation, meaning that enhancers are not necessarily located in the depicted genomic distribution.

## TOWARDS AN ACTIVE REGULATORY LANDSCAPE

One of the most fascinating properties of enhancers is their functional autonomy, meaning the ability to retain their transcription-activating function outside their endogenous locus (Banerji et al. 1981). Taking advantage of their functional autonomy, we have proven DRRE activation in the cells surrounding the wound; however, it is still a challenge to understand how enhancer activation occurs.

Accumulating evidence from many model systems indicates, that the **combinatorial interplay** of multiple TFs, each with its own partially overlapping temporal window of expression, is a prominent regulator of context-specific binding. These variations in the spectrum of TFs generate different target gene expression profiles in different cells. Hence, TFs can induce

induce alterations and enable patterning through development in the same cell type (Reviewed in: Spitz and Furlong 2012; Zabidi and Stark 2016). Accordingly, we have found that different DRREs present different enrichment for motifs of TFs that are not only expressed but upregulated in regeneration. Such TFs are located downstream of signaling pathways, required for organ regeneration in fly, zebrafish, and mouse (i.e. Stat92E, Sd and Myb) (Sun et al. 2016; Goldman et al. 2017). The existence of regions capable of responding to damage or stress that become active in the presence of specific TFs and are shared across organ regeneration in different species, may have profound implications for our understanding of tissue regeneration. Still, one key question remains: Is DRREs accessibility caused by TF binding or is TFs binding a consequence of being already accessible? Further experiments in this regard will be required to disentangle causes and consequences. For instance, comparative analysis with ATAC-seq (or even single ATAC-qPCR) in regeneration, upon depletion of TFs would allow to understand if TF presence itself is enough to open chromatin or not.

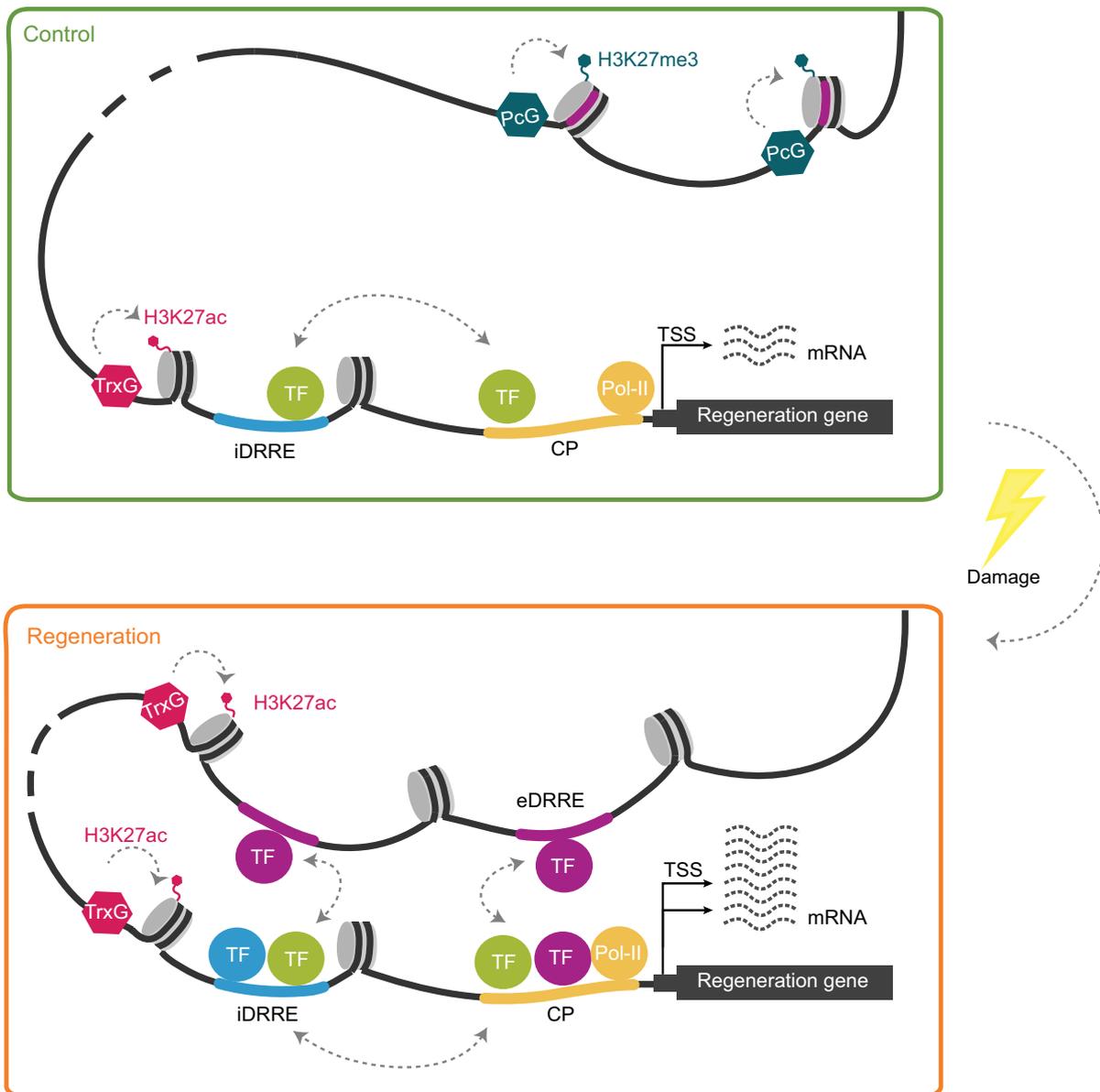
Additionally, we have found that DRREs are enriched with histone marks associated to activated enhancers, yet almost half of all DRREs remained unmarked. Indeed, amongst the validated DRREs, one eDRRE (GMR24G07) and one iDRRE (GMR17D09), did not present any type of active features, even though they are precisely active surrounding the damaged domain. Furthermore, although previous studies have pointed to a role for chromatin modifying factors in regeneration, none of them has demonstrated the requirement of activation marks, but the need of losing repressive ones. Shawn and Martin demonstrated, that H3K27-demethylases are required in blastema cells of injured mice skin to counteract PcG (Shaw and Martin 2009). This has also been observed by Stewart and colleagues, but in this case, to change the balance of bivalent promoters towards an active state upon injury in zebrafish (Stewart et al. 2009). Besides, studies in flies described, that JNK signalling weakens PcG silencing after cell death induction (Lee et al. 2005) and that regeneration ability is lost upon maturation, in part through silencing mediated by the PcG of the WNT enhancer (Harris et al. 2017). In agreement with these findings, we have found that eDRREs are in closed chromatin enriched with H3K27me3 in wild type L3 discs, which is lost upon injury. Altogether, it seems that the regulatory landscape needs to lose the silent default state, carried out by PcG. Still, we are unable to discriminate, whether there is or not a direct relationship between chromatin modifying factors and DRREs.

Furthermore, an active chromatin landscape is not only defined by the opening and activation of enhancers, but also via enhancer mediated transcription through chromatin loops (Reviewed in Catarino and Stark 2018). Recently, it has been described that the 3D genome in *Drosophila* is organized in small 3D-compartments: the active (A) and repressive (B) ones. A compartments present high levels of transcriptional activity, chromatin accessibility, active histone modifications, and chromatin looping, whereas B compartments present exactly the opposite (Rowley et al. 2017). In the same study, it was observed, that domains are directly defined by transcription activity and that they can be shifted easily upon transcriptional

perturbations. Additionally, it has also been demonstrated, that enhancer-promoter loops in *Drosophila* development can be preset, even before gene activation occurs (Ghavi-Helm et al. 2014). It is difficult to discern the general trend in regeneration without having access to HiC data (Chromosome Conformation Capture followed by high throughput sequencing). Based on the results obtained here, we suggest that during regeneration a combination of both previously mentioned mechanisms occurs. It seems, that the activation of the transcriptional state produced by damage events, switches the balance between A and the B compartments towards the A one. This contributes to enhance the occurrence of preset loops between DRREs.

Besides the physical constraints imposed by the genome architecture, enhancers cannot regulate all promoters indiscriminately. Enhancers and promoters need to be biochemically compatible: the specificity in contacts is encoded in enhancer sequences and depends on specific TF motifs. Such motifs recruit different TFs and cofactors to activate different promoters (Zabidi et al. 2015; Zabidi and Stark 2016). For instance, the DRE motif is essential for the enhancer-CP function in housekeeping enhancers. As a consequence, it has been proposed that occupancy of either Dref and Boundary element-associated factor of 32kD (BEAF-32) is a key contributor to enhancer-promoter specificity (Gurudatta et al. 2013; Zabidi et al. 2015; Cubeñas-Potts et al. 2017). Such motif is enriched in DRREs and in the CP of upregulated genes, and both, Dref and BEAF-32, are upregulated during regeneration. Hence, both genes could possibly mediate loops between DRREs and CPs of upregulated genes. Similarly, the Trl motif is found in both CP and DRREs as well. Trl has been recently described to be required in the formation of repressive chromatin loops mediated by PcG for gene silencing in *Drosophila* development (Ogiyama et al. 2018). However, Trl itself mediates transcription activation through enhancer-CP loops (Mahmoudi et al. 2002). Hence, the shift in the balance of the chromatin state could lead to a loss of PcG gene silencing and to transcription activation through Trl. Actually, it could be a putative mechanisms to explain how reused-DRREs are coopted, presenting the Trl motif as the most enriched binding site.

Combining all the obtained pieces of knowledge, both from previous studies and this work, we would like to propose the following model for DRRE activation: upon damage, TFs represent the bottom line for DRREs activation, assisted by the coordinated action of chromatin modifying enzymes, allowing a more flexible chromatin state through loss of the repressive chromatin state. This, would lead to a shift in the A/B compartmental balance, towards the A one, thus allowing chromatin rearrangements and the establishment of loops between DRREs and CPs to trigger gene expression (Fig. 66).



**Figure 66 - A model for enhancer activation.** Illustration showing an hypothetical model for DRRE activation; the model does not reflect a real situation, meaning that enhancers are not necessarily located in the presented genomic distribution. Upon regeneration, DRREs become accessible through direct binding of TFs or loss of repressive marks, and acquisition of active ones, also allowing TF binding. DRREs are then able to contact each other as well as regeneration gene CPs, to promote enhanced expression.

## A BURST OF ACTIVE TRANSCRIPTION

The ultimate goal of DRREs activation is to trigger the transcriptional outcome, that will drive regeneration. We have found, that the regeneration transcriptome is best described by **expression modulation** of already expressed genes, rather than initiation of gene transcription *de novo*. Actually, it has been previously described, that the ability to regenerate is based on the ability to modulate gene expression. For instance, in newt lens regeneration, RNA-seq results showed, that dorsal and ventral irides, which present different regeneration capabilities,

mostly differ in the amount of transcripts, rather than their uniqueness (Sousounis et al. 2013). Moreover, we have found that such gene modulation is a **burst of transcription**, that mainly occurs at the **early regeneration stage**, and is recovered with time. Other time-course transcriptomic analysis, performed in zebrafish heart regeneration, have also demonstrated a similar transcriptional behaviour over time (Rodius et al. 2016).

It is of note, that transcriptional bursting already occurs under homeostatic conditions. Transcription does not seem to be a constant process but occurs in waves, with bursts of transcription initiation separated by inactive intervals (Golding et al. 2005; Chubb et al. 2006; Raj et al. 2006; Dar et al. 2012). Interestingly, although such bursting is regulated by enhancers, the burst size and length is more likely determined by the core promoter sequence (Carey et al. 2013; Tantale et al. 2016; Catarino and Stark 2018). In regeneration, the burst of transcription is not only characterized by a higher number of upregulated genes, but also by the type of genes: **TFs** and **transcription related genes** are upregulated during the early stage. Indeed, the upregulation of just 5 TFs (*Jra*, *atf-3*, *Dref*, *grh* and *Trl*) could explain the upregulation of 50% of the early stage genes, based on the putative binding of such TFs to the CP of genes. Hence, these 5 TFs could **determine the strength of the burst** set by DRREs.

Besides, the occurrence of these precise TFs is not arbitrary. *Jra* is the effector of the JNK pathway, which has been extensively described to be rapidly activated after damage, and required during wound healing and regenerative growth (Bosch et al. 2005; Mattila et al. 2005; Lee et al. 2005; Bosch et al. 2008; Smith-Bolton et al. 2009; Bergantiños et al. 2010; Blanco et al. 2010). *Atf-3* is one of the effectors of the p-38 MAPK pathway, which at the moment, is the only pathway found to be activated in living cells in the wing disc after damage (Sanabábara-Ruiz et al. 2015). It is not known, which of the effectors of this pathway could be the one (or the ones) required in regeneration of the wing disc. Here, we propose *Atf-3* as a candidate for this role. On contrary, no previous role has been proposed for *Dref* in regeneration. As aforementioned, *Dref* is required in enhancer-promoter looping (Zabidi et al. 2015; Cubeñas-Potts et al. 2017), however other roles could be attributed to *Dref* based on its homeostatic functions. In previous studies, it has been demonstrated that *Dref* can physically interact with Akt and S6k (Vinayagam et al. 2014), which we have proven to be activated and required in regeneration. This suggests, that *Dref* could be possibly regulated by the mTOR pathway. Similar to *Dref*, many other functions can be attributed to *Trl* besides its role in chromatin looping (Mahmoudi et al. 2002, Ogiyama et al 2018). Among all functions, *Trl* drives the expression of growth genes in the wing disc, acting as a partner of *Yki* and *Cbt* (Oh et al. 2013; Ruiz-Romero et al. 2015). Both genes are not only upregulated but required in regeneration of the wing disc (Blanco et al. 2010; Sun and Irvine 2015). *Trl* also acts as a partner of *Sd* to promote growth (Bayarmagnai et al. 2012). Although *sd* has not been studied in regeneration in the wing disc, it has been described to be required in other models, such as in muscle regeneration in the mouse (Joshi et al. 2017). In the same study, it was postulated, that *Sd* could be precisely required for enhancer activation. Indeed, we have found the *Sd*

motif enriched in DRREs. Hence, Trl could be acting together with Yki, Cbt, and Sd through DRREs to modulate expression of growth genes. Finally, **grh** is a key transcription factor responsible for epidermal barrier formation and for epidermal wound repair in the fly embryo and in mouse (Ting et al. 2005; Mace et al. 2005; Caddy et al. 2010). It has been described, that Grh action in regeneration is ERK dependent and controls growth and proliferation upon its phosphorylation (Kim and McGinnis 2011). Thus, we propose a similar function in wing disc regeneration as well.

Although this study is only pointing to the putative role of the top 5 TFs bound to the CP, we have found a set of 195 TFs, which are likely contributing to regeneration. Further experiments are needed to decipher, firstly, if they are required in regeneration, and secondly, to highlight their function in regeneration.

## THE CORE SET OF REGENERATION GENES

Conservation analysis have enabled us to discover, that genes implicated in fly regeneration present higher levels of homology with humans, mice, and zebrafish compared to fly genes overall. This pinpoints the relevance of gene regulatory networks required for regeneration. Besides, the **comparative transcriptomic analysis** has helped to highlight the conserved core of genes participating in the process.

The transcriptomes selected for this comparative study belong to three different regeneration models (zebrafish heart, mouse liver, and fly imaginal disc) that undergo the same regeneration type based on the level of biological organization: **organ regeneration**. The mechanisms used to sense the damage in the three models are similar (Reviewed in Hariharan and Serras 2016; González-Rosa et al. 2017; López-Luque and Fabregat 2018), however, they use **different mechanisms** to achieve restoration of the damaged organ. In the wing imaginal disc, transdifferentiation events drive intercalary growth (Bryant et al. 1981; Repiso et al. 2013). In the heart, dedifferentiation events and stem cells seem to play a major role in replacement of lost heart mass (Jopling et al. 2010; Sánchez-Iranzo et al. 2018). In the mouse liver, the remaining hepatocytes undergo hyperplasia and hypertrophy, which account as compensatory growth to recover the liver mass (Reviewed in Michalopoulos and DeFrances, 1997). One could imagine, that the three models transcriptionally modulate the expression of three different sets of genes. Despite, there is a set of **common genes**, which is enriched in **transcription related** genes. Among them, there is a set of 21 TFs that is conserved across the three species, which could constitute the **core set of regeneration genes**. Moreover, we have found the effectors of signalling pathways required in the onset of regeneration, as well as other pathway operators, among the core genes, indicating that a more sustained activity is necessary to achieve regeneration.

We selected three genes from the core set and demonstrated, that they are not only

upregulated but also required to regenerate. **STAT92E** is the effector of the Jak-STAT pathway. Its requirement has been widely studied in the three species. In mouse liver, for instance, STAT3 (the mammalian ortholog of STAT92E) is required in the priming phase of regeneration, where hepatocytes resting in proliferative quiescence re-enter into cell cycle (Reviewed in Cienfuegos et al. 2014). **Dif** has not previously been studied in regeneration of the fly, however the activation of its orthologs (the Rel family of TFs) has been demonstrated to be essential in zebrafish heart regeneration. Rel activation is in charge of NF-κB signaling (Karra et al. 2015). It is of note, that this pathway has been proposed to modulate the fetal reprogramming of cardiomyocytes, required in regeneration in the mouse heart (Maier et al. 2012). Although we have found the Dif motif enriched in iDRRE, we cannot discard, that it could play a role in reused enhancers of other species. Finally, in contrast to STAT92E and Dif, no previous role in regeneration has been attributed to **lilli** in any of the three species, neither in development of mouse and zebrafish. In fly development, lilli acts downstream of the FGFR pathway in cytoskeleton regulation, segmentation, and morphogenesis (Tang et al. 2001; Zhu et al. 2005). Hence, **lilli** represents a good candidate for further experiments in the three species.

Deeper analysis in signalling pathway genes would help to shed light over their specific roles in regeneration, and the processes each of them triggers. For instance, ChIP-seq analysis of the effectors of the pathways, combined with ATAC-seq data already obtained, would help to decipher which DRREs are under the control of each pathway. In a similar way, RNA-seq analysis in regenerating cells, upon depletion of effectors, will shed light on the genetic network, transcriptionally controlled by each pathway.

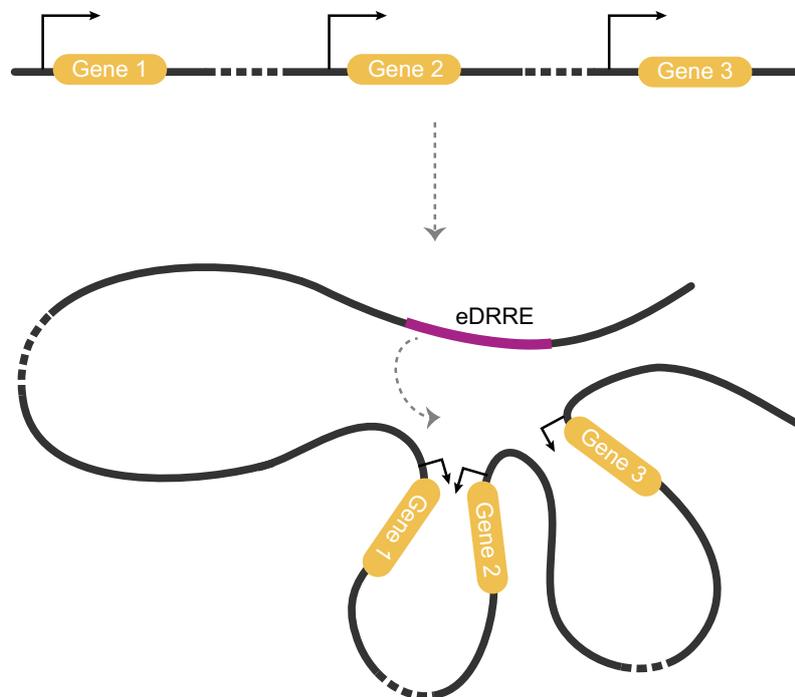
## MAKING THE SYSTEM EFFICIENT: GENE CO-REGULATION

Previous studies have described, that putative spatial clustering between co-expressed genes is compatible with the view that transcription proceeds in “factories” (Rieder et al. 2012). This is reinforced by a recent study in *Drosophila*, showing, that multiple active genes coalesce into compact structures, where transcription is more efficient (Corrales et al. 2017). We have found, that **genes upregulated** during fly regeneration can be located close to one another in the linear genome, creating genomic **clusters of co-regulation**.

These clusters are enriched in genes operating in signaling pathways, with some clusters containing members from different pathways. Such distributions could represent an **efficient regulatory strategy**, meaning, that many genes required for the regenerative process are turned on at once, in bulk, by cluster co-regulation. It is not very likely that gene distribution on the genome is based on regeneration, however, evolution could have somehow favoured a genomic distribution in which genes, that participate in the same biological processes, are positioned close to each other on the genome. Thus, genes from different pathways, playing similar biological roles would be located in the same cluster. For instance, *mer*, *dome* and *cdc42* belong to different pathways, but the three of them are linked to cell proliferation and

growth (LaJeunesse et al. 1998; Buchon et al. 2010; Nakamura et al. 2017). It would be very interesting to investigate whether signaling genes, in general, tend to appear in random positions across the fly genome, or are spotted in precise locations.

Moreover, our results obtained using conformation capture experiments suggest, that global **cluster regulation** could be triggered by a **well-positioned enhancer element**. In this regard, we have found, that a specific eDRRE preferentially interacts with upregulated genes inside a cluster, even though other upregulated genes outside the cluster can be found in closer proximity to the respective eDRRE in the linear genome. Further experiments should be done to confirm if this result is somehow a general trait. First, the deletion of pivotal enhancers with techniques such as CRISPR-cas9 would allow to make sure if cluster regulation is triggered by DRREs. We expect, that basal gene expression remains after DRREs deletion, but the increase of expression after injury would no longer occur. Secondly, HiC analysis or 4C, with specific DRREs as a bait, would allow to recognize all the interactions between genes and enhancers and subsequently, clusters and enhancers. This type of analysis could help to understand, if the regulation of entire gene clusters through a single enhancer is a generalized phenomenon or not. Actually, the integration of a third dimension into our analysis would also help to gain insights into other gene co-regulation events. Since chromatin loops help to shorten genomic distances, genes could also be co-regulated by spatial proximity emerging through genome folding, creating active chromatin regeneration hubs (Fig. 67).



**Figure 67 - Chromatin regeneration hubs.** Illustration showing a chromatin regeneration hub. Spatial proximity generated by chromatin loops brings three genes, located in distal genomic locations in close contact, so they can be co-regulated by a single eDRRE.

## THE RELEVANCE OF COMPARATIVE GENOMICS IN REGENERATION

Thanks to comparative genomics, we have been able to identify common genes between species with regenerative skills, the conservation extent of DRREs through insect species, and the preservation degree of the regulatory logic of DRREs. The role of associative and comparative learning is to allow to discriminate similarities and differences between two given situations. Regeneration is a stepwise process in which failure of one step leads to failure of the whole process (Reviewed in Rhoel et al. 2018). Hence, it is also important to understand, how and when processes fail. Through the use of comparative genomics with additional species lacking regenerative skills, we could be one step closer to answer why some animals are able to regenerate, whereas others are not. What if we take an organism that does not regenerate, perform genome-wide experiments, and compare the results of both good and bad regenerators? What makes one organism stand out from a another one? Which are the genes and enhancers activated in the good regenerator, that are not longer activated in the bad one?

This kind of rational can be applied to many other questions related to the regeneration ability. Comparison between individuals that lose this ability upon maturation would also be helpful, to understand why is it lost. *Drosophila*, could be one of those examples. Again, wWhich are the genes and enhancers activated activated in the larvae upon injury, that are no longer activated in the adult? Using the same genomic approach, that we followed in larvae, but in the adult, we could compare both stages and highlight differences. In the case of zebrafish, the heart regeneration capability after infarction is outstanding, as well as the co-option rates obtained for DRREs. Almost 50% of emerging enhancers are used in the embryo and contribute to the restoration of a complete wild type heart after injury (Kang et al. 2016; Goldman et al. 2017). Thus, could reactivation of embryonic enhancers in the human heart, after infarction allow to prevent the formation of fibrotic tissue? The ability to reactivate silent developmental enhancers in mature structures or organs could be crucial for regenerative medicine. Ectopic activation of regeneration enhancers could potentially be exploited to stimulate the regenerative capacity of organs and tissues that, in principle, are not able to regenerate.

## FUTURE PERSPECTIVES: WHAT IS NEXT?

The future of the regeneration field will basically rely on the kind of research we are willing to do. From my point of view, nowadays, there are two main types of researchers, that will guide the next steps in regeneration research: researchers willing to artificially create and researchers willing to really reconstruct. Although these two approaches differ in many aspects, at the end, both strategies pursuit the same aim. I believe, that none of them will succeed by itself, and that the future is based on the union of both strategies.

In the field of researchers willing to create, biomedical engineers play the major role. With the occurrence of new biomaterials and bioengineering studies, regenerative medicine is

more focused on artificially creating lost structures, than forcing real regeneration. For instance, the Hartford Engineering a Limb (HEAL) project, aims to generate bioartificial functional human limbs by 2030. It is a reality, that, in the near future, advances in the technological field will be more helpful and easily applicable to treat patients, than activating regeneration itself.

This thesis and related works are focused on the research field that is willing to make real regeneration happen. However, this is a long-term perspective, that we do not even know, if we will be able to achieve. Hence, it is easier to think in short-term projects, which basically rely on understanding the regeneration process and try to make it more efficient. In my opinion, it is crucial to explore regeneration taking advantage of all the emerging techniques, that are lately appearing. In our case, trying to understand the regulatory genome, using highthroughput techniques, has allowed us to set the bases for countless new experiments, that for sure, will give rise to very exciting results. But not everything is about genomics. In vivo imaging techniques would tell us, for example, which cells are needed for regeneration and precisely when and where. Biophysics and computational modeling are making great advances in unravelling how structures will be restored, depending, for example, on gene regulatory networks. Drug screenings upon regeneration could highlight soluble factors and molecular cocktails, that make regeneration faster and more efficient. The model system used in this work could also be useful for such screenings, since flies are already being used as personalized avatars to design the best drug combination in cancer therapies. Maybe, *Drosophila* could be used to create regeneration-like avatars as well.

While strategies in the past decade have focused on potential therapies involving treatment with cell populations, tapping into natural regeneration programs and boosting the endogenous capacity of tissue to regenerate or rejuvenate is a prerogative for the field today. This is where the discovery science of developmental and regeneration biology and the application of regenerative medicine must meet and work together closely.

# CONCLUSIONS





The main conclusions of this work can be summarized as:

1. Upon damage, there is a burst of active transcription that mainly occurs at the early stage. Such burst is characterized by increased gene transcription and increased chromatin accessibility.
2. There is co-expression and co-regulation of genes involved in regeneration: upregulated genes in mid an early stage tend to be located close to one another in the linear genome. Cluster co-regulation could represent an efficient regulatory mechanism as genes can be turned on, at once, in bulk.
3. Damage Responsive Regulatory Elements (DRREs) are enhancer elements activated upon damage. They are enriched in active enhancer features, can drive gene expression after different types of damage and can create contacts between them even through long distances. Moreover, DRREs sequence is conserved throughout insect species.
4. The combinatorial action of three different types of DRREs is required in regeneration. Novel DRREs acting exclusively in the damaged tissue cooperate with DRREs co-opted from other tissues and developmental stages, and with endogenous DRREs that show increased activity after injury. Such combinatorial action is conserved in zebrafish heart regeneration.
5. There is a core of regeneration genes conserved across metazoans. Such core is enriched in transcription related genes as well as in signalling pathways genes required for patterning and growth. Binding sites for conserved TFs are found in DRREs.



# BIBLIOGRAPHY





## A

- Arnold CD, Gerlach D, Spies D, Matts JA, Sytnikova YA, Pagani M, Lau NC, Stark A. 2014. Quantitative genome-wide enhancer activity maps for five *Drosophila* species show functional enhancer conservation and turnover during cis-regulatory evolution. *Nat Genet* **46**: 685–692.
- Arnosti DN, Kulkarni MM. 2005. Transcriptional enhancers: Intelligent enhanceosomes or flexible billboards? *J Cell Biochem* **94**: 890–898.

## B

- Baguñà J, Saló E, Romero R. 1989. Effects of activators and antagonists of the neuropeptides substance P and substance K on cell proliferation in planarians. *Int J Dev Biol* **33**: 261–6.
- Banerji J, Rusconi S, Schaffner W. 1981. Expression of a beta-globin gene is enhanced by remote SV40 DNA sequences. *Cell* **27**: 299–308.
- Barozzi I, Simonatto M, Bonifacio S, Yang L, Rohs R, Ghisletti S, Natoli G. 2014. Coregulation of transcription factor binding and nucleosome occupancy through DNA features of mammalian enhancers. *Mol Cell* **54**: 844–857.
- Barrio R, de Celis JF. 2004. Regulation of spalt expression in the *Drosophila* wing blade in response to the Decapentaplegic signaling pathway. *Proc Natl Acad Sci U S A* **101**: 6021–6.
- Bayarmagnai B, Nicolay BN, Islam ABMMK, Lopez-Bigas N, Frolov M V. 2012. *Drosophila* GAGA factor is required for full activation of the dE2f1-Yki/Sd transcriptional program. *Cell Cycle* **11**: 4191–202.
- Bely AE, Nyberg KG. 2010. Evolution of animal regeneration: re-emergence of a field. *Trends Ecol Evol* **25**: 161–70.
- Bergantiños C, Corominas M, Serras F. 2010. Cell death-induced regeneration in wing imaginal discs requires JNK signalling. *Development* **137**: 1169–79.
- Bernstein BE, Mikkelsen TS, Xie X, Kamal M, Huebert DJ, Cuff J, Fry B, Meissner A, Wernig M, Plath K, et al. 2006. A bivalent chromatin structure marks key developmental genes in embryonic stem cells. *Cell* **125**: 315–26.
- Bieli D, Kanca O, Requena D, Hamaratoglu F, Gohl D, Schedl P, Affolter M, Slattery M, Müller M, Estella C. 2015. Establishment of a Developmental Compartment Requires Interactions between Three Synergistic Cis-regulatory Modules. *PLOS Genet* **11**: e1005376.
- Blackwood EM, Kadonaga JT. 1998. Going the distance: a current view of enhancer action.

*Science* **281**: 60–3.

- Blanco E, Ruiz-Romero M, Beltran S, Bosch M, Punset A, Serras F, Corominas M. 2010. Gene expression following induction of regeneration in *Drosophila* wing imaginal discs. Expression profile of regenerating wing discs. *BMC Dev Biol* **10**: 94.
- Bolstad BM, Irizarry RA, Astrand M, Speed TP. 2003. A comparison of normalization methods for high density oligonucleotide array data based on variance and bias. *Bioinformatics* **19**: 185–93.
- Bonn S, Zinzen RP, Girardot C, Gustafson EH, Perez-Gonzalez A, Delhomme N, Ghavi-Helm Y, Wilczyński B, Riddell A, Furlong EEM. 2012. Tissue-specific analysis of chromatin state identifies temporal signatures of enhancer activity during embryonic development. *Nat Genet* **44**: 148–156.
- Boone E, Colombani J, Andersen DS, Léopold P. 2016. The Hippo signalling pathway coordinates organ growth and limits developmental variability by controlling *dilp8* expression. *Nat Commun* **7**: 13505.
- Bosch M, Baguñà J, Serras F. 2008. Origin and proliferation of blastema cells during regeneration of *Drosophila* wing imaginal discs. *Int J Dev Biol* **52**: 1043–50.
- Bosch M, Serras F, Martín-Blanco E, Baguñà J. 2005. JNK signaling pathway required for wound healing in regenerating *Drosophila* wing imaginal discs. *Dev Biol* **280**: 73–86.
- Boutanaev AM, Kalmykova AI, Shevelyov YY, Nurminsky DI. 2002. Large clusters of co-expressed genes in the *Drosophila* genome. *Nature* **420**.
- Boyle AP, Davis S, Shulha HP, Meltzer P, Margulies EH, Weng Z, Furey TS, Crawford GE. 2008. High-Resolution Mapping and Characterization of Open Chromatin across the Genome. *Cell* **132**: 311–322.
- Brand AH, Perrimon N. 1993. Targeted gene expression as a means of altering cell fates and generating dominant phenotypes. *Development* **118**: 401–15.
- Brockes JP, Kumar A. 2008. Comparative aspects of animal regeneration. *Annu Rev Cell Dev Biol* **24**: 525–49.
- Bryant S V., French V, Bryant PJ. 1981. Distal Regeneration and Symmetry. *Science* (80- ) **212**: 993–1002.
- Buchon N, Broderick NA, Kuraishi T, Lemaitre B. 2010. *Drosophila* EGFR pathway coordinates stem cell proliferation and gut remodeling following infection. *BMC Biol* **8**: 152.
- Buenrostro JD, Giresi PG, Zaba LC, Chang HY, Greenleaf WJ. 2013. Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. *Nat Methods* **10**: 1213–1218.

- Buenrostro JD, Wu B, Chang HY, Greenleaf WJ. 2015. ATAC-seq: A Method for Assaying Chromatin Accessibility Genome-Wide. In *Current Protocols in Molecular Biology*, Vol. **109** of, p. 21.29.1-21.29.9
- Bullard KM, Longaker MT, Lorenz HP. 2003. Fetal wound healing: current biology. *World J Surg* **27**: 54–61.
- Burzyn D, Kuswanto W, Kolodin D, Shadrach JL, Cerletti M, Jang Y, Sefik E, Tan TG, Wagers AJ, Benoist C, et al. 2013. A Special Population of Regulatory T Cells Potentiates Muscle Repair. *Cell* **155**: 1282–1295.

## C

- Caddy J, Wilanowski T, Darido C, Dworkin S, Ting SB, Zhao Q, Rank G, Auden A, Srivastava S, Papenfuss TA, et al. 2010. Epidermal Wound Repair Is Regulated by the Planar Cell Polarity Signaling Pathway. *Dev Cell* **19**: 138–147.
- Calo E, Wysocka J. 2013. Modification of enhancer chromatin: what, how, and why? *Mol Cell* **49**: 825–37.
- Carey LB, van Dijk D, Sloot PMA, Kaandorp JA, Segal E. 2013. Promoter Sequence Determines the Relationship between Expression Level and Noise ed. R. Singer. *PLoS Biol* **11**: e1001528.
- Catarino RR, Stark A. 2018. Assessing sufficiency and necessity of enhancer activities for gene expression and the mechanisms of transcription activation. *Genes Dev* **32**: 202–223.
- Charoensawan V, Janga SC, Bulyk ML, Babu MM, Teichmann SA. 2012. DNA Sequence Preferences of Transcriptional Activators Correlate More Strongly than Repressors with Nucleosomes. *Mol Cell* **47**: 183–192.
- Chen C-H, Poss KD. 2017. Regeneration Genetics. *Annu Rev Genet* **51**: annurev-genet-120116-024554.
- Chubb JR, Trcek T, Shenoy SM, Singer RH. 2006. Transcriptional Pulsing of a Developmental Gene. *Curr Biol* **16**: 1018–1025.
- Cienfuegos JA, Rotellar F, Baixauli J, Martínez-Regueira F, Pardo F, Hernández-Lizoáin JL. 2014. Liver regeneration--the best kept secret. A model of tissue injury response. *Rev Esp Enferm Dig* **106**: 171–94.
- Cohen SM. 1993. Imaginal disc development. pp. 747–841 (Bate M, Martinez Arias A, eds. 1993. *The Development of Drosophila melanogaster*, Vol. 2. Plainview, NY: Cold Spring Harbor Lab)

- Colombani J, Andersen DS, Leopold P. 2012. Secreted Peptide Dilp8 Coordinates *Drosophila* Tissue Growth with Developmental Timing. *Science* (80- ) **336**: 582–585.
- Corrales M, Rosado A, Cortini R, van Arensbergen J, van Steensel B, Filion GJ. 2017. Clustering of *Drosophila* housekeeping promoters facilitates their expression. *Genome Res* **27**: 1153–1161.
- Cubeñas-Potts C, Rowley MJ, Lyu X, Li G, Lei EP, Corces VG. 2017. Different enhancer classes in *Drosophila* bind distinct architectural proteins and mediate unique chromatin interactions and 3D architecture. *Nucleic Acids Res* **45**: 1714–1730.
- Cummings SG, Bode HR. 1984. Head regeneration and polarity reversal in *Hydra attenuata* can occur in the absence of DNA synthesis. *Wilhelm Roux's Arch Dev Biol* **194**: 79–86.

## D

- Dar RD, Razoooky BS, Singh A, Trimeloni T V., McCollum JM, Cox CD, Simpson ML, Weinberger LS. 2012. Transcriptional burst frequency and burst size are equally modulated across the human genome. *Proc Natl Acad Sci* **109**: 17454–17459.
- Davie K, Jacobs J, Atkins M, Potier D, Christiaens V, Halder G, Aerts S. 2015. Discovery of transcription factors and regulatory regions driving in vivo tumor development by ATAC-seq and FAIRE-seq open chromatin profiling. ed. P. McKinnon. *PLoS Genet* **11**: e1004994.
- De Santa F, Barozzi I, Mietton F, Ghisletti S, Polletti S, Tusi BK, Muller H, Ragoussis J, Wei C-L, Natoli G. 2010. A Large Fraction of Extragenic RNA Pol II Transcription Sites Overlap Enhancers ed. J.S. Mattick. *PLoS Biol* **8**: e1000384.
- De Wit E, de Laat W. 2012. A decade of 3C technologies: insights into nuclear organization. *Genes Dev* **26**: 11–24.
- Dekker J, Marti-Renom MA, Mirny LA. 2013. Exploring the three-dimensional organization of genomes: interpreting chromatin interaction data. *Nat Rev Genet* **14**: 390–403.
- Deplancke B, Alpern D, Gardeux V. 2016. The Genetics of Transcription Factor DNA Binding Variation. *Cell* **166**: 538–554.
- Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, Gingeras TR. 2013. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**: 15–21.
- Dorigi KM, Swigut T, Henriques T, Bhanu N V, Scruggs BS, Nady N, Still CD, Garcia BA, Adelman K, Wysocka J. 2017. Mll3 and Mll4 Facilitate Enhancer RNA Synthesis and Transcription from Promoters Independently of H3K4 Monomethylation. *Mol Cell* **66**: 568–576.e4.

## E

- Fogarty CE, Diwanji N, Lindblad JL, Tare M, Amcheslavsky A, Makhijani K, Brückner K, Fan Y, Bergmann A. 2016. Extracellular Reactive Oxygen Species Drive Apoptosis-Induced Proliferation via *Drosophila* Macrophages. *Curr Biol* **26**: 575–584.
- French V. 1978. Intercalary regeneration around the circumference of the cockroach leg. *J Embryol Exp Morphol* **47**: 53–84.
- French V. 1981. Pattern regulation and regeneration. *Philos Trans R Soc Lond B Biol Sci* **295**: 601–17.
- Fristrom D, Fristrom JW. 1993. The metamorphic development of the adult epidermis. pp. 843–97 ((Bate M, Martinez Arias A, eds. 1993. *The Development of Drosophila melanogaster*, Vol. 2. Plainview, NY: Cold Spring Harbor Lab)

## G

- Garcia-Bellido A, Ripoll P, Morata G. 1973. Developmental compartmentalisation of the wing disk of *Drosophila*. *Nat New Biol* **245**: 251–3.
- Gauron C, Rampon C, Bouzaffour M, Ipendey E, Teillon J, Volovitch M, Vríz S. 2013. Sustained production of ROS triggers compensatory proliferation and is required for regeneration to proceed. *Sci Rep* **3**: 2084.
- Gehrke AR, Schneider I, de la Calle-Mustienes E, Tena JJ, Gomez-Marin C, Chandran M, Nakamura T, Braasch I, Postlethwait JH, Gómez-Skarmeta JL, et al. 2015. Deep conservation of wrist and digit enhancers in fish. *Proc Natl Acad Sci* **112**: 803–808.
- Ghavi-Helm Y, Klein FA, Pakozdi T, Ciglar L, Noordermeer D, Huber W, Furlong EEM. 2014. Enhancer loops appear stable during development and are associated with paused polymerase. *Nature* **512**: 96–100.
- Gibson MC, Schubiger G. 1999. Hedgehog is required for activation of engrailed during regeneration of fragmented *Drosophila* imaginal discs. *Development* **126**: 1591–9.
- Golding I, Paulsson J, Zawilski SM, Cox EC. 2005. Real-Time Kinetics of Gene Activity in Individual Bacteria. *Cell* **123**: 1025–1036.
- Goldman JA, Kuzu G, Lee N, Karasik J, Gemberling M, Foglia MJ, Karra R, Dickson AL, Sun F, Tolstorukov MY, et al. 2017. Resolving Heart Regeneration by Replacement Histone Profiling. *Dev Cell* **40**: 392–404.e5.
- González-Rosa JM, Burns CE, Burns CG. 2017. Zebrafish heart regeneration: 15 years of discoveries. *Regen (Oxford, England)* **4**: 105–123.

- Graveley BR, Brooks AN, Carlson JW, Duff MO, Landolin JM, Yang L, Artieri CG, van Baren MJ, Boley N, Booth BW, et al. 2011. The developmental transcriptome of *Drosophila melanogaster*. *Nature* **471**: 473–479.
- Grusche FA, Degoutin JL, Richardson HE, Harvey KF. 2011. The Salvador/Warts/Hippo pathway controls regenerative tissue growth in *Drosophila melanogaster*. *Dev Biol* **350**: 255–266.
- Guenther CA, Wang Z, Li E, Tran MC, Logan CY, Nusse R, Pantalena-Filho L, Yang GP, Kingsley DM. 2015. A distinct regulatory region of the *Bmp5* locus activates gene expression following adult bone fracture or soft tissue injury. *Bone* **77**: 31–41.
- Gurudatta B, Yang J, Van Bortle K, Donlin-Asp P, Corces V. 2013. Dynamic changes in the genomic localization of DNA replication-related element binding factor during the cell cycle. *Cell Cycle* **12**: 1605–1615.

## H

- Hariharan IK, Serras F. 2017. Imaginal disc regeneration takes flight. *Curr Opin Cell Biol* **48**: 10–16.
- Harris RE, Setiawan L, Saul J, Hariharan IK. 2016. Localized epigenetic silencing of a damage-activated WNT enhancer limits regeneration in mature *Drosophila* imaginal discs. *Elife* **5**.
- Haynie JL, Bryant PJ. 1977. The effects of X-rays on the proliferation dynamics of cells in the imaginal wing disc of *Drosophila melanogaster*. *Wilhelm Roux's Arch Dev Biol* **183**: 85–100.
- He Q, Bardet AF, Patton B, Purvis J, Johnston J, Paulson A, Gogol M, Stark A, Zeitlinger J. 2011. High conservation of transcription factor binding and evidence for combinatorial regulation across six *Drosophila* species. *Nat Genet* **43**: 414–420.
- Henry JJ, Tsonis PA. 2010. Molecular and cellular aspects of amphibian lens regeneration. *Prog Retin Eye Res* **29**: 543–55.
- Herranz H, Pérez L, Martín FA, Milán M. 2008. A Wingless and Notch double-repression mechanism regulates G1–S transition in the *Drosophila* wing. *EMBO J* **27**: 1633–1645.
- Herrera SC, Martín R, Morata G. 2013. Tissue homeostasis in the wing disc of *Drosophila melanogaster*: immediate response to massive damage during development. ed. N. Perrimon. *PLoS Genet* **9**: e1003446.
- Herrera SC, Morata G. 2014. Transgressions of compartment boundaries and cell reprogramming during regeneration in *Drosophila*. *Elife* **3**: e01831.

- Herrmann C, Van de Sande B, Potier D, Aerts S. 2012. i-cisTarget: an integrative genomics method for the prediction of regulatory features and cis-regulatory modules. *Nucleic Acids Res* **40**: e114–e114.
- Huang DW, Sherman BT, Lempicki RA. 2009. Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Res* **37**: 1–13.
- Huang DW, Sherman BT, Lempicki RA. 2008. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc* **4**: 44–57.
- Huang GN, Thatcher JE, McAnally J, Kong Y, Qi X, Tan W, DiMaio JM, Amatruda JF, Gerard RD, Hill JA, et al. 2012. C/EBP transcription factors mediate epicardial activation during heart development and injury. *Science* **338**: 1599–603.

## I

- Iismaa SE, Kaidonis X, Nicks AM, Bogush N, Kikuchi K, Naqvi N, Harvey RP, Husain A, Graham RM. Comparative regenerative mechanisms across different mammalian tissues.
- Iten LE, Bryant S V. 1973. Forelimb regeneration from different levels of amputation in the newt, *Notophthalmus viridescens*: Length, rate, and stages. *Wilhelm Roux Arch Entwickl Mech Org* **173**: 263–282.

## J

- Janky R, Verfaillie A, Imrichová H, Van de Sande B, Standaert L, Christiaens V, Hulselmans G, Herten K, Naval Sanchez M, Potier D, et al. 2014. iRegulon: From a Gene List to a Gene Regulatory Network Using Large Motif and Track Collections ed. H.J. Bussemaker. *PLoS Comput Biol* **10**: e1003731.
- Jaszczak JS, Halme A. 2016. Arrested development: coordinating regeneration with development and growth in *Drosophila melanogaster*. *Curr Opin Genet Dev* **40**: 87–94.
- Jaszczak JS, Wolpe JB, Dao AQ, Halme A. 2015. Nitric Oxide Synthase Regulates Growth Coordination During *Drosophila melanogaster* Imaginal Disc Regeneration. *Genetics* **200**: 1219–28.
- John S, Sabo PJ, Thurman RE, Sung M-H, Biddie SC, Johnson TA, Hager GL, Stamatoyannopoulos JA. 2011. Chromatin accessibility pre-determines glucocorticoid receptor binding patterns. *Nat Genet* **43**: 264–8.
- Johnson DS, Mortazavi A, Myers RM, Wold B. 2007. Genome-wide mapping of in vivo protein-DNA interactions. *Science* **316**: 1497–502.

- Jopling C, Sleep E, Raya M, Martí M, Raya A, Belmonte JCI. 2010. Zebrafish heart regeneration occurs by cardiomyocyte dedifferentiation and proliferation. *Nature* **464**: 606–609.
- Joshi S, Davidson G, Le Gras S, Watanabe S, Braun T, Mengus G, Davidson I. 2017. TEAD transcription factors are required for normal primary myoblast differentiation in vitro and muscle regeneration in vivo. ed. G.S. Barsh. *PLoS Genet* **13**: e1006600.
- Jung YH, Sauria MEG, Lyu X, Cheema MS, Ausio J, Taylor J, Corces VG. 2017. Chromatin States in Mouse Sperm Correlate with Embryonic and Adult Regulatory Landscapes. *Cell Rep* **18**: 1366–1382.
- Junion G, Spivakov M, Girardot C, Braun M, Gustafson EH, Birney E, Furlong EEM. 2012. A Transcription Factor Collective Defines Cardiac Cell Fate and Reflects Lineage History. *Cell* **148**: 473–486.

## K

- Kanehisa M, Furumichi M, Tanabe M, Sato Y, Morishima K. 2017. KEGG: new perspectives on genomes, pathways, diseases and drugs. *Nucleic Acids Res* **45**: D353–D361.
- Kanehisa M, Goto S. 2000. KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res* **28**: 27–30.
- Kanehisa M, Sato Y, Kawashima M, Furumichi M, Tanabe M. 2016. KEGG as a reference resource for gene and protein annotation. *Nucleic Acids Res* **44**: D457–D462.
- Kang J, Hu J, Karra R, Dickson AL, Tornini VA, Nachtrab G, Gemberling M, Goldman JA, Black BL, Poss KD. 2016. Modulation of tissue repair by regeneration enhancer elements. *Nature* **532**: 201–6.
- Karra R, Knecht AK, Kikuchi K, Poss KD. 2015. Myocardial NF- $\kappa$ B activation is essential for zebrafish heart regeneration. *Proc Natl Acad Sci U S A* **112**: 13255–60.
- Katsuyama T, Comoglio F, Seimiya M, Cabuy E, Paro R. 2015. During Drosophila disc regeneration, JAK/STAT coordinates cell proliferation with Dilp8-mediated developmental delay. *Proc Natl Acad Sci U S A* **112**: E2327-36.
- Katsuyama T, Paro R. 2011. Epigenetic reprogramming during tissue regeneration. *FEBS Lett* **585**: 1617–24.
- Khan SJ, Abidi SNF, Skinner A, Tian Y, Smith-Bolton RK. 2017. The Drosophila Duox maturation factor is a key component of a positive feedback loop that sustains regeneration signaling ed. G. Bosco. *PLOS Genet* **13**: e1006937.
- Kharchenko P V, Tolstorukov MY, Park PJ. 2008. Design and analysis of ChIP-seq experiments

for DNA-binding proteins. *Nat Biotechnol* **26**: 1351–1359.

Kim T-K, Hemberg M, Gray JM, Costa AM, Bear DM, Wu J, Harmin DA, Laptewicz M, Barbara-Haley K, Kuersten S, et al. 2010. Widespread transcription at neuronal activity-regulated enhancers. *Nature* **465**: 182–187.

Koch F, Fenouil R, Gut M, Cauchy P, Albert TK, Zacarias-Cabeza J, Spicuglia S, de la Chapelle AL, Heidemann M, Hintermair C, et al. 2011. Transcription initiation platforms and GTF recruitment at tissue-specific enhancers and promoters. *Nat Struct Mol Biol* **18**: 956–963.

Koenecke N, Johnston J, Gaertner B, Natarajan M, Zeitlinger J. 2016. Genome-wide identification of *Drosophila* dorso-ventral enhancers by differential histone acetylation analysis. *Genome Biol* **17**: 196.

## L

La Fortezza M, Schenk M, Cosolo A, Kolybaba A, Grass I, Classen A-K. 2016. JAK/STAT signalling mediates cell survival in response to tissue stress. *Development* **143**: 2907–2919.

LaJeunesse DR, McCartney BM, Fehon RG. 1998. Structural Analysis of *Drosophila* Merlin Reveals Functional Domains Important for Growth Control and Subcellular Localization. *J Cell Biol* **141**: 1589–1599.

Lam MTY, Cho H, Lesch HP, Gosselin D, Heinz S, Tanaka-Oishi Y, Benner C, Kaikkonen MU, Kim AS, Kosaka M, et al. 2013. Rev-Erbs repress macrophage gene expression by inhibiting enhancer-directed transcription. *Nature* **498**: 511–515.

Landt SG, Marinov GK, Kundaje A, Kheradpour P, Pauli F, Batzoglou S, Bernstein BE, Bickel P, Brown JB, Cayting P, et al. 2012. ChIP-seq guidelines and practices of the ENCODE and modENCODE consortia. *Genome Res* **22**: 1813–31.

Lee N, Maurange C, Ringrose L, Paro R. 2005. Suppression of Polycomb group proteins by JNK signalling induces transdetermination in *Drosophila* imaginal discs. *Nature* **438**: 234–7.

Levin M. 2009. Bioelectric mechanisms in regeneration: Unique aspects and future perspectives. *Semin Cell Dev Biol* **20**: 543–556.

Li B, Dewey CN. 2011. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics* **12**: 323.

Li X-Y, Thomas S, Sabo PJ, Eisen MB, Stamatoyannopoulos JA, Biggin MD. 2011. The role of chromatin accessibility in directing the widespread, overlapping patterns of *Drosophila*

transcription factor binding. *Genome Biol* **12**: R34.

Liu S-Y, Selck C, Friedrich B, Lutz R, Vila-Farré M, Dahl A, Brandl H, Lakshmanaperumal N, Henry I, Rink JC. 2013. Reactivating head regrowth in a regeneration-deficient planarian species. *Nature* **500**: 81–84.

Long HK, Prescott SL, Wysocka J. 2016. Ever-Changing Landscapes: Transcriptional Enhancers in Development and Evolution. *Cell* **167**: 1170–1187.

López-Luque J, Caballero-Díaz D, Martínez-Palacián A, Roncero C, Moreno-Càceres J, García-Bravo M, Grueso E, Fernández A, Crosas-Molist E, García-Álvaro M, et al. 2016. Dissecting the role of epidermal growth factor receptor catalytic activity during liver regeneration and hepatocarcinogenesis. *Hepatology* **63**: 604–619.

López-Luque J, Fabregat I. 2018. Revisiting the liver: from development to regeneration - what we ought to know! *Int J Dev Biol* **62**: 441–451.

Loubière V, Delest A, Thomas A, Bonev B, Schuettengruber B, Sati S, Martínez A-M, Cavalli G. 2016. Coordinate redeployment of PRC1 proteins suppresses tumor formation during *Drosophila* development. *Nat Genet* **48**: 1436–1442.

Lykke-Andersen S, Jensen TH. 2015. Nonsense-mediated mRNA decay: an intricate machinery that shapes transcriptomes. *Nat Rev Mol Cell Biol* **16**: 665–77.

## M

Mace KA, Pearson JC, McGinnis W. 2005. An epidermal barrier wound repair pathway in *Drosophila* is mediated by grainy head. *Science* **308**: 381–5.

Mahmoudi T, Katsani KR, Verrijzer CP. 2002. GAGA can mediate enhancer function in trans by linking two separate DNA molecules. *EMBO J* **21**: 1775–1781.

Maier HJ, Schips TG, Wietelmann A, Krüger M, Brunner C, Sauter M, Klingel K, Böttger T, Braun T, Wirth T. 2012. Cardiomyocyte-specific I $\kappa$ B kinase (IKK)/NF- $\kappa$ B activation induces reversible inflammatory cardiomyopathy and heart failure. *Proc Natl Acad Sci U S A* **109**: 11794–9.

Marco-Sola S, Sammeth M, Guigó R, Ribeca P. 2012. The GEM mapper: fast, accurate and versatile alignment by filtration. *Nat Methods* **9**: 1185–1188.

Martín FA, Morata G. 2006. Compartments and the control of growth in the *Drosophila* wing imaginal disc. *Development* **133**: 4421–6.

Mattila J, Omelyanchuk L, Kyttälä S, Turunen H, Nokkala S. 2005. Role of Jun N-terminal Kinase (JNK) signaling in the wound healing and regeneration of a *Drosophila melanogaster*

- wing imaginal disc. *Int J Dev Biol* **49**: 391–9.
- Maurange C, Lee N, Paro R. 2006. Signaling meets chromatin during tissue regeneration in *Drosophila*. *Curr Opin Genet Dev* **16**: 485–489.
- McClure KD, Sustar A, Schubiger G. 2008. Three genes control the timing, the site and the size of blastema formation in *Drosophila*. *Dev Biol* **319**: 68–77.
- McGuire SE, Le PT, Osborn AJ, Matsumoto K, Davis RL. 2003. Spatiotemporal rescue of memory dysfunction in *Drosophila*. *Science* **302**: 1765–8.
- McKay DJ, Lieb JD. 2013. A common set of DNA regulatory elements shapes *Drosophila* appendages. *Dev Cell* **27**: 306–18.
- Merika M, Thanos D. 2001. Enhanceosomes. *Curr Opin Genet Dev* **11**: 205–8.
- Meserve JH, Duronio RJ. 2015. Scalloped and Yorkie are required for cell cycle re-entry of quiescent cells after tissue damage. *Development* **142**: 2740–2751.
- Michalak P. 2008. Coexpression, coregulation, and cofunctionality of neighboring genes in eukaryotic genomes. *Genomics* **91**: 243–8.
- Michalopoulos GK, DeFrances MC. 1997. Liver regeneration. *Science* **276**: 60–6.
- Mikhaylichenko O, Bondarenko V, Harnett D, Schor IE, Males M, Viales RR, Furlong EEM. 2018. The degree of enhancer or promoter activity is reflected by the levels and directionality of eRNA transcription. *Genes Dev* **32**: 42–57.
- Montavon T, Soshnikova N, Mascrez B, Joye E, Thevenet L, Splinter E, de Laat W, Spitz F, Duboule D. 2011. A regulatory archipelago controls Hox genes transcription in digits. *Cell* **147**: 1132–45.
- Moreira S, Stramer B, Evans I, Wood W, Martin P. 2010. Prioritization of Competing Damage and Developmental Signals by Migrating Macrophages in the *Drosophila* Embryo. *Curr Biol* **20**: 464–470.
- Morgan TH: Regeneration. New York: Macmillan; 1901.

## N

- Nakamura M, Verboon JM, Parkhurst SM. 2017. Prepatterning by RhoGEFs governs Rho GTPase spatiotemporal dynamics during wound repair. *J Cell Biol* **216**: 3959–3969.
- Narciso C, Wu Q, Brodskiy P, Garston G, Baker R, Fletcher A, Zartman J. 2015. Patterning of wound-induced intercellular Ca(2+) flashes in a developing epithelium. *Phys Biol* **12**: 056005.

Nègre N, Brown CD, Ma L, Bristow CA, Miller SW, Wagner U, Kheradpour P, Eaton ML, Loriaux P, Sealfon R, et al. 2011. A cis-regulatory map of the *Drosophila* genome. *Nature* **471**: 527–531.

Neph S, Kuehn MS, Reynolds AP, Haugen E, Thurman RE, Johnson AK, Rynes E, Maurano MT, Vierstra J, Thomas S, et al. 2012. BEDOPS: high-performance genomic feature operations. *Bioinformatics* **28**: 1919–20.

Niethammer P, Grabher C, Look AT, Mitchison TJ. 2009. A tissue-scale gradient of hydrogen peroxide mediates rapid wound detection in zebrafish. *Nature* **459**: 996–999.

## O

Ogiyama Y, Schuettengruber B, Papadopoulos GL, Chang J-M, Cavalli G. 2018. Polycomb-Dependent Chromatin Looping Contributes to Gene Silencing during *Drosophila* Development. *Mol Cell* **71**: 73–88.e5.

Oh H, Slattery M, Ma L, Crofts A, White KP, Mann RS, Irvine KD. 2013. Genome-wide Association of Yorkie with Chromatin and Chromatin-Remodeling Complexes. *Cell Rep* **3**: 309–318.

Orphanides G, Lagrange T, Reinberg D. 1996. The general transcription factors of RNA polymerase II. *Genes Dev* **10**: 2657–83. Ostuni R, Piccolo V, Barozzi I, Polletti S, Termanini A, Bonifacio S, Curina A, Prosperini E, Ghisletti S, Natoli G. 2013. Latent Enhancers Activated by Stimulation in Differentiated Cells. *Cell* **152**: 157–171.

## P

Pastor-Pareja JC, Wu M, Xu T. 2008. An innate immune response of blood cells to tumors and tissue damage in *Drosophila*. *Dis Model Mech* **1**: 144–54; discussion 153.

Pérez-Lluch S, Blanco E, Carbonell A, Raha D, Snyder M, Serras F, Corominas M. 2011. Genome-wide chromatin occupancy analysis reveals a role for ASH2 in transcriptional pausing. *Nucleic Acids Res* **39**: 4628–39.

Peters AHFM, Mermoud JE, O’Carroll D, Pagani M, Schweizer D, Brockdorff N, Jenuwein T. 2001. Histone H3 lysine 9 methylation is an epigenetic imprint of facultative heterochromatin. *Nat Genet* **30**: 77–80.

Petrie TA, Strand NS, Tsung-Yang C, Rabinowitz JS, Moon RT, Moon RT. 2014. Macrophages modulate adult zebrafish tail fin regeneration. *Development* **141**: 2581–2591.

Pignatelli M, Serras F, Moya A, Guigo R, Corominas M. 2009. CROC: finding chromosomal

clusters in eukaryotic genomes. *Bioinformatics* **25**: 1552–1553.

Pohl A, Beato M. 2014. bwtool: a tool for bigWig files. *Bioinformatics* **30**: 1618–9.

Pollex T, Furlong EEM. 2017. Correlation Does Not Imply Causation: Histone Methyltransferases, but Not Histone Methylation, SET the Stage for Enhancer Activation. *Mol Cell* **66**: 439–441.

Porrello ER, Mahmoud AI, Simpson E, Hill JA, Richardson JA, Olson EN, Sadek HA. 2011. Transient Regenerative Potential of the Neonatal Mouse Heart. *Science* **331**: 1078–1080.

Poss KD, Wilson LG, Keating MT. 2002. Heart Regeneration in Zebrafish. *Science* **298**: 2188–2190.

## Q

Quinlan AR, Hall IM. 2010. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**: 841–2.

## R

Rada-Iglesias A, Bajpai R, Swigut T, Brugmann SA, Flynn RA, Wysocka J. 2011. A unique chromatin signature uncovers early developmental enhancers in humans. *Nature* **470**: 279–283.

Raj A, Peskin CS, Tranchina D, Vargas DY, Tyagi S. 2006. Stochastic mRNA Synthesis in Mammalian Cells ed. U. Schibler. *PLoS Biol* **4**: e309.

Razzell W, Evans IR, Martin P, Wood W. 2013. Calcium flashes orchestrate the wound inflammatory response through DUOX activation and hydrogen peroxide release. *Curr Biol* **23**: 424–9.

Reiter F, Wienerroither S, Stark A. 2017. Combinatorial function of transcription factors and cofactors. *Curr Opin Genet Dev* **43**: 73–81.

Repiso A, Bergantinos C, Serras F. 2013. Cell fate respecification and cell division orientation drive intercalary regeneration in *Drosophila* wing discs. *Development* **140**: 3541–3551.

Restrepo S, Basler K. 2016. *Drosophila* wing imaginal discs respond to mechanical injury via slow InsP3R-mediated intercellular calcium waves. *Nat Commun* **7**: 12450.

Rickels R, Herz H-M, Sze CC, Cao K, Morgan MA, Collings CK, Gause M, Takahashi Y, Wang L, Rendleman EJ, et al. 2017. Histone H3K4 monomethylation catalyzed by Trr

and mammalian COMPASS-like proteins at enhancers is dispensable for development and viability. *Nat Genet* **49**: 1647–1653.

Rodius S, Androsova G, Götz L, Liechti R, Crespo I, Merz S, Nazarov P V, de Klein N, Jeanty C, González-Rosa JM, et al. 2016. Analysis of the dynamic co-expression network of heart regeneration in the zebrafish. *Sci Rep* **6**: 26822.

Roeder RG. 1996. The role of general initiation factors in transcription by RNA polymerase II. *Trends Biochem Sci* **21**: 327–35.

Roehl HH. 2018. Linking wound response and inflammation to regeneration in the zebrafish larval fin. *Int J Dev Biol* **62**: 473–477.

Rowley MJ, Corces VG. 2016. The three-dimensional genome: principles and roles of long-distance interactions. *Curr Opin Cell Biol* **40**: 8–14.

Rowley MJ, Nichols MH, Lyu X, Ando-Kuri M, Rivera ISM, Hermetz K, Wang P, Ruan Y, Corces VG. 2017. Evolutionarily Conserved Principles Predict 3D Chromatin Organization. *Mol Cell* **67**: 837–852.e7.

Ruiz-Romero M, Blanco E, Paricio N, Serras F, Corominas M. 2015. Cabut/dTIEG associates with the transcription factor Yorkie for growth control. *EMBO Rep* **16**: 362–369.

## S

Sánchez-Iranzo H, Galardi-Castilla M, Minguillón C, Sanz-Morejón A, González-Rosa JM, Felker A, Ernst A, Guzmán-Martínez G, Mosimann C, Mercader N. 2018. Tbx5a lineage tracing shows cardiomyocyte plasticity during zebrafish heart regeneration. *Nat Commun* **9**: 428.

Sánchez Alvarado A, Tsonis PA. 2006. Bridging the regeneration gap: genetic insights from diverse animal models. *Nat Rev Genet* **7**: 873–84.

Santabárbara-Ruiz P, López-Santillán M, Martínez-Rodríguez I, Binagui-Casas A, Pérez L, Milán M, Corominas M, Serras F. 2015. ROS-Induced JNK and p38 Signaling Is Required for Unpaired Cytokine Activation during Drosophila Regeneration ed. G.P. Copenhaver. *PLOS Genet* **11**: e1005595.

Schubiger G, Hadorn E. 1968. Auto- und allotypische Differenzierungen ausin vivo-kultivierten Vorderbeinblastemen von Drosophila melanogaster. *Dev Biol* **17**: 584–602.

Schuettengruber B, Bourbon H-M, Di Croce L, Cavalli G. 2017. Genome Regulation by Polycomb and Trithorax: 70 Years and Counting. *Cell* **171**: 34–57.

Schuster KJ, Smith-Bolton RK. 2015. Taranis Protects Regenerating Tissue from Fate Changes

- Induced by the Wound Response in *Drosophila*. *Dev Cell* **34**: 119–128.
- Schwartz YB, Cavalli G. 2017. Three-Dimensional Genome Organization and Function in *Drosophila*. *Genetics* **205**: 5–24.
- Seifert AW, Kiama SG, Seifert MG, Goheen JR, Palmer TM, Maden M. 2012a. Skin shedding and tissue regeneration in African spiny mice (*Acomys*). *Nature* **489**: 561–565.
- Seifert AW, Monaghan JR, Smith MD, Pasch B, Stier AC, Michonneau F, Maden M. 2012b. The influence of fundamental traits on mechanisms controlling appendage regeneration. *Biol Rev* **87**: 330–345.
- Sexton T, Cavalli G. 2015. The Role of Chromosome Domains in Shaping the Functional Genome. *Cell* **160**: 1049–1059.
- Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T. 2003. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res* **13**: 2498–504.
- Shaw T, Martin P. 2009. Epigenetic reprogramming during wound healing: loss of polycomb-mediated silencing may enable upregulation of repair genes. *EMBO Rep* **10**: 881–886.
- Shlyueva D, Stampfel G, Stark A. 2014. Transcriptional enhancers: from properties to genome-wide predictions. *Nat Rev Genet* **15**: 272–286.
- Simon JA, Kingston RE. 2009. Mechanisms of Polycomb gene silencing: knowns and unknowns. *Nat Rev Mol Cell Biol* **10**: 697–708.
- Simpson P, Berreur P, Berreur-Bonnenfant J. 1980. The initiation of pupariation in *Drosophila*: dependence on growth of the imaginal discs. *J Embryol Exp Morphol* **57**: 155–65.
- Skinner A, Khan SJ, Smith-Bolton RK. 2015. Trithorax regulates systemic signaling during *Drosophila* imaginal disc regeneration. *Development* **142**: 3500–3511.
- Slack JM. 2017. Animal regeneration: ancestral character or evolutionary novelty? *EMBO Rep* **18**: 1497–1508.
- Smith-Bolton RK, Worley MI, Kanda H, Hariharan IK. 2009. Regenerative Growth in *Drosophila* Imaginal Discs Is Regulated by Wingless and Myc. *Dev Cell* **16**: 797–809.
- Sousounis K, Looso M, Maki N, Ivester CJ, Braun T, Tsonis PA. 2013. Transcriptome Analysis of Newt Lens Regeneration Reveals Distinct Gradients in Gene Expression Patterns ed. X. Zhang. *PLoS One* **8**: e61445.
- Spitz F, Furlong EEM. 2012. Transcription factors: from enhancer binding to developmental control. *Nat Rev Genet* **13**: 613–26.
- Sproul D, Gilbert N, Bickmore WA. 2005. The role of chromatin structure in regulating the

expression of clustered genes. *Nat Rev Genet* **6**: 775–781.

Stark A, Lin MF, Kheradpour P, Pedersen JS, Parts L, Carlson JW, Crosby MA, Rasmussen MD, Roy S, Deoras AN, et al. 2007. Discovery of functional elements in 12 *Drosophila* genomes using evolutionary signatures. *Nature* **450**: 219–232.

Stewart S, Tsun Z-Y, Belmonte JCI. 2009. A histone demethylase is necessary for regeneration in zebrafish. *Proc Natl Acad Sci* **106**: 19889–19894.

Sun G, Irvine KD. 2011. Regulation of Hippo signaling by Jun kinase signaling during compensatory cell proliferation and regeneration, and in neoplastic tumors. *Dev Biol* **350**: 139–151.

Sun X, Chuang J-C, Kanchwala M, Wu L, Celen C, Li L, Liang H, Zhang S, Maples T, Nguyen LH, et al. 2016. Suppression of the SWI/SNF Component Arid1a Promotes Mammalian Regeneration. *Cell Stem Cell* **18**: 456–466.

Sunderland ME. 2010. Regeneration: Thomas Hunt Morgan's Window into Development. *J Hist Biol* **43**: 325–361.

Supek F, Bošnjak M, Škunca N, Šmuc T. 2011. REVIGO summarizes and visualizes long lists of gene ontology terms. ed. C. Gibas. *PLoS One* **6**: e21800.

Svaren J, Klebanow E, Sealy L, Chalkley R. 1994. Analysis of the competition between nucleosome formation and transcription factor binding. *J Biol Chem* **269**: 9335–44.

## T

Tanaka EM, Reddien PW. 2011. The Cellular Basis for Animal Regeneration. *Dev Cell* **21**: 172–185.

Tang AH, Neufeld TP, Rubin GM, Müller HA. 2001. Transcriptional regulation of cytoskeletal functions and segmentation by a novel maternal pair-rule gene, lilliputian. *Development* **128**: 801–13.

Tantale K, Mueller F, Kozulic-Pirher A, Lesne A, Victor J-M, Robert M-C, Capozzi S, Chouaib R, Bäcker V, Mateos-Langerak J, et al. 2016. A single-molecule view of transcription reveals convoys of RNA polymerases and multi-scale bursting. *Nat Commun* **7**: 12248.

Ting SB, Caddy J, Hislop N, Wilanowski T, Auden A, Zhao L-L, Ellis S, Kaur P, Uchida Y, Holleran WM, et al. 2005. A Homolog of *Drosophila* grainy head Is Essential for Epidermal Integrity in Mice. *Science* **308**: 411–413.

Tuan D, Kong S, Hu K. 1992. Transcription of the hypersensitive site HS2 enhancer in erythroid cells. *Proc Natl Acad Sci U S A* **89**: 11219–23.

Tyner C, Barber GP, Casper J, Clawson H, Diekhans M, Eisenhart C, Fischer CM, Gibson D, Gonzalez JN, Guruvadoo L, et al. 2017. The UCSC Genome Browser database: 2017 update. *Nucleic Acids Res* **45**: D626–D634.

## U

Umesono Y, Tasaki J, Nishimura Y, Hrouda M, Kawaguchi E, Yazawa S, Nishimura O, Hosoda K, Inoue T, Agata K. 2013. The molecular logic for planarian regeneration along the anterior–posterior axis. *Nature* **500**: 73–76.

## V

Venkatesh S, Workman JL. 2015. Histone exchange, chromatin structure and the regulation of transcription. *Nat Rev Mol Cell Biol* **16**: 178–189.

Verghese S, Su TT. 2016. Drosophila Wnt and STAT Define Apoptosis-Resistant Epithelial Cells for Tissue Regeneration after Irradiation. ed. B.A. Edgar. *PLoS Biol* **14**: e1002536.

Vinayagam A, Zirin J, Roesel C, Hu Y, Yilmazel B, Samsonova AA, Neumüller RA, Mohr SE, Perrimon N. 2014. Integrating protein-protein interaction networks with phenotypes reveals signs of interactions. *Nat Methods* **11**: 94–99.

Vogg MC, Wenger Y, Galliot B. 2016. How Somatic Adult Tissues Develop Organizer Activity. *Curr Top Dev Biol* **116**: 391–414. .

Vriz S, Reiter S, Galliot B. 2014. Cell Death. A Program to Regenerate. *Curr Top Dev Biol* **108**: 121-51

## W

Walter PP, Owen-Hughes TA, Côté J, Workman JL. 1995. Stimulation of transcription factor binding and histone displacement by nucleosome assembly protein 1 and nucleoplasmin requires disruption of the histone octamer. *Mol Cell Biol* **15**: 6178–87.

Wei Y, Gokhale RH, Sonnenschein A, Montgomery KM, Ingersoll A, Arnosti DN. 2016. Complex *cis* -regulatory landscape of the insulin receptor gene underlies the broad expression of a central signaling regulator. *Development* **143**: 3591–3603.

Worley MI, Setiawan L, Hariharan IK. 2012. Regeneration and Transdetermination in *Drosophila* Imaginal Discs. *Annu Rev Genet* **46**: 289–310.

Wynn TA, Vannella KM. 2016. Macrophages in Tissue Repair, Regeneration, and Fibrosis. *Immunity* **44**: 450–462.

## Y

Yates A, Akanni W, Amode MR, Barrell D, Billis K, Carvalho-Silva D, Cummins C, Clapham P, Fitzgerald S, Gil L, et al. 2016. Ensembl 2016. *Nucleic Acids Res* **44**: D710–D716.

Yoo SK, Pascoe HG, Pereira T, Kondo S, Jacinto A, Zhang X, Hariharan IK. 2016. Plexins function in epithelial repair in both *Drosophila* and zebrafish. *Nat Commun* **7**: 12282.

## Z

Zabidi MA, Arnold CD, Schernhuber K, Pagani M, Rath M, Frank O, Stark A. 2015. Enhancer-core-promoter specificity separates developmental and housekeeping gene regulation. *Nature* **518**: 556–9.

Zhang Y, Liu T, Meyer CA, Eeckhoute J, Johnson DS, Bernstein BE, Nussbaum C, Myers RM, Brown M, Li W, et al. 2008. Model-based Analysis of ChIP-Seq (MACS). *Genome Biol* **9**: R137.

Zhu MY, Wilson R, Leptin M, Rubin GM. 2005. A screen for genes that influence fibroblast growth factor signal transduction in *Drosophila*. *Genetics* **170**: 767–77.

# ANNEX





# **Annex I: Experiment genotypes**



## Experiment Genotypes

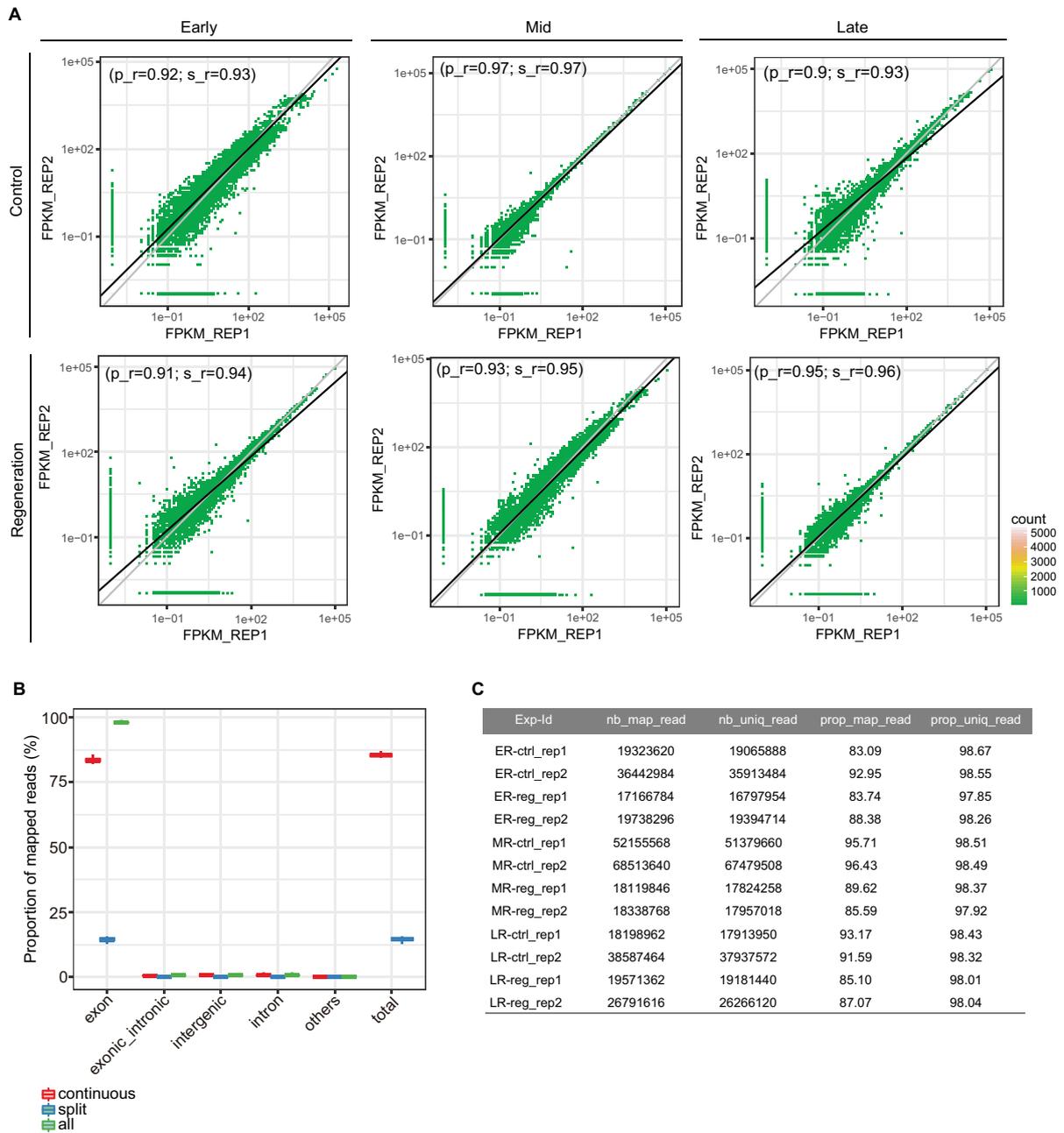
	Control	w;salm-Gal4;tub-Gal80 <sup>TS</sup>
	Regeneration	wUAS-rpr;salm-Gal4;tub-Gal80 <sup>TS</sup>
Fig. 37D	Cell Death ON, DN OFF	ci-Gal4/+;lexO-rpr/sal <sup>E/PV</sup> -LHG:tub-Gal80 <sup>TS</sup>
	Cell Death OFF, DN ON	ci-Gal4/UAS-S6K.KQ;sal <sup>E/PV</sup> -LHG:tub-Gal80 <sup>TS</sup> /+
	CellDeath OFF, DN ON	ci-Gal4/UAS-PI3K[D954A];sal <sup>E/PV</sup> -LHG:tub-Gal80 <sup>TS</sup> /+
	Cell Death ON, DN ON	ci-Gal4/UAS-S6K.KQ;lexO-rpr/sal <sup>E/PV</sup> -LHG:tub-Gal80 <sup>TS</sup>
	Cell Death ON, DN ON	ci-Gal4/UAS-PI3K[D954A];lexO-rpr/sal <sup>E/PV</sup> -LHG:tub-Gal80 <sup>TS</sup>
Fig. 52A	Control - NegCtrl	UAS-mCD8GFP/+;GMR36C06-Gal4/+
	Control - NegCtrl	UAS-mCD8GFP/+;GMR47D05-Gal4/+
	Control - NegCtrl	UAS-mCD8GFP/+;GMR88H01-Gal4/+
	Injury - NegCtrl	UAS-mCD8GFP/+;GMR36C06-Gal4/+
	Injury - NegCtrl	UAS-mCD8GFP/+;GMR47D05-Gal4/+
	Injury - NegCtrl	UAS-mCD8GFP/+;GMR88H01-Gal4 /+
Fig. 52B	Control - iDRRE	UAS-mCD8GFP/+;GMR25D02-Gal4 /+
	Control - iDRRE	UAS-mCD8GFP/+;VT39456-Gal4/+
	Control - iDRRE	UAS-mCD8GFP/+; GMR26G03-Gal4/+
	Control - iDRRE	UAS-mCD8GFP/+;GMR21F09-Gal4/+
	Control - iDRRE	UAS-mCD8GFP/+;GMR35A10-Gal4/+
	Control - iDRRE	UAS-mCD8GFP/+;GMR17D09-Gal4 /+
	Injury - iDRRE	UAS-mCD8GFP/+;GMR25D02-Gal4 /+
	Injury - iDRRE	UAS-mCD8GFP/+;VT39456-Gal4/+
	Injury - iDRRE	UAS-mCD8GFP/+; GMR26G03-Gal4/+
	Injury - iDRRE	UAS-mCD8GFP/+;GMR21F09-Gal4 /+
	Injury - iDRRE	UAS-mCD8GFP/+;GMR35A10-Gal4/+
	Injury - iDRRE	UAS-mCD8GFP/+;GMR17D09-Gal4/+
Fig. 52C	Control - eDRRE	UAS-mCD8GFP/+;GMR32B11-Gal4/+
	Control - eDRRE	UAS-mCD8GFP/+;GMR85E02-Gal4/+
	Control - eDRRE	UAS-mCD8GFP/+;GMR24G07-Gal4/+
	Control - eDRRE	UAS-mCD8GFP/+;GMR42G10-Gal4/+
	Control - eDRRE	UAS-mCD8GFP/+;GMR69F06-Gal4/+
	Control - eDRRE	UAS-mCD8GFP/+;GMR41E03-Gal4/+
	Injury - eDRRE	UAS-mCD8GFP/+;GMR32B11-Gal4/+
	Injury - eDRRE	UAS-mCD8GFP/+;GMR85E02-Gal4/+
	Injury - eDRRE	UAS-mCD8GFP/+;GMR24G07-Gal4/+
	Injury - eDRRE	UAS-mCD8GFP/+;GMR42G10-Gal4/+
	Injury - eDRRE	UAS-mCD8GFP/+;GMR69F06-Gal4/+
	Injury - eDRRE	UAS-mCD8GFP/+;GMR41E03-Gal4/+
Fig. 53A	Control - iDRRE	UAS-mCD8GFP/+;GMR25D02-Gal4 /+
	Control - iDRRE	UAS-mCD8GFP/+;VT39456-Gal4/+
	Control - iDRRE	UAS-mCD8GFP/+; GMR26G03-Gal4/+
	Control - eDRRE	UAS-mCD8GFP/+;GMR32B11-Gal4/+
	Control - eDRRE	UAS-mCD8GFP/+;GMR85E02-Gal4/+
	Control - eDRRE	UAS-mCD8GFP/+;GMR24G07-Gal4/+
	Genetic Ablation - iDRRE	UAS-mCD8GFP/lexO-rpr;GMR25D02-Gal4/sal <sup>E/PV</sup> -LHG:tub-Gal80 <sup>TS</sup>
	Genetic Ablation - iDRRE	UAS-mCD8GFP/lexO-rpr;VT39456-Gal4 /sal <sup>E/PV</sup> -LHG:tub-Gal80 <sup>TS</sup>
	Genetic Ablation - iDRRE	UAS-mCD8GFP/lexO-rpr;GMR26G03-Gal4 /sal <sup>E/PV</sup> -LHG:tub-Gal80 <sup>TS</sup>
	Genetic Ablation - eDRRE	UAS-mCD8GFP/lexO-rpr;GMR32B11-Gal4/sal <sup>E/PV</sup> -LHG:tub-Gal80 <sup>TS</sup>
	Genetic Ablation - eDRRE	UAS-mCD8GFP/lexO-rpr;GMR85E02-Gal4/sal <sup>E/PV</sup> -LHG:tub-Gal80 <sup>TS</sup>
	Genetic Ablation - eDRRE	UAS-mCD8GFP/lexO-rpr;GMR24G07-Gal4/sal <sup>E/PV</sup> -LHG:tub-Gal80 <sup>TS</sup>

Fig. 53B	Control - NegCtrl	UAS-mCD8GFP/+;GMR88H01-Gal4/+
	Genetic Ablation - NegCtrl	UAS-mCD8GFP/lexO-rpr;GMR88H01-Gal4/sal <sup>E/PV</sup> -LHG:tub-Gal80 <sup>TS</sup>
Fig. 56B	eDRRE	UAS-mCD8GFP/+;GMR32B11-Gal4/+
	eDRRE	UAS-mCD8GFP/+;GMR85E02-Gal4/+
	eDRRE	UAS-mCD8GFP/+;GMR24G07-Gal4/+
	eDRRE	UAS-mCD8GFP/+;GMR42G10-Gal4/+
	eDRRE	UAS-mCD8GFP/+;GMR69F06-Gal4/+
	eDRRE	UAS-mCD8GFP/+;GMR41E03-Gal4/+
Fig. 61	Cell Death ON, RNAi OFF	ci-Gal4/+;lexO-rpr/sal <sup>E/PV</sup> -LHG:tub-Gal80 <sup>TS</sup>
	Cell Death OFF, RNAi ON	ci-Gal4/RNAi-Dif;sal <sup>E/PV</sup> -LHG:tub-Gal80 <sup>TS</sup> /+
	Cell Death OFF, RNAi ON	ci-Gal4/RNAi-Stat92-E;sal <sup>E/PV</sup> -LHG:tub-Gal80 <sup>TS</sup> /+
	Cell Death OFF, RNAi ON	ci-Gal4/RNAi-Lilli;sal <sup>E/PV</sup> -LHG:tub-Gal80 <sup>TS</sup> /+
	Cell Death ON, RNAi ON	ci-Gal4/RNAi-Dif;lexO-rpr/sal <sup>E/PV</sup> -LHG:tub-Gal80 <sup>TS</sup>
	Cell Death ON, RNAi ON	ci-Gal4/RNAi-Stat92-E;lexO-rpr/sal <sup>E/PV</sup> -LHG:tub-Gal80 <sup>TS</sup>
	Cell Death ON, RNAi ON	ci-Gal4/RNAi-Lilli;lexO-rpr/sal <sup>E/PV</sup> -LHG:tub-Gal80 <sup>TS</sup>

**Annex II:**  
**Statistics and replicate analysis**



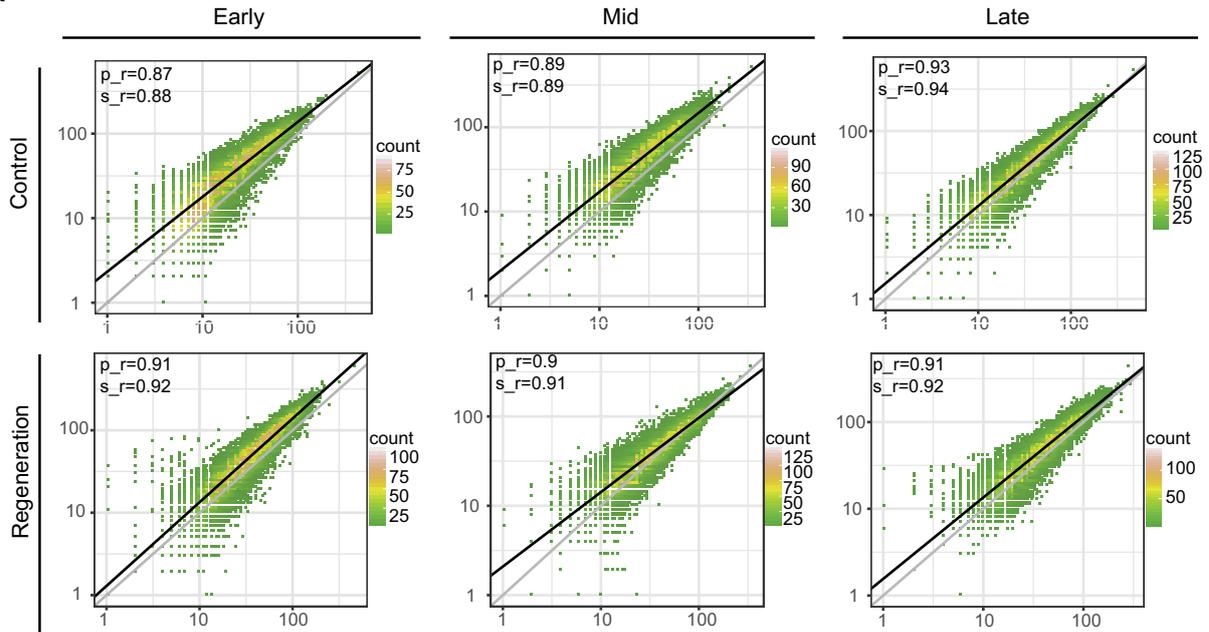
## Statistics and replicate analysis of RNA-seq



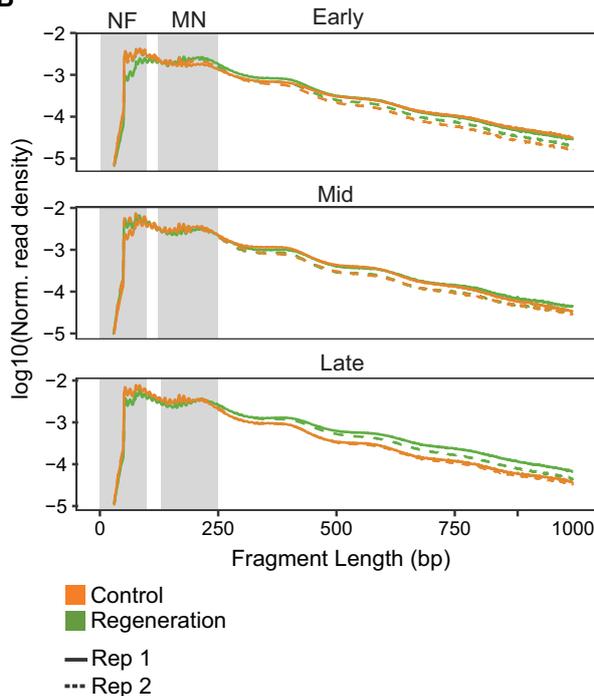
**Figure 70 - Statistics and replicate analyses of RNA-seq.** RNA-seq was performed from two independent biological replicates from each time point and condition. (A) Scatter plots showing high correlation of gene expression levels between replicates (Pearson and Spearman correlation coefficients higher than 0.9, denoted by  $p_r$  and  $s_r$ , respectively). (B) Mapped genomic reads were classified as: exonic if reads map entirely within exons, exonic-intronic if reads map both in exons and introns, intergenic if reads map outside genes and intronic if reads map entirely within a gene but not within annotated exons. Split reads were reads mapping to splice junctions. (C) RNA-seq mapping statistics. Number and proportion of mapped reads and unique mapped reads are shown. Most reads (98%) map to the exons.

# Statistics and replicate analysis of ATAC-seq

**A**



**B**



**C**

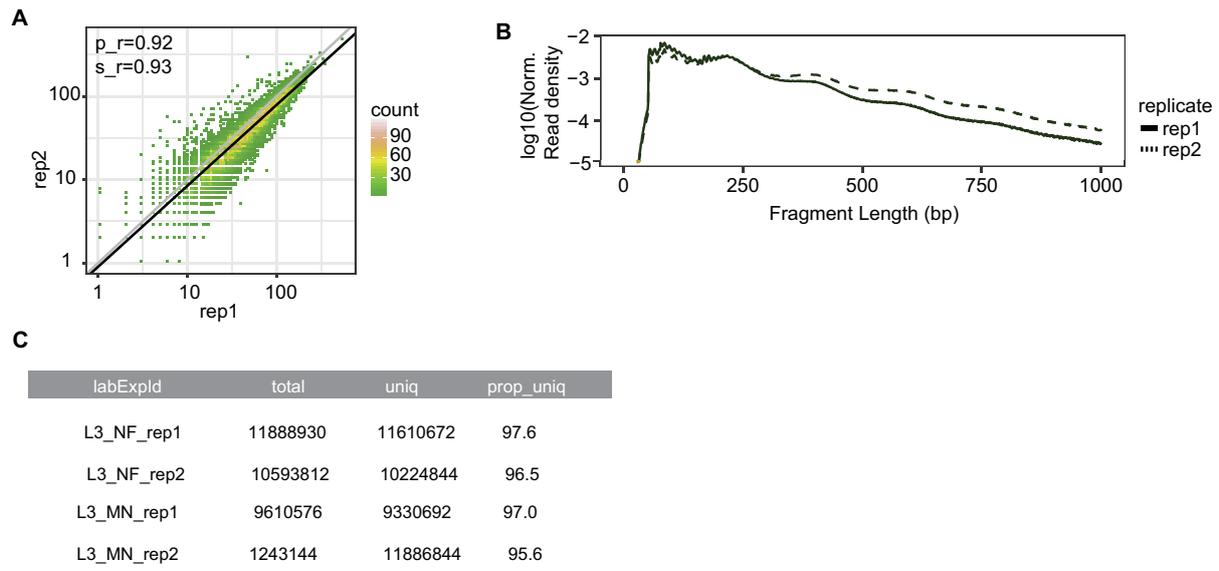
labExpId	total	uniq	prop_uniq
ER-ctrl_NF_rep1	5341626	5199086	97.3
ER-ctrl_NF_rep2	10242146	9975064	97.3
ER-reg_NF_rep1	10775260	10466302	97.1
ER-reg_NF_rep2	11857222	11559318	97.4
MR-ctrl_NF_rep1	8261020	7960502	96.3
MR-ctrl_NF_rep2	12779268	12368228	96.7
MR-reg_NF_rep1	9441288	9178988	97.2
MR-reg_NF_rep2	12362538	11997774	97.0
LR-ctrl_NF_rep1	9936622	9672454	97.3
LR-ctrl_NF_rep2	11242928	10976686	97.6
LR-reg_NF_rep1	10745466	10382710	96.6
LR-reg_NF_rep2	13657716	13253602	97.0

**D**

labExpId	total	uniq	prop_uniq
ER-ctrl_MN_rep1	10876938	110526732	96.7
ER-ctrl_MN_rep2	10286054	998006	97.0
ER-reg_MN_rep1	8023184	766414	96.7
ER-reg_MN_rep2	10384794	10065668	96.9
MR-ctrl_MN_rep1	7126620	6811380	95.5
MR-ctrl_MN_rep2	10716514	10293398	96.0
MR-reg_MN_rep1	12144392	11646416	95.8
MR-reg_MN_rep2	8268676	7992674	96.6
LR-ctrl_MN_rep1	11086670	10755750	97.0
LR-ctrl_MN_rep2	11732148	11411698	97.2
LR-reg_MN_rep1	8603818	8192000	95.2
LR-reg_MN_rep2	10699724	10326310	96.5

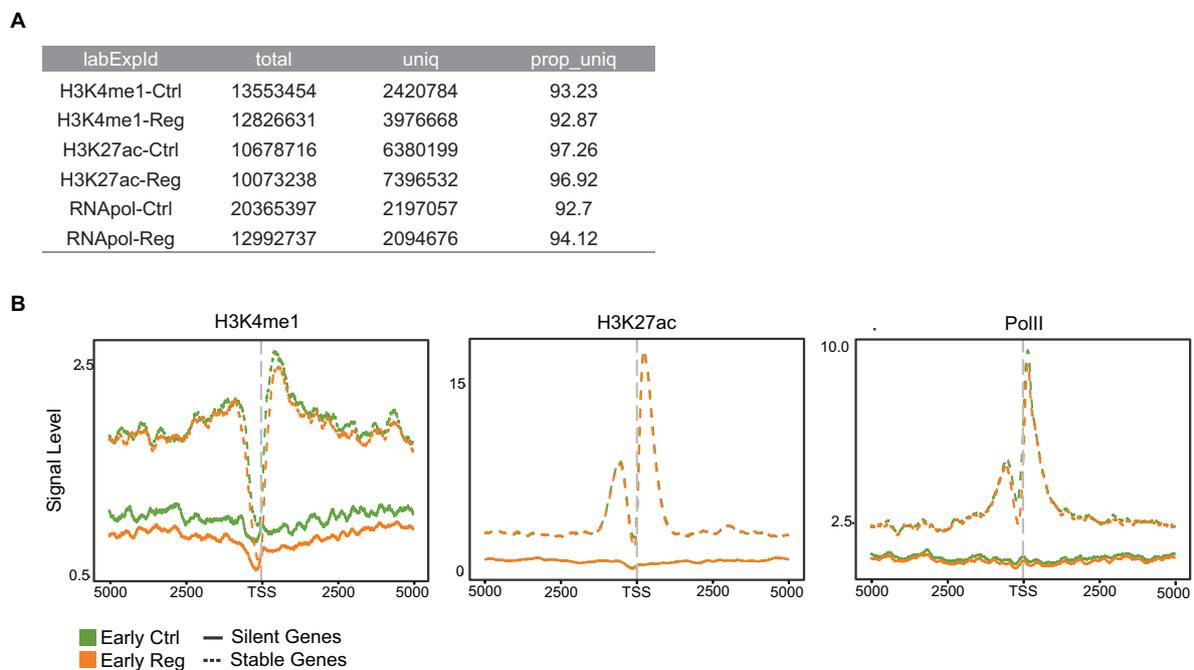
**Figure 71 - Statistics and replicate analyses of ATAC-seq.**

ATAC-seq was performed from two independent biological replicates from each time point and condition. (A) Scatter plots showing high correlation of peak heights between replicates (Pearson and Spearman correlation coefficients higher than 0.85, denoted by  $p_r$  and  $s_r$ , respectively). (B) Line plot showing read density per fragment length. Fragments belonging to nucleosome-free region (NF) fall in 0 to 100bp, meanwhile mononucleosome fraction (MN) fall in 180 to 247bp. (C) NF mapping statistics. (D) MN mapping statistics.



**Figure 72 - Statistics and replicate analyses of third instar larval ATAC-seq.** (A) Scatter plot showing high correlation of peak height between replicates in L3 ATAC-seq (Pearson and Spearman correlation coefficients higher than 0.9, denoted by  $p_r$  and  $s_r$ , respectively). (B) Line plot showing read density per fragment length. Fragments belonging to NF will fall in 0 to 100bp meanwhile MN fraction will fall in 180 to 247bp. (C) NF and MN mapping statistics for L3.

## Statistics and analysis of ChIP-seq



**Figure 73 - Statistics and analysis of ChIP-seq.** (A) ChIP-seq mapping statistics for early control and regeneration. (B) Quality check for ChIP-seq based on the signal level at the TSS of modEncode stable and silent genes. Silent genes should have flat signal whereas stable genes should have clear signal. Average profile of H3K4me1, H3K27ac and PolII around  $\pm 5000$  bp of the TSS of L3 wing disc modENCODE silent and active genes.

## Mapping statistics of shared regenerative genes

Orthologous genes ratios			
		Zebrafish	Mouse
1	Mapped fly genes	7357	7304
2	All fly orthologous genes up-regulated in	2476	2347
	Ratio 2/1	33,66%	32,13%
3	Up-regulated fly genes mapped	1288	1287
4	Up-regulated fly genes mapped to an up-regulated gene in	419	431
	Ratio 3/4	32,53%	33,49%

**Table 8 - Mapping statistics of regenerative genes.** Table showing ratios based on mapping statistics of shared regenerative genes.

**Annex III:  
Clusters and hotspots lists**



## Cluster list

Table showing the coordinates, the *p*-value and gene names of differentially expressed genes in each cluster.

Early Upregulated Clusters				
chr	Begin	End	<i>p</i> -value	Upregulated genes inside cluster
chr2L	273836	297449	0.0442808	CG11555 CG17078 smo Hop
chr2L	415067	425321	0.000625668	CG11885 CG13690 BBS8 CG13692 ebi
chr2L	810577	851071	0.0243328	Nnf1b Dbp21E2 CG3662 drongo
chr2L	2359142	2380506	0.0319127	Rab5 CG3609 Atxn7
chr2L	2869652	2954406	0.00672512	NTPase CG8813 CG8814 lilli
chr2L	2977122	2987723	0.0470252	CG3165 CG9643 Chd1
chr2L	3159950	3171092	0.0319127	gkt Hydr2 CG44002
chr2L	6022951	6053877	0.00128281	CG12393 CG9135 CG31643 CG9107 mtm ade2
chr2L	6920704	6948808	0.019811	Hrb27C Fgop2 ihog
chr2L	7986786	8026898	0.000625668	Wwox Tg CG12560 Sirup Spn28Dc pes
chr2L	8071812	8116531	0.0107667	Bsg Trf MED20
chr2L	8305586	8311195	0.0470252	CG7840 CG7818 CG7810
chr2L	8370454	8411868	0.0176077	CG17292 CG31898 CG13384 CG13397 Wdr82 Rcd4 l(2)k12914
chr2L	8515096	8541439	0.011079	ImgA CG17834 ImgB PIG-U
chr2L	8977834	8999816	0.011079	CG13101 alien rost Try29F
chr2L	9984645	10002736	0.00468345	CG31875 bib SoYb
chr2L	10054213	10063925	0.019811	Pen Spn31A Cpr31A
chr2L	10207085	10269904	0.00128281	CG5731 me31B CG5694 CG5708 CG4953 CG31717 Prosalpha6 CG4957 Ror eEF1delta
chr2L	10362989	10393228	0.00390894	pim lft Cdk1 mRpS7 da
chr2L	10730255	10740762	0.0470252	CG17118 Dpy-30L1 CG12299
chr2L	10987806	11001476	0.0470252	CG12253 CG16833 aub
chr2L	11095150	11113061	0.0319127	Ge-1 l(2)gd1 CG6201
chr2L	12046711	12066775	0.00367539	RpL7-like Jhl-21 CG14946 Plzf CG34164
chr2L	12693872	12707711	0.0243328	CG5780 CG15484 spict PICK1
chr2L	13165564	13182596	0.00390894	Sfmbt CG5287 Sir2 CG31849 CG5439
chr2L	13286777	13293415	0.019811	Uvrag CG31729 CG16824
chr2L	15044963	15070316	0.0107667	ck Tflls vig
chr2L	17410641	17474773	0.0319127	Sgt BicD Dif CG5050 CG15141 dl
chr2L	18700606	18711297	0.00367539	Grip71 CG10343 CG10376 Faf
chr2L	19114401	19179758	0.00233854	CG10561 l(2)37Cg Ddc l(2)37Cb brat
chr2L	20058163	20097253	0.011079	mRpS18B barr CG10730 CG10747 fok vls
chr2R	6653217	6715892	0.00283315	Opbp Mob4 CG3270 Hsepi Tdc1 sced CG3409 geminin CG45092
chr2R	7510622	7532981	0.00797182	CG2144 az2 Orc1 Drat
chr2R	8087396	8123001	0.00206005	Cul4 CG8712 sut2 CG11210 Asap
chr2R	8914524	8939789	0.00480942	lin CG8248 Spt CG13749 FANCI
chr2R	9841080	9850642	0.0130748	Orc6 Vamp7 CG1667 CG1671
chr2R	10096480	10121451	0.0122905	Jra 14-3-3zeta dmpd
chr2R	10420414	10465701	0.0007807	Prx2540-1 CG33474 CG11825 CG12895 Galphao
chr2R	11280484	11309829	0.038958	CG9062 Fbl6 CG18336 CG7745
chr2R	12138695	12144552	0.0225209	CG18343 CG8378 OSCP1
chr2R	13160063	13198337	0.0198641	GLaz Kdm4B Nrk CG33137
chr2R	13573618	13593339	0.0225209	CG42808 CG6191 CG45088
chr2R	13954195	14060356	0.0130748	CG42806 CG6701 Ctf4 mRpL53 mam CG33155
chr2R	18658418	18673536	0.0225209	CG42306 Gint3 edl

chr2R	18833372	18862078	0.0130748	CG15098 CG18190 Jheh1 Jheh3 CG15084
chr2R	19430659	19485686	0.0130748	EloC Fak par-1 CG7744
chr2R	21666687	21691139	0.000219251	CG30285 CG42365 NC2alpha CG42497 Tim10 Tbp CG10307 CG10306
chr2R	24092079	24114292	0.0122905	CG4049 gek enok
chr3L	593246	606569	0.0283603	Reg-2 CG13893 MED14
chr3L	3803832	3810196	0.00687162	ntc IntS10 CG32262
chr3L	5803033	5813093	0.00687162	mad2 CG5537 CG42272
chr3L	7321099	7353622	0.0220256	RhoGEF4 mus312 lark CG14834 qm
chr3L	7970588	7977849	0.0283603	HP4 RNaseX25 CG8209
chr3L	8340314	8356296	0.0283603	DNApol-alpha50 CG7083 CG17352
chr3L	8402168	8472669	0.00116251	Cbl mkg-p CG33057 CG32354 Unr Gug CG13667
chr3L	8718949	8726108	0.0451377	SrpRbeta CG32022 CG6511
chr3L	9446864	9456281	0.0376634	CG3529 CG3448 ghi phol
chr3L	9834367	9859889	0.000233146	Cdk8 CG10809 RasGAP1 I-2
chr3L	10623066	10640136	0.0376634	CG42521 E(z) CG8009 CG18628
chr3L	11061571	11073348	0.0265021	Blos2 CG32075 CG6310 Mocs1
chr3L	11693879	11728118	0.0175921	crim CG11652 CG14132 CG11658 Bmcp
chr3L	11820338	11833578	0.0451377	CG11597 CG32095 CycA
chr3L	12116101	12138584	0.00438221	ssp yps Atg12 ND-SGDH I(3)2D3 Lsp2
chr3L	14043694	14083229	0.00284456	CG13484 CG6833 Pex1 blue btl
chr3L	14747929	14761844	0.0283603	CG42507 Trl CG33260
chr3L	16085221	16112626	0.0376634	Taspase1 CG5235 CG5414 Zn72D
chr3L	17621824	17654280	0.0175921	Ndfip Krn CG43085 CG7510
chr3L	18112959	18145418	0.0283603	CG32195 CG7341 geko
chr3L	18745856	18801309	0.0283603	MED11 Atg3 ftz-f1
chr3L	19798975	19836413	0.00599033	CG14103 CG14102 trc CG34116 CG32221 Deaf1 Fibp
chr3L	20353449	20371746	0.0175921	CG5618 DNApol-alpha60 in CG13247
chr3L	22721203	22748614	0.0265021	mael CG11367 CG11241 CG7369
chr3R	4314455	4339914	0.0358337	Syt14 CG9795 CG1103
chr3R	5633627	5656336	0.0281829	CG12171 CG31546 MTA1-like MED27
chr3R	6350896	6380328	0.0358337	Pak CG42564 CG1024
chr3R	8712933	8735837	0.0385776	CG8112 nom M1BP CG8202
chr3R	8990645	9020022	0.00739358	E(var)3-9 CG11975 CG11983 CG11986 tgo
chr3R	9330885	9337459	0.0223317	CG8121 CG8866 CG16749
chr3R	9561583	9591002	0.0385776	CG9393 CG16789 CG9399 bocks
chr3R	9742877	9768984	0.00176352	CG31415 Snap24 MED6 PpD3 CG8412 CG34409
chr3R	10854501	10882188	0.000389722	CG14694 Tpc1 CG4511 Sodh-2 SelR CG4570 CG6574
chr3R	11620090	11650706	0.00788735	Csk CG14712 CG14710 CG18764
chr3R	11877235	11949129	0.0358337	sad mthl5 Cad87A
chr3R	12348987	12376572	0.000659951	GstD9 GstD10 CG10038 GstD1 GstD5 GstD3 CG10041 GstD4
chr3R	14302428	14320494	0.0223317	CG3259 RpII15 pr-set7
chr3R	15214762	15246614	0.00739358	CG31344 tefu Caf1 CG12241 Hsc70-4 SIDL
chr3R	16342318	16359466	0.000318671	Manf CG14879 CG17931 pad CG10311
chr3R	16432998	16458233	0.00739358	ema CG14881 CG14883 CG14882 CG17565
chr3R	18400373	18545586	0.0223317	CG12333 CG7685 fru
chr3R	20617883	20644279	0.0121845	CG5466 Pi3K92E Lrrk
chr3R	21362968	21373575	0.0358261	pre-mod(mdg4)-AB pre-mod(mdg4)-AE pre-mod(mdg4)-X pre-mod(mdg4)-AA pre-mod(mdg4)-O
chr3R	22375923	22392349	0.0385776	CG18596 CG34149 CG43342 CG7059

chr3R	22411301	22424929	0.0121845	lqfR mats pinta
chr3R	24760530	24814012	0.0358337	CG5807 Ude CG5805 AstA CG13631
chr3R	24871202	24894605	0.0223317	CG3744 CG31381 CG31121
chr3R	25053369	25102146	0.00285448	CG11920 CG11836 CG33095 CG31111 CG9996 CG31109
chr3R	25997621	26098711	0.000579755	CG42498 E(spl)m2-BFM E(spl)m4-BFM E(spl)m3-HLH E(spl)m6-BFM E(spl)malpha-BFM CG14550 dys E(spl)mdelta-HLH E(spl)m5-HLH E(spl)mbeta-HLH CG14551 gro
chr3R	26842462	26871995	0.00203738	Lerp IntS12 His2Av BM-40-SPARC
chr3R	27256595	27282661	0.0196619	woc CG14262 CG5934 mrt
chr3R	29928701	29934587	0.0196619	CG42558 Cog7 alpha-Man-Ic CG42557
chr3R	31236035	31254900	0.0358337	qlless mRpL32 CG1750
chrX	1329518	1357601	0.0346818	Tsp2A png CG12773 CG11409
chrX	1890256	1930517	0.0227564	mRpL16 arm CG32801 Rbcn-3B CG11596 Ocri
chrX	2617758	2679553	0.00140475	Tsp3A Seipin PI4KIIIalpha sgg
chrX	5690002	5738031	0.0346818	cv CG4096 CG32758
chrX	7725996	7739978	0.00377028	Tom40 NELF-B CG12155 Rab39
chrX	9216872	9232378	0.0141067	CG12121 CG15369 CG15370 t
chrX	10764878	10824277	0.000691772	CG15211 Atg8a Imp BTBD9 Ant2
chrX	11089290	11094947	0.0346818	Vago CG2076 CG2061
chrX	11144120	11164670	0.0340111	CG15201 Ran Lint-1 Dlic
chrX	11355881	11409776	0.000221981	hop dlg1 dsh Pa1 Tim8
chrX	11561643	11573640	0.00960545	CG11699 PGRP-SA Kmn1 CG11697 RpII215
chrX	11792533	11837460	0.019226	CkIIbeta CG1578 rudhira
chrX	11990118	12020719	0.00787497	CG1492 CG18130 CG1806 CG15735
chrX	13288051	13329303	0.0346818	CG15744 CG1622 IP3K2
chrX	15446179	15466812	0.0227564	Scamp shtd CG6294 CG6299
chrX	15697494	15712309	0.0340111	CG9281 Pis CG8128 CG15601
chrX	16385278	16399850	0.0346818	CG4239 mei-41 TH1
chrX	16422831	16448302	0.0346818	hang rngo CG34015
chrX	18078491	18101314	0.0346818	CG7536 dik CG7192
chrX	18843663	18888860	0.00787497	CG34401 CG34422 CG7332 CG7326 Usp39
chrX	19676061	19701755	0.000120922	Ubqn CG14227 CG14231 dome Cdc42 CG14229
chrX	21023390	21041467	0.019226	r-cup CG1529 Ntf-2

#### Early Downregulated Clusters

chr	Begin	End	p-value	Downregulated genes inside cluster
chr2R	8430115	8436095	0.000125486	Lcp1 Lcp3 Lcp2
chr2R	9237512	9253255	2,31E+00	Cyp4p2 Cyp4p1 CG30343
chr3L	9371511	9374964	0.000240928	Hsp67Bb Hsp67Bc Hsp22
chr3R	12502510	12511461	5,30E-01	Hsp70Bb Hsp70Bc Hsp70Bbb

## Mid Upregulated Clusters

chr	Begin	End	p-value	Upregulated genes inside cluster
chr2L	252589	282167	0.00484885	CG3645 mbm CG17078 smo
chr2L	404285	418536	0.0171815	Rpl135 AP-2alpha ebi
chr2L	1129316	1166425	0.0154686	MFS3 Vps29 IntS14 CG4552 Tfb4
chr2L	2372455	2384747	0.0485001	CG3609 CG9870 CG15390
chr2L	4361925	4386062	0.010482	Traf4 CG17612 CG3338
chr2L	4442870	4453038	0.0362114	morgue Elp3 CG15439
chr2L	4821774	4850419	0.0257572	CG15628 tank mRpS2
chr2L	4955509	4981592	0.0432287	Marcal1 CG34125 mRpL28 CG8892
chr2L	5521281	5546642	0.0272583	Cap-D3 CG7371 CG6907 Lam
chr2L	6908083	6918743	0.010482	nop5 Wee1 Rat1
chr2L	6945464	6964074	0.00221584	CG13775 SA sip2 ihog CG3430
chr2L	7426866	7445882	0.00559711	CG5973 CG5261 CG5958
chr2L	7986786	8007210	0.0362114	pes Wwox Spn28Dc
chr2L	8363477	8382807	0.00676446	CG13392 Aats-ala Wdr82 Pp2A-29B RpS13
chr2L	8509069	8528618	6.17E+00	CG13097 ImgA CG13090 ImgB CG13096 PIG-U
chr2L	9954325	9968645	0.0362114	CG4709 Dref RpL13
chr2L	10302173	10323102	0.011051	CG4972 CG5381 Usp14 nmd
chr2L	10373369	10393228	0.0485001	CG33303 CG5096 da
chr2L	10970443	10997897	0.0362114	YL-1 CG16833 abo
chr2L	12690336	12712952	0.011051	CG5776 CG5787 CG6153 CG5525
chr2L	13812700	13831576	0.0362114	CG16888 Arpc1 Orc5
chr2L	15749319	15768349	0.0257572	ND-B17 l(2)35Df wek
chr2L	19110154	19131954	0.00753788	l(2)37Cc Aats-asn amd l(2)37Cb
chr2L	19435850	19453490	0.0171815	CG10237 RanGAP Top2
chr2L	20058163	20096253	0.00484885	barr nesd pr neb lok
chr2L	21084605	21102429	0.0362114	CG33509 ppk13 CG12050
chr2L	21154293	21182812	0.0272583	E2f2 Nbr CG9246 Mcm10
chr2R	5754517	5774324	0.00352019	CG7791 l(2)09851 Gp210
chr2R	6074471	6085625	0.00815841	Ars2 CG14590 CG7845
chr2R	6633647	6656594	0.00254924	Trap1 Debcl Opbp geminin
chr2R	6693430	6715892	0.0364891	CG3270 CG3409 CG45092
chr2R	7440706	7477162	0.00398605	CG11125 Aldh-III CG11123 sPLA2 Inos
chr2R	7496702	7539696	0.000130792	CG2144 CG1603 CG1598 Orc1 dpa Drat CG1602
chr2R	7791901	7815486	0.02457	Gapdh1 CG1550 CG1882
chr2R	8102260	8156654	0.000814588	Nup50 Cul4 coil Pbp49 kermit Obp44a
chr2R	8421033	8447164	0.0120292	CSN7 Lcp4 Lcp1 Cyp4e2
chr2R	8898729	8951086	0.00202928	CG8243 CG8229 FANCI CG8230 Pgi ana2 CG30349 CG8258 CG8235 MrgBP
chr2R	10063385	10084210	0.0364891	CG2292 cdc2rk egr
chr2R	10875748	10886375	0.02457	nclb Taf5 CG18004
chr2R	11886067	11895762	0.00815841	Ef1alpha48D cuff ERp60
chr2R	12453190	12471722	0.00352019	CG8545 dgt5 Lac
chr2R	13580313	13596557	0.02457	fand CG45088 CG6191
chr2R	14243778	14270658	0.00254924	Hsc70-5 CG8503 SelD CG8531
chr2R	16322186	16347500	0.0364891	CG4282 CG7997 wcd
chr2R	17420407	17430963	0.0364891	Bap55 CG6550 l(2)k01209
chr2R	20550334	20562331	0.02457	CG13430 lms Mgat1

chr2R	21658770	21688545	8,99E+00	CG30285 CG9865 CG42380 CG42379 NC2alpha CG42381 Rae1 Tbp CG42364
chr2R	22636063	22661389	0.0102752	RpS24 Ugt58Fa CG43326 bonsai Cdk9 CG42565
chr2R	23055474	23074725	0.00118431	PPO3 CG44252 I(2)k09913 Fib CG9890
chr2R	23847161	23877442	0.00853629	Thiolase eIF6 gbb CG5569 ytr ken
chr2R	24133989	24186343	0.0151325	Nop60B spag Zfrp8 CG3328 Dat
chr3L	677879	699144	0.03057	CG17129 ebd1 CG13894
chr3L	1534929	1576330	0.0204825	Pex10 CG12099 CG12091 CG12004
chr3L	1862881	1891462	0.000115796	Tmhs dre4 CG12025 CG13937 HBS1
chr3L	3145405	3172545	0.00948528	CG14963 CG15812 CG32281 Asciz Usp5
chr3L	3899259	3935617	7,09E+00	ida Ubi-p63E mge Sc2 Eip63F-1
chr3L	4134860	4152473	0.0204825	CG1299 Rop RfC4
chr3L	4242734	4288349	0.0204825	CG1309 TflIEbeta pav CG11583 CG1311 CG15011
chr3L	8110368	8151813	0.03057	Uba2 CG13676 ldr CG7927
chr3L	8176135	8191118	0.0427896	Nmt ERR Atg18a
chr3L	8340314	8360906	0.0125519	DNApol-alpha50 Idh CG7182
chr3L	8402168	8422205	0.03057	mkg-p Exo70 CG32354
chr3L	9334332	9358935	0.0125519	PGRP-LA PGRP-LC UGP
chr3L	9717786	9733807	0.0138402	CG18178 CG14174 defl CG6749
chr3L	10623066	10640136	0.00186422	CG18628 CG8009 CG42521 E(z) hay
chr3L	11070526	11100578	0.03057	Mocs1 CG7839 APP-BP1
chr3L	13009399	13044234	0.0192817	SRm160 CG11267 Zmynd10 CG11261 ste14 RpS4
chr3L	16415292	16496560	0.00673285	CG4573 CG33158 Aats-tyr
chr3L	16583028	16615472	0.0192817	spd-2 CG32164 Rpn12 Prosbeta6 Mo25
chr3L	19544345	19584568	0.03057	I(3)76BDr CG9300 Shal
chr3L	19589072	19612199	0.0427896	CG14100 Taf6 Lon
chr3L	19912405	19927175	0.00613081	Rpn1 Su(z)12 Grasp65 Mtr3
chr3L	20356807	20371746	0.0427896	kin17 DNApol-alpha60 CG13247
chr3L	21821997	21836685	0.0427896	CycH CG7148 Nopp140
chr3L	22263369	22292810	0.0204825	Ddx1 CG11523 Aats-ile
chr3R	5589601	5601136	0.018689	Vha26 kra noi
chr3R	8240747	8279274	0.0011648	RpA-70 CG9636 CG9630 ato CG33722 Mcm2
chr3R	8801564	8827279	0.0157157	Tcp-1eta CG9839 CG8369 CG8379
chr3R	9763650	9789323	0.0413836	CG8478 MED6 MtnA CG8500
chr3R	10766622	10791150	0.00828791	Art1 TflIFbeta Mcm5 CG45076 mRpL37
chr3R	10853374	10878257	0.0011648	CG6567 CG4511 Sodh-2 SelR CG4570 CG6574
chr3R	11650971	11679690	0.018689	Elp1 CG14715 Taf12
chr3R	12352143	12374824	0.0161674	MBD-R2 GstD9 GstD1 GstD2 GstD4
chr3R	12403340	12429282	0.0312003	Cyp9f2 CG5167 Pp1-87B CG5641
chr3R	12456573	12465298	0.00339803	CG12267 CG5608 CG5961 trus
chr3R	12502510	12511461	0.0101392	Hsp70Bc Hsp70Bbb Hsp70Bb
chr3R	14892392	14923083	0.0157157	CG7265 Nup93-2 CG3817 Sdr
chr3R	15209204	15230516	0.00558043	mRpS10 Caf1 CG4203 CG12241 SIDL
chr3R	15343653	15367665	0.0117863	CG34404 CG42727 CG31301 CG42726 Surf4
chr3R	16264271	16304768	0.0101392	Mat89Ba asun gish
chr3R	16432998	16445010	0.0226603	CG14881 ema CG14882 CG17565
chr3R	16640205	16649978	0.0445141	CG10324 nonA-I Cctgamma
chr3R	17532325	17546531	0.000578437	Prx3 CG5823 P5cr-2 I(3)07882
chr3R	18216142	18242159	4,38E-01	mTerf5 Non3 CG7183 CG7988 CG7156 Mdh2 CG14314
chr3R	20535323	20564067	0.018689	CG10877 Stat92E att-ORFA CG5180

chr3R	22408566	22423302	0.0445141	CG13850 Nop56 mats
chr3R	22687542	22700307	0.018689	wfs1 Nup133 Cyp6d4
chr3R	23752876	23767488	0.00356271	Rpt5 CG10217 CG10214 Lsd-1 Plip
chr3R	24128049	24167763	8,93E-02	Apc2 CG5463 Kal1 Slimp Sec10 Tsc1 p38a p38c
chr3R	24913216	24934919	0.0312003	PIG-S sud1 Mink CG11771
chr3R	25117218	25131155	0.0011648	Mocs2 CG31510 vig2 Clbn CG42503
chr3R	26041736	26059267	0.00558043	CG42498 gro CG14550 RpL34a CG14543
chr3R	26715165	26725216	0.0101392	CG6425 CG6420 CG5484
chr3R	27913777	27924888	0.0101392	CG5003 CG33213 RpL4 CG12259
chr3R	27948711	27988384	0.00558043	CG4884 Gp93 CG4849 betaTub97EF CG4951
chr3R	29034573	29062975	0.0413836	Cpsf100 Slu7 CG11837 Cul5
chr3R	29116391	29130972	0.018689	dgt6 Atg14 yem
chrX	1028291	1047009	0.0146617	CDC45L CG14630 CG42259
chrX	1457175	1472325	0.0402943	Lrpprc2 Ns3 pck O-fut2
chrX	4097592	4117487	0.0493901	CG6379 Nsun2 Tip60
chrX	6276876	6301285	0.0310991	Ubi-p5E Rpt4 CG3815 CG12219
chrX	6674387	6698567	0.0115438	Pink1 Mcm6 CG14440 CG3184
chrX	7725996	7739978	0.0146617	Tom40 NELF-B CG12155
chrX	8117012	8137046	0.00243792	CG2116 CG2120 CG2260 CG1575
chrX	8461479	8469012	0.0354244	Es2 Sprt CG12116
chrX	10852446	10877893	0.0146617	feo sofe CG1637
chrX	15328342	15344625	0.0238295	dah CG9123 CG12608
chrX	16296671	16308667	0.0354244	CG9914 CG3632 CG3679
chrX	18085761	18116922	0.0354244	dik Taf8 wupA
chrX	18648182	18657286	0.0146617	CG6659 lng3 CG6540
chrX	19475063	19509350	0.0232709	CG12200 Nat1 ND-18 Pfrx
chrX	19731004	19750167	0.000769137	Rpp20 COX6B AP-1-2beta Alr CG33932
chrX	20148124	20173514	0.00243792	CG9572 Phf7 AnxB10 CG9577 CG9581
chrX	20239855	20251093	0.0238295	Peritrophin-A obst-A CG17068

#### Mid Downregulated Clusters

chr	Begin	End	p-value	Downregulated genes inside cluster
chr2L	16292010	16302954	0.000119925	GMF CG13258 Cyp303a1
chr2R	7823943	7830429	0.000282583	CG18853 CG30383 phr
chr3L	9358999	9374964	0.000573051	CG32039 Hsp67Bc Hsp67Bb
chr3R	21354445	21373575	4,11E-03	pre-mod(mdg4)-J pre-mod(mdg4)-O pre-mod(mdg4)-N pre-mod(mdg4)-P pre-mod(mdg4)-G pre-mod(mdg4)-Y
chr4	1172469	1196848	0.00211495	pho CG33521 PIP4K

Late Upregulated Clusters				
chr	Begin	End	p-value	Upregulated genes inside cluster
chr2L	3692123	3713105	0.000627931	CG31955 CG16712 CG16704 CG16713 CG31777
chr2L	8190297	8215239	0.0179706	CG34134 CG8360 CG8353
chr2L	10681030	10689271	6,11E+00	CG17105 CG7296 CG7299 CG7294
chr2L	13965012	13976769	0.000107351	NimC1 He NimB3 NimB2
chr2L	14737090	14745210	0.000267612	CG34166 CG42586 CG31775 CG42587
chr2R	6693430	6715892	0.00373669	CG3409 CG3270 CG45092
chr2R	7042656	7057548	0.0129285	Tsp42Ei Tsp42Ep Tsp42Ek
chr2R	8430115	8447164	2,93E+00	Lcp1 Lcp4 Cyp4e2 Lcp3 Lcp2
chr2R	11248126	11275286	0.0012476	CG34227 Cpr47Eb Cpr47Ec Cpr47Ee TpnC47D
chr2R	17136451	17140865	0.00235815	CG43107 CG43103 Gbp CG11395
chr2R	22640526	22668848	3,01E-02	RpS24 CG4250 Ugt58Fa CG43326 CG3746 CG42565 CG30196 CG30269 Cyp6d2
chr2R	23356372	23360266	0.00751336	CG3500 CG34423 CG34424
chr3L	3291459	3317144	0.00255298	CG12009 CG14968 Drsl5
chr3L	8202792	8222953	0.00124375	CG13679 CG8012 CG13678
chr3L	16271282	16303178	0.000291458	CG13044 CG13067 CG13046 CG4962 CG13045
chr3R	11983098	12003516	0.00103978	mfas Ect3 CG3397
chr3R	12360008	12408107	1,14E+00	GstD2 GstD9 GstD10 CG5167 GstD1 CG4115 lig3 Cyp9f2 GstD4
chr3R	12682186	12697379	0.0040073	NijC CG31347 kar
chr3R	21354445	21370249	0.000996138	pre-mod(mdg4)-W pre-mod(mdg4)-AD pre-mod(mdg4)-C pre-mod(mdg4)-B pre-mod(mdg4)-I
chr3R	24859900	24867651	8,10E+00	CG13641 CG13640 CG7016 Elal CG13639
chr3R	25291370	25316116	0.0121753	CG31098 CG31370 CG10559
chrX	1031496	1052631	0.000553943	CG14630 CG42259 CG14629
chrX	1464363	1483413	0.00165064	Rab27 CG14777 sta O-fut2
chrX	10839230	10848534	0.00182998	CG15209 CG17333 CG15210
chrX	11555050	11573640	0.000596142	CG11699 PGRP-SA CG1572 CG11697
chrX	15450261	15475400	0.00769044	Ahcy13 CG15642 mh

Late Downregulated Clusters				
chr	Begin	End	p-value	Downregulated genes inside cluster
chr2L	7030493	7065837	0.000439812	CG31907 Nuf2 Mnn1 Rab30
chr2L	10207085	10261049	6,45E-01	CG5731 mthl15 Pten me31B CG5694 CG5708 eEF1delta
chr2L	10362989	10386262	0.00618328	pim lft Cdk1
chr2L	10414017	10441047	0.00123746	Klp31E TflIB Nse4 CHIP
chr2L	21614232	21647964	0.000439812	CG2201 Df31 Ac3
chr2R	5339706	5354723	0.000106304	CG10465 CG10395 COX4L
chr2R	7823943	7828678	0.00199695	Prosalph1 CG30383 phr
chr2R	12593604	12623197	0.000259887	Amph CG45086 Galphaq
chr3L	7357275	7373711	0.0039058	pst akirin Sh3beta
chr3L	8424954	8449377	0.000214337	Cbl Unr CG6983
chr3L	9371511	9382765	0.0147055	Hsp67Bb Hsp67Bc Hsp23
chr3L	15818734	15837482	0.000128239	dbo CG15715 CG18081 DCP2
chr3L	21507852	21533713	0.00267853	M6 SAK Mkrn1
chr3R	21355692	21373575	0.00216529	pre-mod(mdg4)-J pre-mod(mdg4)-X pre-mod(mdg4)-O
chr4	1172469	1224467	0.00258956	CG33521 Mitf PIP4K pho
chrX	15841769	15866175	0.00150001	Aats-arg AlkB Gbeta13F

## Hotspots list

Table showing the coordinates, the p-value of hotspots.

Early Upregulated Hotspots			
chr	Begin	End	p-value
chr2L	273836	851071	0.338655
chr2L	2359142	3171092	0.0693084
chr2L	7986786	8999816	0.1844
chr2L	9984645	11113061	0.192654
chr2L	12693872	13293415	0.192654
chr2R	9841080	10465701	0.192654
chr2R	13160063	14060356	0.192654
chr2R	18658418	19485686	0.0693084
chr3L	7970588	8726108	0.48233
chr3L	11061571	12138584	0.0693084
chr3R	8712933	9768984	0.09532
chr3R	11620090	12376572	0.338655
chr3R	24760530	25102146	0.48233
chrX	10764878	12020719	0.0211217
chrX	15446179	16448302	0.313069

Mid Upregulated Hotspots			
chr	Begin	End	p-value
chr2L	252589	1166425	0.489332
chr2L	4361925	5546642	0.029133
chr2L	6908083	8528618	0.344619
chr2L	9954325	10997897	0.801204
chr2L	19110154	20096253	0.344619
chr2R	5754517	8951086	0.0710268
chr3L	3145405	4288349	0.0710268
chr3L	8110368	9358935	0.443173
chr3L	19544345	20371746	0.196708
chr3R	10766622	12511461	0.319839
chr3R	14892392	15367665	0.196708
chr3R	16264271	16649978	0.344619
chr3R	24128049	25131155	0.616452
chrX	7725996	8469012	0.196708
chrX	19475063	20251093	0.0710268

Late Upregulated Hotspots			
chr	Begin	End	p-value
chr3R	11983098	12697379	0.0348965

Late Upregulated Hotspots			
chr	Begin	End	p-value
chr2L	10207085	10441047	0.012161

**Annex IV:  
DRRE motif discovery through time**



Tables showing all the TFs putatively binding to an enriched motif in DRREs at the different time points. For the motifs found at the early stage, TFs that are upregulated in regeneration are shown in bold and TFs that are unique to each DRRE and underlined.

Early stage: iDRRE			Early stage: reused eDRRE		
	TF	Score		TF	score
1	<b>trl</b>	9,95	1	<b>trl</b>	5,47
2	<b>in, fd68A, jum, CHES-1-like</b>	7,2	2	<b>Homeodomain</b>	4,76
3	<b>C2H2-ZF</b>	7,06	3	Mad	4,36
4	grh	6,59	4	<b>C2H2-ZF</b>	4,34
5	CG5245, CG4360, Meics	6,45	5	Adf1	4,22
6	<b>Homeodomain</b>	6,12	6	<b>Aef1</b>	4,22
7	gem	6	7	dm	4,17
8	<b>Myb, zfh1</b>	5,25	8	inv, en	4,06
9	foxo	5,25	9	<b>sd, Clk</b>	4,04
10	Psi	5,11	10	<b>mof</b>	3,77
11	E(bx)	5,06	11	CG4360	3,77
12	<b>DNApol-iota</b>	5	12	opa	3,72
13	Cnx99A	4,99	13	<b>crp</b>	3,71
14	Tet	4,93	14	lola	3,69
15	<b>Kdm2</b>	<b>4,65</b>	15	<b>caup</b>	3,67
16	<b>Rpd3</b>	4,53	16	pnt	3,64
17	<b>tgo</b>	<b>4,29</b>	17	<b>fd68A</b>	3,57
18	sima	4,28	18	<b>Stat92E</b>	3,55
19	Adf1	4,22	19	xbp1	3,3
20	<b>E2F1</b>	4,21	20	<b>zfh1</b>	3,13
21	Mad	4,17	21	<b>Max</b>	3,11
22	Sox102F	4,16			
23	<b>jim</b>	4,15			
24	nej, Sirt6	4,09			
25	CG4360	4,08			
26	G9a	3,95			
27	<b>Aef1</b>	3,91			
28	<b>Taf1</b>	<b>3,71</b>			
29	lola	3,71			
30	trx	3,3			
31	usf	3,29			
32	<b>Dif, dl</b>	3,28			
33	CrebB	3,24			
34	<b>dpn</b>	3,23			
35	pnt	3,21			
36	Hsf	3,11			
			Early stage: novel eDRRE		
	TF	Score		TF	score
1	grn, <b>pnr, pho, phol</b>	5,01	1	grn, <b>pnr, pho, phol</b>	5,01
2	CG12236	4,91	2	CG12236	4,91
3	<b>Stat92E</b>	4,67	3	<b>Stat92E</b>	4,67
4	pan	4,57	4	pan	4,57
5	<b>CG8319</b>	4,44	5	<b>CG8319</b>	4,44
6	<b>da</b>	4,41	6	<b>da</b>	4,41
7	<b>Deaf1</b>	4,39	7	<b>Deaf1</b>	4,39
8	hth	4,26	8	hth	4,26
9	CG17829	4,21	9	CG17829	4,21
10	<b>pzg</b>	4,16	10	<b>pzg</b>	4,16
11	<b>ftz-1</b>	4,16	11	<b>ftz-1</b>	4,16
12	Awh	4,03	12	Awh	4,03
13	E5, ems	3,93	13	E5, ems	3,93
14	<b>homeobox</b>	3,76	14	<b>homeobox</b>	3,76
15	foxo	3,72	15	foxo	3,72
16	TfIIIB	3,6	16	TfIIIB	3,6
17	<b>p53</b>	3,57	17	<b>p53</b>	3,57
18	esg	3,44	18	esg	3,44
19	<b>vis, achi</b>	3,35	19	<b>vis, achi</b>	3,35
20	<b>unpg</b>	3,25	20	<b>unpg</b>	3,25
21	<b>Max, crp</b>	3,23	21	<b>Max, crp</b>	3,23
22	CG17209	3,22	22	CG17209	3,22
23	<b>svp</b>	3,02	23	<b>svp</b>	3,02

## Mid Stage: iDRRE

	TF	score
1	trl	9,33
2	mad	6,54
3	adf1	6,42
4	fd68A, jumu, CHES-1-like	6,41
5	C2H2 ZF	5,36
6	CG5245, CG4360, Meics	5,28
7	grh	5,21
8	Myb, zfh1	4,99
9	gem	4,93
10	Usf	4,89
11	E(bx)	4,78
12	DNApol-iota	4,7
13	foxo	4,66
14	Taf1	4,64
15	Rpd3	4,4
16	Kdm2	4,19
17	lola	4,07
18	jim	3,91
19	Aef1	3,9
20	E2f1	3,69
21	ewg	3,51
22	tj	3,34
23	odd	3,32
24	trx	3,1

## Mid Stage: reused eDRRE

	TF	score
1	aef1	9,13
2	CG11504	8,64
3	luna, dar1, CG42741	7,84
4	CG5846	7,54
5	snf	7,04
6	Hsf	6,61
7	tj	6,09
8	Pdi	5,85
9	Adf1	5,76
10	C2H2 ZF	5,72
11	crp	5,48
12	pzg	5,3
13	usf	5,13
14	vis, achi	4,8
15	CTCF	4,73
16	sin3A	4,58
17	sd	4,51
18	wcd	4,09
19	fd68A	4,06
20	mad	3,68
21	tap	3,22

## Mid Stage: novel eDRRE

	TF	score
1	vis, achi	4,06
2	homeobox	4,05
3	tj, E2f1	3,92
4	bow1	3,82
5	smox	3,7
6	CG6276	3,65
7	Mes2	3,61
8	Dif,dl	3,52
9	C2H2 ZF	3,45
10	sry-beta	3,41
11	gem	3,33
12	usf	3,24
13	TfIIA-S, Tbp, Trf, Trf2, TfIIIB	3,14
14	tap	3,14
15	dpn	3,09
16	CG5245, CG4360, Meics	3,07
17	crol	3,01
18	nub, foxo	3,01
19	CG42741	3

Late Stage: iDRRE

	TF	score
1	Trl	6,63
2	E(bx)	5,49
3	grh	5,32
4	Myb, zfh1	4,81
5	Psi	4,77
6	gem	4,58
7	jim	4,41
8	bol	3,71
9	usf	3,69
10	Sox15	3,65
11	adh	3,54
12	fd68A, jummu, CHES-1-like	3,48
13	rn	3,41
14	nej, Sirt6	3,36
15	Adf1	3,2
16	foxo	3,19
17	lola	3,14
18	bs	3,12
19	CTCF	3,05
20	ewg	3,02

Late Stage: Novel eDRRE

	TF	score
1	CG9727	5,09
2	E2f1	4,36
3	Sin3A	4,27
4	Ets98B	3,48
5	zfh2	3,41
6	su(H)	3,4
7	CG7101, CG123	3,36
8	Stat92E	3,19
9	tgo	3,19
10	ttk	3,18
11	sd	3,17
12	pnr	3,14
13	nub	3,13
14	kn	3,07
15	Hsf	3,07

Late Stage: reused eRRE

	TF	score
1	luna, dar1, CG42741	4,86
2	z	4,64
3	wcd	4,59
4	kay	4,42
5	trl	4,24
6	Hr4	4,19
7	Hsf	4,1
8	Myb	3,96
9	lola	3,96
10	fd68A, jummu, CHES-1-like	3,74
11	ubx	3,71
12	aop	3,51
13	emc	3,47
14	CrebB	3,43
15	Mad	3,32
16	pnt	3,27
17	CG17829	3,21
18	Pdi	3,19
19	Mef2	3,17
20	Adf1	3,17
21	rn	3,15
22	twi	3,09
23	CG8243	3,09
24	tj	3,02