



**A network of bHLH factors controls  
self-renewal and bipotential differentiation  
in the *Drosophila* intestine**

Aleix Puig Barbé

Thesis submitted for the award of Doctor of Philosophy

September 2018

## DECLARATION

This work has not been submitted in substance for any other degree or award at this or any other university or place of learning, nor is being submitted concurrently in candidature for any degree or other award.

Signed ..... (candidate)      Date .....

## STATEMENT 1

This thesis is being submitted in partial fulfillment of the requirements for the degree of PhD

Signed ..... (candidate)      Date .....

## STATEMENT 2

This thesis is the result of my own independent work/investigation, except where otherwise stated.

Other sources are acknowledged by explicit references. The views expressed are my own.

Signed ..... (candidate)      Date .....

## STATEMENT 3

I hereby give consent for my thesis, if accepted, to be available for photocopying and for inter-library loan, and for the title and summary to be made available to outside organisations.

Signed ..... (candidate)      Date .....

## STATEMENT 4: PREVIOUSLY APPROVED BAR ON ACCESS

I hereby give consent for my thesis, if accepted, to be available for photocopying and for inter-library loans **after expiry of a bar on access previously approved by the Academic Standards & Quality Committee.**

Signed ..... (candidate)      Date .....

## Acknowledgements

I would like to start expressing all my gratitude to my supervisor, Dr Joaquín de Navascués. First of all, for giving me the opportunity to join you in the lab. Second, because I had your support from the first moment, inside and outside the lab. I could count on you when I had any problem, and you cannot imagine how important this is when you live far from home for three years and a half. It has been a pleasure to work and learn by your side all these years and I am very proud to be the first PhD student (the first of many!) from your lab. I wish you all the luck in the future and I know that we will keep in touch.

I want to thank Cardiff University for supporting my PhD studentship.

I would like to express my gratitude to Dr Sonia López de Quinto and Dr Tony D. Southall for agreeing to be the examiners for my viva.

During the course of this thesis we initiated some experiments using the DamID method. This are ongoing at the time of writing, and are not included in the thesis, but I would like to thank Dr Tony Southall for his invaluable help with the DamID protocol and troubleshooting.

In the last four years I have been very lucky to be surrounded by wonderful people that supported me, in the lab and outside; in Cardiff or from the distance, and I would like to take the opportunity to thank them:

Sonia, Helen, Mike, Wynand: for all your suggestions and feedback during the fly meetings, that made this thesis much better. Especialmente quiero agradecerle, Sonia, que nos hayas hecho la vida mucho más fácil todos estos años, encargándote de que la comida de moscas nunca faltara, ayudándome a organizar los meetings y mil cosas más.

David, Simona and Angelo: I have really missed you the last year and a half. Friday's evenings at Woodville, discovering sushi places, watching Champions league matches (especially Barça-Roma!), ... I wish you all the best in Verona and Valencia, and I hope we can meet again in the future.

María Ángeles: la fiesta en Cardiff ya nunca fue lo mismo desde que te fuiste. Llegaste poco después que yo y me ayudaste muchísimo en el laboratorio. Pero sobretodo quiero agradecerte que me insistieras para que saliera de mi zona de confort en muchas ocasiones y recordarme que había más cosas aparte del lab.

Ramon Arens and Suzanne: for accepting me in your lab when I had no experience. Only now I can realize all the patience you had with me.

Marco, Ana y todos los milanos: por introducirme en el mundo de la mosca y enseñarme tanto. También por acogerme con los brazos abiertos cada vez que voy de visita.

Jo: I really cannot imagine how this lab would work without you. You are always helping everyone with a smile on your face.

Mark: for helping me every time the confocal decided not to work (or any machine in the lab!)

I would like to thank all the ECSCRI members for all the moments that we shared in the lab.

Adri: por todos esos partidos al Fifa, luchas con el Mortal Kombat, juegos de mesa, cenas en kaspas, chinos y japoneses, scape rooms y partidas al poker. Espero que ahora que volvemos los dos para España podamos seguir viéndonos.

Raquel: por todas las risas durante estos años, ya fueran en tu casa que la pared se caía a trozos o saliendo a cenar de vez en cuando. Te agradezco que pudieras hacerme huecos en tu tan apretada agenda, y aunque hubieras trabajado mil horas en la cafetería o tuvieras un deadline en breves.

Juan: per aguantar-me durant tots aquests anys. Encara no entenc que volguessis mudar-te a la mateixa casa (de fet sí, no podies dir que no a una casa que estava literalment a 25 m del lab). Sempre podrem recordar els teus sucs de taronja...

Leigh Anne: I really enjoyed getting to know you, especially in the last couple of years. If you come to Catalonia with Juan, it is obligated to come to visit me in Barcelona!

Ana: Literalmente no recuerdo cómo era el lab antes de que vinieras. ¿Con quién cotilleaba yo? ¿Quién aguantaba mis quejas por cualquier tontería o mis paranoias? Realmente nos hemos reído mucho todos estos años, y no te imaginas cuanto lo echaré de menos. ¡Ahora te toca a ti terminar la tesis para que nos podamos ir a Florida!

Irina: Thank you for always taking care of me, like an older sister. And also thank you for being my Portuguese mentor. You are the only one who speaks understandable Portuguese.

Pedro: ¡por ser el que siempre intenta engañar al resto para jugar a GOT! Ha sido un placer conocerte y compartir tantos momentos. También quiero destacar la gran labor que haces para intentar que Ana sea puntual, es un trabajo que no está lo suficientemente reconocido.

Joana: lab mate, house mate, image assessor, food supplier, ... I am really glad that it was you who moved in the house when Fabio left. I really miss living with you, so just find a job to Barcelona and we can be housemates again!

Carlotta: just thanks for being the way you are. You bring joy to the lab and whenever we meet outside. Also, I would like to let you know that when I grow up, I would like to be like you, working so hard in the lab, and somehow you find the time to do thousands of activities!

Fabio: Obrigado! Queres sumo? I really enjoyed living with you, playing PS4 until late or watching TV shows (although you know how I am with slow TV shows). And now inviting me from time to time to have brunch eating your famous pancakes.

Giusy: you are one of the people who made my integration in the lab easy. I feel really lucky to do my PhD in the same institute as you. Your mentoring was priceless, and I wish you all the best to you and your beautiful son in the future

Laura: mi hermana (mayor, cómo me recordarías tú). Sé que puedo contar contigo siempre que lo necesite. Que si tengo un día de bajón solo me hace falta llamarte para que mejore. Simplemente, quiero agradecerte que siempre, siempre estés ahí, incluso cuando hay miles de quilómetros de distancia.

Als meus pares: pel vostre suport incondicional, que m'estimeu tant tot i que sigui tant esquerp (inclús més que la Lana). No hauria arribat fins aquí sense vosaltres, ni seria la persona que sóc. Gràcies de tot cor.

Al meu germà i a tota la meva família: Per tot el vostre suport i estima.

## Table of contents

ACKNOWLEDGEMENTS.....	1
TABLE OF CONTENTS.....	5
ABSTRACT .....	8
LIST OF ABBREVIATIONS.....	9
<b>SECTION 1: INTRODUCTION.....</b>	<b>12</b>
A. ADULT STEM CELLS.....	14
B. THE <i>DROSOPHILA</i> MIDGUT.....	15
<i>B1. Parallelisms and differences with the mammalian intestine.....</i>	<i>16</i>
<i>B2. Signalling pathways regulating the differentiation of the ISC.....</i>	<i>18</i>
B2.1. The Notch pathway controls intestinal stem cell fate .....	18
B2.1.3. Notch signalling.....	18
B2.1.2. Notch signalling in the <i>Drosophila</i> midgut.....	20
B2.1.3. Notch signalling in the mammalian midgut .....	22
B2.2. The Jak/Stat pathway promotes both ISC proliferation and differentiation .....	23
<i>B3. Control of ISC proliferation .....</i>	<i>24</i>
<i>B4. Maintenance of the progenitor state .....</i>	<i>26</i>
<i>B5. Secretory differentiation.....</i>	<i>28</i>
B5.1. <i>Drosophila</i> secretory differentiation .....	28
B5.2. Secretory differentiation in the mammalian system.....	31
C. BHLH FACTORS .....	33
<i>C1. General remarks .....</i>	<i>33</i>
C1.1. Class I bHLH transcription factors.....	35
C1.2. Class II bHLH transcription factors.....	36
C1.3. Class V HLH factors .....	36
<i>C2. Notch and bHLH factors in the development of the peripheral nervous system .....</i>	<i>37</i>
<i>C3. bHLH proteins in the <i>Drosophila</i> developing eye .....</i>	<i>38</i>
<i>C4. bHLH proteins in the mammalian gut.....</i>	<i>40</i>
<b>SECTION 2: MATERIALS AND METHODS.....</b>	<b>42</b>
1. Fly stocks.....	43
Driver lines:.....	43
UAS-transgenes: .....	43
Expression reporter genes:.....	44
Mosaic Analysis with a Repressible Cell Marker (MARCM) stocks: .....	44
Mutants: .....	44

RNAi stocks .....	44
FlipOut Lineage tracing.....	45
<i>Drosophila melanogaster</i> husbandry.....	45
<b>2. Lineage Tracing experiments.....</b>	<b>45</b>
MARCM clones.....	45
FlipOut .....	46
<b>3. Immunohistofluorescence and imaging.....</b>	<b>47</b>
Paraformaldehyde fixation .....	47
Formaldehyde-Methanol fixation.....	47
Heat fixation .....	47
<b>4. Cell counting and quantification.....</b>	<b>49</b>
<b>5. Quantification in UAS-sc expression .....</b>	<b>49</b>
<b>6. Quantification of Delta expression .....</b>	<b>50</b>
Extraction of Pros <sup>+</sup> cells pixel values.....	50
Extraction of DI <sup>+</sup> cells pixel values in clones .....	50
Calculating cell intensity .....	51
<b>7. Esg-DamID binding regions .....</b>	<b>51</b>
<b>8. Statistical analysis .....</b>	<b>51</b>
<b>9. Generation of the <i>crp</i><sup>pl21.4</sup> mutant allele.....</b>	<b>52</b>
<b>SECTION 3: RESULTS .....</b>	<b>53</b>
CHAPTER 1 A: A NETWORK OF BHLH FACTORS CONTROLS SELF-RENEWAL AND BIPOTENTIAL DIFFERENTIATION IN THE INTESTINE: EMC.....	54
<b>1.A.1 Introduction.....</b>	<b>54</b>
<b>1.A.2. Aims.....</b>	<b>55</b>
<b>1.A.3. Results .....</b>	<b>55</b>
1.A.3.1. Expression of <i>emc</i> in the midgut .....	55
1.A.3.2. Loss of <i>emc</i> arrests terminal dedifferentiation .....	57
1.A.3.3. Loss of <i>emc</i> induces apoptosis .....	63
1.A.3.4. Emc is necessary to inhibit dedifferentiation of EBs .....	65
1.A.3.5. Expression of <i>emc</i> selects for absorptive fate .....	69
1.A.3.6. <i>emc</i> acts downstream of Notch activation .....	69
<b>1.A.4. Conclusions.....</b>	<b>74</b>
CHAPTER 1B: A NETWORK OF BHLH FACTORS CONTROLS SELF-RENEWAL AND BIPOTENTIAL DIFFERENTIATION IN THE INTESTINE: DA AND SC .....	75
<b>1.B.1. Introduction.....</b>	<b>75</b>
<b>1.B.2. Aims.....</b>	<b>76</b>

1.B.3. Results.....	76
1.B.3.1. Emc opposes Da function .....	76
1.B.3.2. Da homodimerizes to maintain the progenitor state .....	82
1.B.3.3. Notch controls the function of Da homodimers.....	85
1.B.3.4. Da homodimers do not auto-regulate and maintain stemness in parallel to Esg .....	88
1.B.3.5. Sc forms heterodimers with Da and functionally opposes Da:Da dimers to promote secretory differentiation.....	93
1.B.3.6. Emc downregulates Dl by antagonising Sc .....	96
1 B.4. Conclusions.....	98
1 A&B.5. Discussion .....	99
1.5.1. Imbalance of the bHLH equilibrium controls proliferation and differentiation.....	99
1.5.2. <i>emc</i> expression is controlled by different signals .....	101
1.5.3. Da:Da regulates the progenitor state .....	102
1.5.4. Sc triggers commitment to secretory fate, but delays terminal EE differentiation .....	103
1.5.5. Sc regulate Dl expression in ISCs and pre-EE cells .....	105
1.5.6. De-regulation of <i>da</i> induces cell death.....	106
1.5.7. Loss of <i>emc</i> in all the tissue as a potential tumor model.....	108
CHAPTER 2 CRP OPPOSES PROLIFERATION AND DIFFERENTIATION IN ISCs.....	110
2.1. Introduction .....	110
2.2. Aims .....	112
2.3. Results .....	112
2.3.1. Crp is expressed in all cell types of the adult posterior midgut.....	112
2.3.2. Generation of a new <i>crp</i> mutant allele .....	114
2.3.4. Overexpression of <i>crp</i> promotes EE cell death/elimination .....	118
2.4. Discussion .....	120
<b>CONCLUSIONS.....</b>	<b>123</b>
<b>APPENDIX 1: CELL COUNTER.....</b>	<b>124</b>
<b>APPENDIX 2: DL SIGNAL QUANTIFICATION .....</b>	<b>132</b>
<b>REFERENCES .....</b>	<b>137</b>

## Abstract

The mechanism that controls adult stem cell commitment is not fully elucidated, although it is known that the transcriptional control plays a major role in this process. Since the discovery of stem cells in the *Drosophila melanogaster* posterior midgut, many transcription factors have been identified to control whether intestinal stem cells (ISC) remain in the stem compartment, commit into the absorptive fate (enterocytes, EC) or differentiate into enteroendocrine (EE) cells. However, the molecular mechanisms governing these cell fate decisions remain to be fully elucidated.

We have identified a network of basic helix-loop-helix (bHLH) transcription factors to be key for ISC fate and the maintenance of the tissue homeostasis. The class I bHLH factor Daughterless (Da) impedes terminal differentiation by forming Da:Da homodimers, accumulating cells as ISCs or as absorptive committed cells (enteroblasts, EB). However, the class V HLH protein Extramacrochaete (Emc) inhibits the formation of Da:Da by forming dimers which lack transcriptional activity. We confirmed that emc is expressed in the posterior midgut and we showed that emc over-expression produces terminal differentiation of ISCs into ECs. Moreover, we found that Emc function is downstream of the Notch-Delta pathway. Importantly, we identified that Emc is also important in EBs to block de-differentiation of EBs into ISCs. This blocking is due to Emc inhibition over a bHLH class II, Scute (Sc). We showed that Sc has a dual role; it drives the expression of stem cells genes, such as Dl, while when being expressed at high levels, Sc initiates secretory differentiation.

Finally, we studied the function of a bHLH-leucine zipper called Cropped (Crp), whose expression produces the complete arrest of cell division and differentiation in ISCs.

Our results indicate that a simple Sc/Da/Emc network of bHLH factors act as a three-position toggle switch, choosing between the stem, secretory and absorptive fates by swapping dimerization partners. This is the first time that such a mechanism can account for all cell fate transitions in the fly gut, and it has direct implications for the maintenance of the mammalian intestine.

## List of abbreviations

<b>Ac</b>	Achaete
<b>AD1</b>	Activation Domain 1
<b>AiA</b>	Apoptosis induced Apoptosis
<b>AS-C</b>	Achaete-Scute complex
<b>Ascl2</b>	Achaete-Scute complex like 2
<b>Ase</b>	Asense
<b>ASF1</b>	Anti-Silencing Factor 1
<b>Ato</b>	Atonal
<b>Atoh1</b>	Atonal homolog 1
<b>BDSC</b>	Bloomington <i>Drosophila</i> Stock Center
<b>bHLH</b>	basic Helix-loop-Helix
<b>CBC</b>	Columnar Base Cell
<b>Chn</b>	Charlatan
<b>CPTI</b>	Cambridge Protein Trap Insertion
<b>CtBP</b>	C-terminal Binding Protein
<b>Da</b>	Daughterless
<b>DGRC</b>	<i>Drosophila</i> Genetic Resource Center
<b>DI</b>	Delta
<b>Dome</b>	Domeless
<b>E(spl)</b>	Enhancer of split
<b>EB</b>	Enteroblast
<b>EC</b>	Enterocyte
<b>EE</b>	Enteroendocrine
<b>EGF</b>	Epidermal Growth Factor
<b>Egr</b>	Eiger
<b>Emc</b>	Extramacrochaetae
<b>EMT</b>	Endothelial-Mesenchymal Transition
<b>Esg</b>	Escargot
<b>Ex</b>	Expanded
<b>Ey</b>	Eyeless
<b>Eyg</b>	Eyegone
<b>Fkh</b>	Fork head
<b>FO</b>	FlipOut

<b>GBE</b>	Grainyhead palindromic binding site
<b>GMC</b>	Ganglion Mother Cell
<b>H</b>	Hairless
<b>Hdc</b>	Headcase
<b>Hh</b>	Hedgehog
<b>HLH</b>	Helix-loop-Helix
<b>Hop</b>	Hopscotch
<b>Hpo</b>	Hippo
<b>Id</b>	Inhibitor of differentiation / Inhibitor of DNA binding
<b>ISC</b>	Intestinal Stem Cell
<b>l'sc</b>	Lethal of scute
<b>Lgr5</b>	Leucin-rich repeat-containing G-protein-coupled receptor 5
<b>LH</b>	Loop Helix
<b>Mam</b>	Mastermind
<b>mcs</b>	Microchaetae
<b>MCs</b>	Macrochaetae
<b>MF</b>	Morphogenetic Furrow
<b>NECD</b>	Notch Extracellular Domain
<b>NEXT</b>	Notch Extracellular Truncation
<b>Ngn3</b>	Neurogenin 3
<b>NICD</b>	Notch Intracellular Domain
<b>NIG</b>	National Institute of Genetics
<b>NRE</b>	Notch Responsive Element
<b>PNS</b>	Peripheral Nervous system
<b>PH3</b>	Phospho Histone 3
<b>Phyl</b>	Phyllopod
<b>PNC</b>	Pro-Neural Cluster
<b>pre-EE</b>	pre-Enteroendocrine
<b>REP</b>	Repression Domain
<b>Sc</b>	Scute
<b>Ser</b>	Serrate
<b>SOP</b>	Sensory Organ Precursor
<b>Spdo</b>	Sanpodo
<b>Su(H)</b>	Supressor of Hairless
<b>TA</b>	Transit Amplifying

<b>TACE</b>	TNF-alfa converting enzyme
<b>TAD</b>	Transactivator Domain
<b>Tap</b>	Target of Poxn
<b>TNF</b>	Tumor Necrosis Factor
<b>Ttk69</b>	Tramtrack69
<b>UAS</b>	Upstream Activated Sequence
<b>UBP</b>	Unknown Binding Partner
<b>Upd</b>	Unpaired
<b>VDRC</b>	Vienna <i>Drosophila</i> Resource Center
<b>VM</b>	Visceral Muscle
<b>Vn</b>	Vein
<b>Wg</b>	Wingless
<b>Yki</b>	Yorkie

During the elaboration of this thesis, I decided to write the Results section using the first person, plural personal pronoun “we”. The aim is to avoid constant changes between the singular and the plural pronouns. All the experiments were planned and discussed by Aleix Puig and Dr. Joaquín de Navascués. Unless otherwise stated, all the experiments, data acquisition and data analysis were performed by Aleix Puig. The development of the scripts to analyse the data were done by Aleix Puig under the supervision of Joaquín de Navascués.

# Section 1: Introduction

“The fixity of the milieu supposes a perfection of the organism such that the external variations are at each instant compensated for and equilibrated” (Bernard, 1878). With this sentence, Claude Bernard described a dynamic process that later Cannon termed as homeostasis: “The coordinated physiological reactions which maintain most of the steady states in the body are so complex, and are so peculiar to the living organism, that it has been suggested that a specific designation for these states be employed — homeostasis” (Cannon, 1929).

Adult tissues maintain homeostasis through a delicate equilibrium between proliferation and differentiation of adult stem cells. In this thesis we will try to comprehend the transcriptional regulation of adult stem cell differentiation in the *Drosophila* midgut.

## A. Adult stem cells

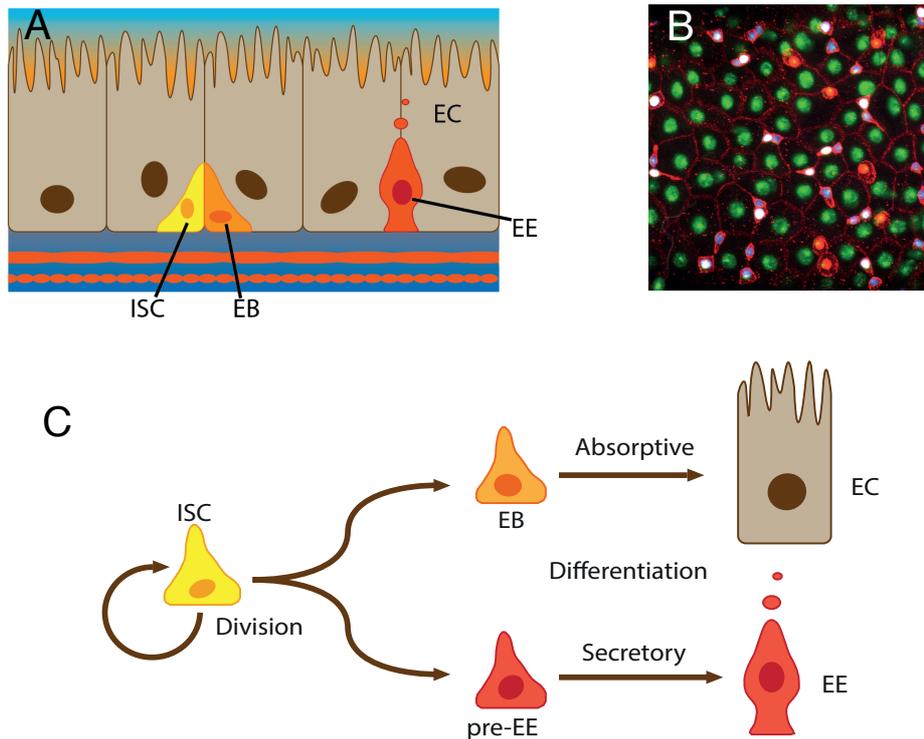
During the first stages of development, embryonic stem cells are the source that will generate all cell types of the different tissues and organs. However, once the individual is formed and has grown, how will all the different organs be maintained? Adult stem cells are a subset of cells that resides in adult tissues, having the capacity to divide in order to generate new stem cells (self-renewal) and differentiate directly or through multiple steps. This capacity to self-replicate and commit is needed to replenish the tissue when differentiated cells are lost. Moreover, stem cells need to maintain the multipotency in order to replace any cell type that has been lost in the tissue. However, stem cell processes are tightly regulated, as an indiscriminate proliferation or lack of it would lead to disease situations such as cancer. Therefore, there is a balance between divisions and cell loss that will maintain the homeostasis of the tissue. Thus, in a tissue injury situation, the number of cell division will increase, followed by differentiation to replenish all cells that are lost.

Some tissues can be maintained with unipotent stem cells, as they only need to replace one cell type. In these tissues, stem cells only control proliferation and when to differentiate, but do not need mechanisms that decides the different cell fates. This is the case of the skeletal muscle (Costamagna *et al.*, 2015; Wosczyzna and Rando, 2018). However, in other organs stem cells can differentiate into a number of different fates. These are the cases of the small intestine, the trachea or the airway epithelia (reviewed in Gehart and Clevers, 2015; Rawlins and Hogan, 2006). Some organs are exposed to insults; for instance, in the midgut the resident microbiota, digestive enzymes and biliary and gastric acids produce the necessary environment to process the nutrients, which can be damaging for the cells exposed to it. Hence, there is an active replacement of the different cell types and after cell proliferation, daughter cells are programmed for correct differentiation. Therefore, the intestine has been raised as a paradigmatic organ where to study the transcriptional regulation of stem cell differentiation, as it has a great biomedical value. The *Drosophila melanogaster* midgut captures the essence of the problem of the multipotent stem cell, as it contains bipotent stem cells whose fate depend on a precise transcriptional regulation. Moreover, the *Drosophila* midgut is the only organ in the fly that contains multipotent cells.

## B. The *Drosophila* midgut

The *D. melanogaster* midgut is a flat tube that contains adult stem cells that sustain the differentiated cells (Micchelli and Perrimon, 2006; Ohlstein and Spradling, 2006). The homeostatic intestine has four cell types: the intestinal stem cells (ISC), cells committed to differentiation, called enteroblasts (EB), the absorptive enterocytes (EC) and secretory enteroendocrine (EE) cells (Micchelli and Perrimon, 2006; Ohlstein and Spradling, 2006) (**fig. I1A, B**). Once an ISC differentiates into an EB, this cell commits to differentiation without further divisions. Some studies propose that EBs differentiate to ECs or EEs (Ohlstein and Spradling, 2006; Ohlstein and Spradling, 2007). However, two recent studies show that there could be two different types of progenitors, one that will differentiate into an EC (EBs) and a different one, named pre-EE cells, which are a very short-lived cell that quickly differentiates to an EE cell (**fig. IC**) (Biteau and Jasper, 2014; Zeng and Hou, 2015). EE cells secrete peptide hormones and there are different subgroups depending on the hormones they secrete, such as allostatin or tachykinin (Beehler-Evans and Micchelli, 2015). ECs contain an actin-rich brush at the apical side to take nutrients. The size of ECs is bigger with respect to the other cell types in the gut due multiple cycles of endoreplication.

Similarly to the mammalian intestine, equivalent ISCs choose stochastically whether to symmetrically self-renew, symmetrically differentiate, or allocate fate asymmetrically, in balanced proportions (de Navascués *et al.*, 2012). A quantitative study of ISC divisions showed that while an 80% of divisions are asymmetric, a 20% are symmetric, 10% giving two ISCs and 10% giving two EBs. de Navascués *et al.* proposed that after ISC division, daughter cells could be uncommitted and their fates would be defined stochastically (de Navascués *et al.*, 2012).



**Figure I1. The *Drosophila* midgut**

In the *Drosophila* posterior midgut one can typically find four different cell types: Intestinal stem cells (ISC), which are the main proliferative population; committed enteroblasts (EB), which do not divide and differentiate into enterocytes (EC), the mature absorptive cells; moreover, the ISCs are bipotent, and can produce secretory enteroendocrine (EE) cells.

**A.** Side view of the midgut epithelium. Cells are depicted with their apical side at the top, facing the lumen of the intestine.

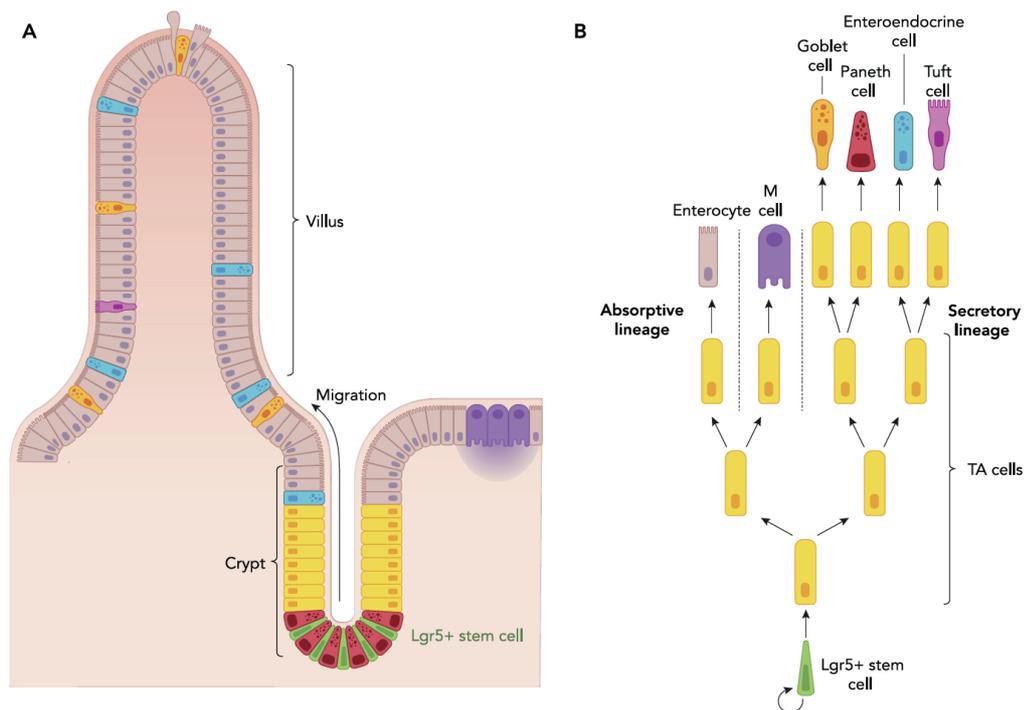
**B.** Micrography (top view) of the *Drosophila* midgut epithelium, stained with Armadillo (red), labelling the cell membrane, Prospero (nuclear red) for EE cells, *escargot-lacZ* (blue) for progenitor cells, *GBE-Su(H):GFP* (white) for EBs and DNA (green) for all nuclei. Scale Bar: 20µm. Adapted from de Navascués *et al.*, 2012.

**C.** ISCs self-renew and differentiate into ECs, going through the long-lived EB transient state; or into EE cells, going through a very short-lived pre-EE state.

## B1. Parallelisms and differences with the mammalian intestine

The mammalian intestinal tract is a folded monolayer that contains multiple units. Each unit is sub-divided in two distinct regions: the villi are evaginations that are more exposed to the lumen and present differentiated cell types for nutrient absorption and digestion; and the crypts of Lieberkühn are invaginations that contains stem cells, progenitor cells and Paneth cells (reviewed in Clevers, 2013). At the base of the crypt we can identify the columnar base cells (CBC), which are stem cells that can be identified with the marker Leucin-rich repeat-containing G-protein-coupled receptor 5 (Lgr5) (**fig. I2A**) (Barker *et al.*, 2007; Snippert *et al.*, 2010). Interestingly, crypts that had complete ablation

of CBCs maintained homeostasis, and the Lgr5<sup>+</sup> cell pool was restored in the crypt (Tian *et al.*, 2011). Clonal analysis performed by Tian *et al.* showed that new CBC's source were cells localized at the +4 position from the base of the crypt. These +4 cells were for long thought to be quiescent stem cells (Muñoz *et al.*, 2012; Potten *et al.*, 1978). However, it was later shown that the +4 cells are secretory precursors that in stress conditions where CBCs are compromised, can de-differentiate into stem cells (Buczacki *et al.*, 2013; Schwitalla *et al.*, 2013). Supporting stem cells, Paneth cells were surrounding CBCs to form the niche with stromal cells, providing Wnt, Notch and EGF signalling to maintain self-renewal and multipotency (Farin *et al.*, 2012; reviewed in Gassler, 2017; Valenta *et al.*, 2016). Moreover, CBCs give rise to the transit amplifying (TA) cells, which rapidly proliferate and migrate to the lumen. These cells are progenitors that will differentiate either into the absorptive fate or the secretory fate (see section B.2.1.3). When progenitor cells reach the villus, they terminally differentiate. Absorptive progenitors only



**Figure 12. The architecture of the mammalian intestinal epithelium**

**A.** CBCs (Lgr5<sup>+</sup> cells) (green) are found at the base of the crypts supported by Paneth cells (red). CBCs can divide and differentiate. Committed progenitor cells (yellow) are specified in the transit amplifying zone, where they are specified into the secretory or absorptive lineage. When cells are completely differentiated, they migrate to the villus. Taken from Anderson-Rolf *et al.*, 2017.

**B.** CBCs self-renew and differentiate into the absorptive fate (ECs), M cells or the secretory fate, which comprises several subtypes (Goblet cells, Paneth cells, EE cells and Tuft cells). Taken from Anderson-Rolf *et al.*, 2017.

differentiate into ECs, while secretory progenitors can differentiate into EE cells, tuft cells, goblet cells or Paneth cells (**fig. I2B**), with the later migrating to the crypt (Marshman *et al.*, 2002).

Therefore, both CBCs in mammals and ISCs in *Drosophila* are multipotent and self-renew. However, while it has been shown that ablation of stem cells in mammals could be reverted with some progenitor cells, to date in *Drosophila* no de-differentiation has been observed when all progenitor cells disappear (Lu and Li, 2015). In both systems differentiated cells go through a progenitor state, although in *Drosophila* absorptive progenitor cells do not divide. It has been reported that EBs can migrate before differentiating into ECs (Antonello *et al.*, 2015), which resembles the migration in the TA compartment.

## **B2. Signalling pathways regulating the differentiation of the ISC**

In the *Drosophila* midgut two pathways have been identified to control whether ISCs can differentiate or not, and to which cell fate: the Notch pathway and the Jak/Stat pathway.

### B2.1. The Notch pathway controls intestinal stem cell fate

Notch pathway is an evolutionary conserved pathway that is involved in the correct cell specification during development and adult tissues. In the *Drosophila* midgut, Notch is the most important pathway for the correct differentiation of ISCs (Micchelli and Perrimon, 2006; Ohlstein and Spradling, 2006; Ohlstein and Spradling, 2007).

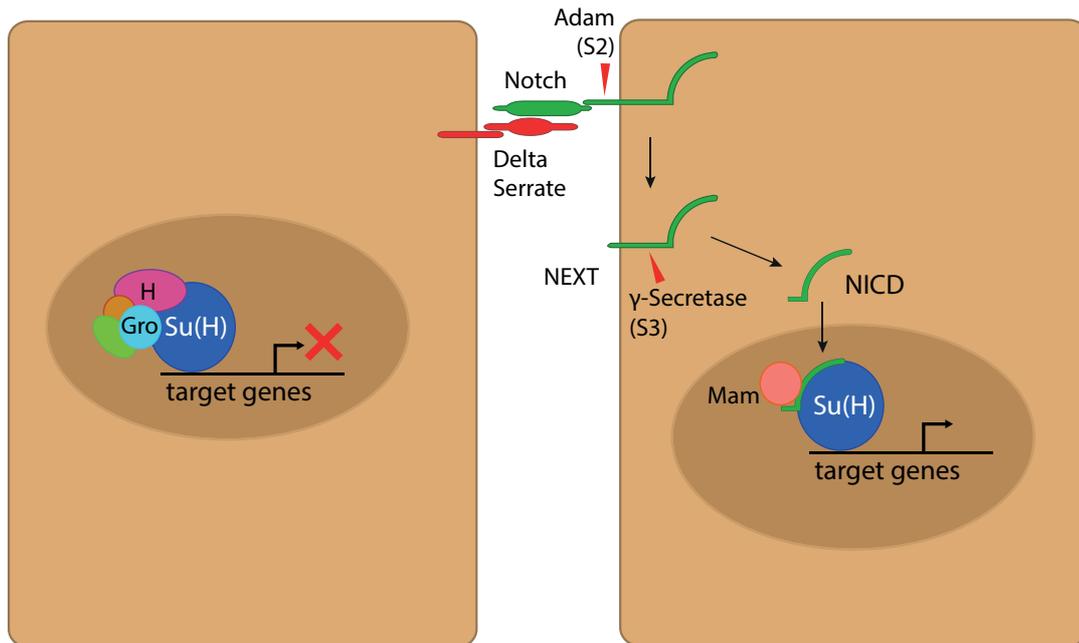
#### *B2.1.3. Notch signalling*

Notch is a transmembrane receptor that owes the name to the notches that fly wings had in heterozygous Notch mutant females (Mohr, 1919; Wharton *et al.*, 1985). Notch, with its two ligands Dl (Dl) and Serrate (Ser), were first identified as neurogenic genes, as when they were lost epidermal cells transformed into neuroblasts (Fleming *et al.*, 1990; Lehmann *et al.*, 1983; Vässin *et al.*, 1987). Both ligands were shown to be type I single-pass transmembrane, and therefore, activation of Notch needed physical contact with a cell expressing the ligands (Fleming *et al.*, 1990; Heitzler and Simpson, 1991; Kopczynski *et al.*, 1988) (**Fig. I3**).

Notch contains a long extracellular domain (NECD) and a shorter intracellular domain (NICD). In the NECD there are 36 tandem epidermal growth factor (EGF) repeats, which are also found in Dl (7) and Ser (12). The N-terminal domain of the ligands directly bind with two specific EGF repeats (11-12) in Notch to initiate the activation (Rebay *et al.*, 1991). Notch receptor contains three cleavage sites (S1, S2 and S3). Cleavage on the S1 site is mediated by Furin and it is necessary for the maturation of the protein during the secretory pathway, independently of ligand binding. After ligand-receptor binding, the TNF- $\alpha$  converting enzyme (TACE) and kuzbanian, two metalloproteases of the ADAM family, cleave the S2 site, releasing the NECD and forming an activated Notch extracellular truncation (NEXT) (Brou *et al.*, 2000; Rooke *et al.*, 1996). Then, the  $\gamma$ -secretase protease complex cleaves the S3 site through its catalytic subunit presenilin (Struhl and Adachi, 2000; Struhl and Greenwald, 1999; Ye *et al.*, 1999), releasing the NICD from the membrane. NICD, which is the domain with transcriptional function, translocates to the nucleus and forms a complex with the DNA binding protein Suppressor of Hairless, Su(H) (Fortini and Artavanis-Tsakonas, 1994). Consequently, this complex can recruit the co-activator Mastermind (Mam) (Petcherski and Kimble, 2000) (**Fig. I3**).

When NICD is not present, Su(H) remains bound to the DNA and to the adaptor protein Hairless (H) (Brou *et al.*, 1994; Furriols and Bray, 2000). H promotes chromatin repression by recruiting the co-repressors Groucho, C-terminal binding protein (CtBP) and anti-silencing factor 1 (ASF1), which is a conserved H3-H4 histone chaperone (Bang and Posakony, 1992; Barolo *et al.*, 2002; Maier *et al.*, 1999).

Notch signaling typically implies two or more cells. Notch signaling is required for acquisition of distinctive fates in neighboring cells that are in contact. Thus, one cell activates the Notch receptor of all surrounding cells, and Notch activation triggers a downstream cascade that prevents their ability to present Notch ligands, and the cell that cannot be inhibited by Notch anymore achieve an alternative fate than the neighbors. This is called lateral inhibition (reviewed in Sjöqvist and Andersson, 2017). After ISC division in the *Drosophila* midgut, lateral Notch controls the daughter fate by lateral inhibition (Guisoni *et al.*, 2017).

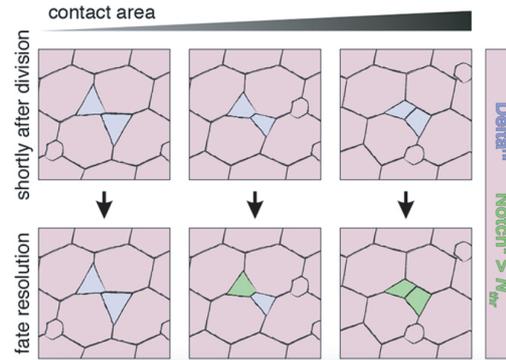


**Figure I3. Notch pathway**

Notch activation starts when the receptor binds with the ligand. The metalloproteases from the ADAM family can cleave the S2 site, generating the NEXT. A second cleavage by the  $\gamma$ -secretase releases the NICD to the cytoplasm, which can enter the nucleus. NICD then can bind with Su(H) and recruit Mam to drive transcription of target genes. In the absence of Notch activation, Su(H) binds H and recruits the co-repressors Gro, CTFP (orange) and ASF1 (green).

#### *B2.1.2. Notch signalling in the Drosophila midgut*

Notch signalling plays a crucial role in the *Drosophila* midgut to control cell lineage and therefore, it has emerged as an excellent model to understand Notch signalling in adult stem cells. ISCs express both Notch and the ligand Delta (Bardin *et al.*, 2010). After ISC divides both cells inherit Dl and Notch in similar amounts (Ohlstein and Spradling, 2007), although in the majority of cases, Notch is activated in one of the two cells and it becomes an EB. Modulating Notch signalling, Perdigoto *et al.* showed that a high level of Notch activity was required for cells to become EBs (Perdigoto *et al.*, 2011). EBs do not present Dl anymore, and therefore their sister cells remain as ISCs (asymmetric divisions) (de Navascués *et al.*, 2012; Micchelli and Perrimon, 2006; Ohlstein and Spradling, 2007). However, there are cases where two daughter cells can achieve the same fate, either ISCs or EBs, and it was proposed that this might depend on the amount of Notch signalling between the two daughter cells, as limited by their contact area (**Fig. I4**) (Guisoni *et al.*, 2017).



**Figure I4 Contact area determines the daughter fate after ISC division**

After ISC division, small contact area between two daughter cells maintain both cells as ISCs, while a large contact area favours Notch activation and differentiation into EBs of both cells. (Guisoni *et al.*, 2017)

Apart from the contact area, asymmetric protein segregation during ISC divisions could also be important for daughter fates. Integrins regulate spindle orientation and the re-localization of the Par complex in the putative ISC (Goulas *et al.*, 2012). Suppression of the Par complex induces ISCs and EE cells large clusters, in a similar manner to loss of function mutations of *Notch* (Goulas *et al.*, 2012). Moreover, the endosomal adaptor protein Numb is also localized to one side of the dividing cell (Sallé *et al.*, 2017). Numb has been identified as a Notch inhibitor in neurogenesis by regulating Notch trafficking to late endosomes and restricting a specific sub-population of Notch that would be recycled (Guo *et al.*, 1996; Johnson *et al.*, 2016). In the midgut, Numb is also found in the new ISC side during mitosis and has been observed to be upstream of Notch signalling (Sallé *et al.*, 2017). In addition, Sara endosomes, which contain Notch and Dl, are inherited by presumptive EBs after mitosis (Montagne and Gonzalez-Gaitan, 2014). Notch could be activated in these endosomes, initiating the acquisition of the fate.

The activation of Notch is key to progress into the absorptive fate and lack of Notch induces ISCs over-proliferation or selection into the secretory fate (Micchelli and Perrimon, 2006; Ohlstein and Spradling, 2006). Moreover, inhibiting Notch signalling with H expression or removing Su(H) also induce an accumulation of ISCs (Bardin *et al.*, 2010). On the other hand, ectopic induction of NICD in ISCs results in differentiation into EC differentiation, and a complete loss of the stem pool (Micchelli and Perrimon, 2006).

It is likely that Notch activation in the gut induces the expression of members of the *Enhancer of split* [*E(spl)*] complex (Celis *et al.*, 1996; Jennings *et al.*, 1994), which are transcriptional regulators of the basic helix-loop-helix (bHLH) family, class VI. In the *Drosophila* midgut, loss of the whole *E(spl)* complex results in accumulation of ISCs (Bardin *et al.*, 2010), without secretory differentiation.

Notch signalling also has a second role in the midgut for secretory specification, as *Notch* loss of function clones increase the number of EE cells (Ohlstein and Spradling, 2006). Guo and Ohlstein reported that in the pupae and later in the adult midgut, after asymmetrical division, the newly formed pre-EE cell express *Dl*, while ISC has Notch activated transiently to halt secretory differentiation (Guo and Ohlstein, 2015).

### *B2.1.3. Notch signalling in the mammalian midgut*

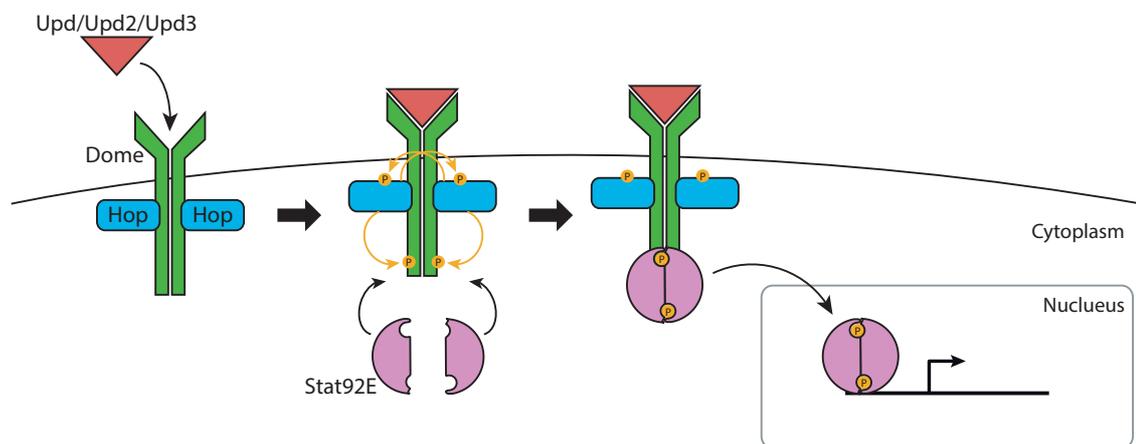
Notch signalling is also important in the mammalian midgut, although it shows some differences with *Drosophila*. Notch activation has a dual role: to promote proliferation and prevent differentiation in CBC and to select for the absorptive phenotype instead of the secretory in TA cells. In *Notch* loss of function mutations or *Dl1* and *Dl4* mutations, all stem cells halted the proliferation and differentiated into EE cells (Pellegrinet *et al.*, 2011; van Es *et al.*, 2005), while activation of Notch signalling in the gut promotes stem cell proliferation and ceases goblet and EE differentiation (Fre *et al.*, 2005). Notch effectors are the *E(spl)* mammalian homologs, the Her/Hes family, which target *atonal homolog 1* (*Atoh1*), a member of the bHLH family, responsible for secretory differentiation (Kazanjian *et al.*, 2010; Kim and Shivdasani, 2011; Milano *et al.*, 2004; Shroyer *et al.*, 2007; VanDussen and Samuelson, 2010; Yang *et al.*, 2001). Due to genetic redundancy, the ablation of only one member of the family, such as *Hes1*, is not enough to stop the proliferation and promote secretory differentiation, and it is needed to inactivate at least *Hes1*, *Hes3* and *Hes5* (Ueo *et al.*, 2012). Interestingly, the inactivation of *Hes1* alone is sufficient to eliminate proliferation in intestinal tumours without affecting the homeostasis of the crypt (Ueo *et al.*, 2012).

In the crypt base, Paneth cell presents *Dl* on the surface to induce activation of Notch in the CBC. In the TA compartment, secretory progenitors signal to the absorptive progenitors to activate Notch and promote differentiation into ECs. Therefore, compared

to the *Drosophila* midgut, Notch activation has opposed effects in ISCs, as in mammals maintains stemness, while in *Drosophila* induces differentiation. However, in both systems Notch signalling selects for absorptive differentiation over secretory fate (reviewed in Fre *et al.*, 2011; Perdigoto and Bardin, 2013).

## B2.2. The Jak/Stat pathway promotes both ISC proliferation and differentiation

The Jak/Stat pathway is a conserved pathway active in numerous developmental processes. In *Drosophila*, there are three leptin-like cytokines (Unpaired (Upd), Upd2 and Upd3) which are the secretable ligands, one transmembrane receptor called Domeless (Dome), one tyrosine kinase called Hopscotch (Hop) and a transcription factor, Stat92E. Binding of Upd, Upd2 or Upd3 to Dome activates the receptor associated kinase Hop, which phosphorylates Dome and itself. This phosphorylation generates docking sites in Dome/Hop dimers, allowing cytoplasmic Stat92E to bind Dome/Hop through the Stat92E SH2 domains. Then, Stat92E is also phosphorylated by itself and forms homodimers with other phosphorylated Stat92E to translocate into the nucleus and drive transcription of target genes, including its own expression to create a positive feedback loop (**fig. I5**) (reviewed in Arbouzova & Zeidler 2006).



**Figure I5 Jak/Stat pathway**

Upd/Upd2/Upd3 bind and activate the receptor Dome, which have Hop associated. Dome activation induces Hop trans-phosphorylation and therefore, Hop activation. Then, Hop phosphorylates Dome, inducing Stat92E recruitment and phosphorylation. Phosphorylated Stat92E dimerise and translocate to the nucleus to bind DNA target sequences to regulate gene transcription.

In the fly midgut, Upd, Dome, Hop and Stat92E are expressed in all progenitor cells (Beebe *et al.*, 2010; Jiang *et al.*, 2009; Lin *et al.*, 2010; Liu *et al.*, 2010). However, it seems that the localization of Stat92E in the majority of these cells is cytoplasmic, and only newly formed EBs shows a strong nuclear intensity (Liu *et al.*, 2010). Loss of function of *hop* or *Stat92E* in progenitor cells caused an inability of these cells to terminally differentiate either to ECs or to EE cells (Beebe *et al.*, 2010; Jiang *et al.*, 2009; Lin *et al.*, 2010; Liu *et al.*, 2010). Moreover, while some studies show that ISCs can still progress to EBs (Jiang *et al.*, 2009; Lin *et al.*, 2010), others show that this commitment step is also arrested (Liu *et al.*, 2010). However, the strong nuclear signal in newly formed EBs supports the hypothesis that Stat92E is also necessary for EB differentiation.

There are multiple sources for Upd signalling to activate the Jak/Stat pathway in ISCs. The activation of JNK in old ECs will lead to apoptosis and induces the expression and secretion of Upd, Upd2 at low levels, and of Upd3 at very high levels in ECs (Jiang *et al.*, 2009; Zhou *et al.*, 2013). Upd has also been shown to be secreted by progenitor cells themselves and from the visceral muscle (VM) (Lin *et al.*, 2010).

The expression of *Stat92E* seems to be under the control of Notch, as loss of *Notch* increase *Stat92E* expression. Interestingly, loss of *Notch* also enhances *upd* expression (Liu *et al.*, 2010). In addition, ectopic expression of NICD in progenitor cells to activate the Notch pathway cannot induce terminal differentiation when the Jak/Stat pathway is not functional (Beebe *et al.*, 2010; Jiang *et al.*, 2009; Lin *et al.*, 2010). It could be possible either that Notch negatively regulates directly the expression of members of the Jak/Stat pathway or activates the expression of Jak/Stat inhibitors, such as *windpipe* (Ren *et al.*, 2015).

### **B3. Control of ISC proliferation**

To maintain the integrity of an adult tissue, when a cell is lost a new cell needs to be generated. The *Drosophila* midgut contains multiple signalling pathways that can sense that a cell is lost, either if is at small scale (an old EC is lost due the wear and tear) or at a big scale (infection, damaging agents) (Amcheslavsky *et al.* 2009). Therefore, the newly formed cell can proceed to differentiation and replace lost cells. Therefore, multiple pathways are responsible to control ISCs divisions.

One of the most important mitogenic pathways in the *Drosophila* midgut is the **Jak/Stat** pathway (Beebe *et al.*, 2010; Buchon *et al.*, 2009; Cordero *et al.*, 2012a; Jiang *et al.*, 2009; Lin *et al.*, 2010; Liu *et al.*, 2010; Suijkerbuijk *et al.*, 2016). However, although inhibition of the Jak/Stat pathway reduces drastically the proliferation when cell death is induced in ECs, it is not absolutely required for ISC self-renewal, as loss of function of *Stat92E* or *hop* clones still contain mitotic ISCs (Beebe *et al.*, 2010).

Another important proliferation pathway is the **EGFR/Ras/MAPK** pathway by which, similarly to the Jak/Stat, ISCs receive mitogenic signals from other cells (Buchon *et al.*, 2009; Jiang *et al.*, 2011). EBs secrete the EGFR ligands Spitz and Keren (Buchon *et al.*, 2010; Xu *et al.*, 2011), ECs secretes Keren at low levels (Jiang *et al.*, 2011) and the VM secretes Vein (Vn) (Biteau and Jasper, 2011; Jiang *et al.*, 2011; Xu *et al.*, 2011) to induce proliferation of ISCs. In conditions of inhibition of *vn* expression, the ISC population still persists or is not affected (Biteau and Jasper, 2011; Jiang *et al.*, 2011; Xu *et al.*, 2011). Some studies show that Upd3 secreted from ECs drives expression of *spitz* in EBs and *vn* in the VM, showing a synergy between both pathways (Zhou *et al.*, 2013). Finally, the activation of the MAPK pathway seems to induce phosphorylation in one of the two necessary phosphorylation sites of Fos (JNK pathway) (Biteau and Jasper, 2011). However, the role of the JNK pathway in ISCs has not been elucidated yet, as so far it has been shown to have a role only in aging, mediating the loss of homeostasis of the tissue in old flies (Biteau *et al.*, 2008). Therefore, it is possible that MAPK has different targets to promote proliferation.

The **Wingless** (Wg) pathway can also affect ISC proliferation. The activation of the pathway resulted in an increased number of ISC mitosis (Lee *et al.*, 2009; Lin *et al.*, 2008). In addition, ectopic expression of *wg* in progenitor cells also induced more cell divisions. However, these two studies differ in whether Wg had an essential role in ISC self-renewal or not, as Lin *et al.* reported that attenuated *wg* signal induced loss of ISC, while Lee *et al.* did not observe any difference in the number of clones when Wg was hyperactivated.

The **Hippo** (Hpo) pathway is implicated in the coordination of organ growth during development (reviewed in Hariharan, 2006). The activation of this pathway results in the phosphorylation and inactivation of the transcription factor Yorkie (Yki) (Huang *et al.*,

2005). Yki promotes proliferation by inducing the expression of cell cycle genes and inhibits apoptosis inducing anti-apoptotic genes (Huang *et al.*, 2005). However, in the homeostatic *Drosophila* gut, inhibition of *yki* does not affect ISC proliferation. When upstream members of the Hpo pathways are not present or Yki is activated, the number of ISC mitoses increase greatly (Karpowicz *et al.*, 2010; Li *et al.*, 2014a; Ren *et al.*, 2010; Shaw *et al.*, 2010; Staley and Irvine, 2010). This indicates that the Hpo pathway is inhibiting Yki in homeostatic conditions. Interestingly, loss of function clones of one of the Yki targets, a microRNA called *bantam*, reduced proliferation (Huang *et al.*, 2014). Although Huang *et al.* claim that *bantam* regulates proliferation downstream of the Hpo pathway, it is likely that another transcription factor can regulate *bantam* expression.

**Hedgehog** (Hh) signalling is also implicated in ISC division regulation, inducing ISC proliferation. During development, Hh is implicated in tissue grow and patterning (Cohen, 2003; Zecca *et al.*, 1995). *hh* is expressed in ISCs, ECs and VM and knock down of *hh* in all these cells (using and ubiquitous promoter) reduced the number of ISC divisions (Li *et al.*, 2014b).

#### **B4. Maintenance of the progenitor state**

Apart from the differentiation and proliferation pathways, several transcriptional regulators have been identified, that are expressed in then intestinal progenitor cells and play a role in the maintenance of the undifferentiated state.

Two members of the Snail family of transcription factors are important in *Drosophila* adult ISCs. Snail proteins are zinc-finger transcription factors normally implicated in epithelial-mesenchymal transition (EMT) processes (Nieto, 2002), but are also involved in stem cell processes. In the *Drosophila* midgut, **Escargot** (*esg*) is expressed in ISCs, EBs and pre-EE cells, but not in EE cells and ECs (Antonello *et al.*, 2015; Micchelli and Perrimon, 2006; Zeng and Hou, 2015). Two recent studies have stated the importance of *Esg* maintaining ISC stemness and blocking differentiation (Korzelius *et al.*, 2014; Lozacoll *et al.*, 2014). They showed that overexpressing *esg* in progenitor cells block terminal differentiation and leads to the formation of clusters of EBs with some ISCs. Loss of *esg* drives ISCs and EBs cells to differentiate into ECs and EE cells (Korzelius *et al.*, 2014). In progenitor cells, *Esg* inhibits the expression of *nubbin*, which is necessary for EC

differentiation (Korzelius *et al.*, 2014). Antonello *et al.* speculated that EBs could sense cell loss or tissue stress and express the micro-RNA *mir-8*, which would inhibit the expression of *esg* and allow terminal differentiation (Antonello *et al.*, 2015). On the other hand, *Esg* inhibits secretory differentiation by directly repressing the expression of *prospero* (*pros*) (Korzelius *et al.*, 2014; Li *et al.*, 2017a). It has also been suggested that *Esg* induces EMT in mature EBs to integrate into the epithelial layer and differentiate (Antonello *et al.*, 2015). **Snail**, another member of the Snail family, is also expressed in progenitor cells and it has been shown to inhibit both differentiation and proliferation, maintaining progenitor cells in a quiescent state (Dutta *et al.*, 2015).

**Charlatan** (*Chn*) is a zinc-finger transcription factor of the C2H2 family important for maintaining the progenitor state (Amcheslavsky *et al.*, 2014a). Over-expression of *chn* in progenitor cells blocks the differentiation into ECs or EE cells and promotes proliferation. However, although loss of *chn* reduces ISC divisions, it does not promote direct differentiation. Cells remain as *Esg*<sup>+</sup>, but neither express the ISC marker *Dl*, nor activates Notch (a trademark of EBs). This suggests that ISCs lose some stem properties upon *chn* loss, but do not differentiate (Amcheslavsky *et al.*, 2014a).

The **Osa**-containing SWI/SNF remodelling complex has also seen to play a role in maintaining the stem fate. Zeng *et al.* showed that knocking down the expression of *osa* in progenitor cells induce accumulation of *Dl*<sup>-</sup> ISCs (identified by the ISC marker *Sanpodo* [*Spdo*]). They identified that *Osa* binds to the promoter regions of *Dl* and *asense* (*ase*) regulating their transcription (Zeng *et al.*, 2013). Therefore, when *osa* is not present, *Dl* and *ase* cannot be expressed. Lack of *Dl* interrupts the Notch pathway to promote the absorptive differentiation, while loss of *ase* also blocks the secretory differentiation (see section B5.1) (Bardin *et al.*, 2010). Therefore, ISCs proliferate and accumulate without differentiating.

A member of the Sox family, **Sox21A**, is expressed in progenitor cells, although there is a disagreement about which cell express higher levels of *sox21A*. While some groups indicate that EBs, specially the more mature ones, express higher levels of *sox21A* than ISCs (Chen *et al.*, 2016), others shows that ISC express higher levels than EBs (Meng and Biteau, 2015; Zhai *et al.*, 2017), and the levels of Sox21A decrease when EBs mature (Zhai

*et al.*, 2017). The inconsistency of the results could be explained by the different detection methods, as Chen *et al.* use an antibody that detects cytoplasmic Sox21A, while Zhai *et al.* use Sox21A reporter. Meng and Biteau developed an independent Sox21A antibody that detects specifically nuclear Sox21A and although they did not quantify the signal in each cell, it seems that ISC has a stronger nuclear signal (Meng and Biteau, 2015). It also would be possible that progenitor cells have different requirements depending on the state of the midguts, as it has been shown that in homeostatic guts, Sox21A limits cell division, whereas in a stress context, Sox21A is required in the ISC to activate proliferation (Chen *et al.*, 2016; Zhai *et al.*, 2017). Moreover, Sox21A also has a differentiation function in EBs, as loss of function of *sox21A* promotes accumulation of EBs (Chen *et al.*, 2016; Meng and Biteau, 2015; Zhai *et al.*, 2015; Zhai *et al.*, 2017). The activation of Sox21A depends on the Jak/Stat pathway and activates expression of *Dl* in ISCs and *GATAe* and *dpp* in differentiating EBs during active regeneration (Zhai *et al.*, 2017).

FoxA transcription factor **Fork head** (Fkh) is a transcription factor highly expressed in ISCs and EBs and has lower expression in EE cells, while ECs do not express it (Lan *et al.*, 2018). Loss of *fkh* in progenitor cells induce terminal absorptive differentiation. The function of *fkh* seems to be, rather than acting as a transcription factor, to keep the chromatin open and allow the binding of other transcription factors to maintain the progenitor pool (Zaret *et al.*, 2010).

Finally, **Daughterless** (Da) is a member of the bHLH class I that has also been shown to be crucial to maintain the progenitor state, as loss of *da* induce terminal differentiation (Bardin *et al.*, 2010). In this thesis we will explore the function of Da extensively.

## **B5. Secretory differentiation**

### B5.1. *Drosophila* secretory differentiation

For long, it was believed that EBs could differentiate either to ECs or to EE cells, and the terminal fate was solely dependent of the Notch signal. EBs that received a strong Notch signal would be selected for absorptive differentiation, while weak Notch activation resulted in secretory differentiation (Ohlstein and Spradling, 2007). However, recent

studies showed that EE cells came from ISCs, although they went through a very short lived transient state, term pre-EE (Biteau and Jasper, 2014; Zeng and Hou, 2015). This pre-EE cell still expresses *Dl* and *esg*, although it also expresses the secretory marker *pros* (**Fig I1C**) (Guo and Ohlstein, 2015; Zeng and Hou, 2015).

It is generally accepted that the main initiators of EE differentiation are two bHLH class II transcription factors encoded in the *achaete-scute* complex (AS-C): **Scute** (Sc) and **Asense** (Ase) (Bardin *et al.* 2010; Amcheslavsky *et al.* 2014), although it has not been investigated if they work in parallel or if one regulates the other. Sc and Ase promote the initiation of the secretory differentiation by expressing the differentiating gene *pros* that will lead to terminal differentiation (Wang *et al.*, 2015; Zeng and Hou, 2015). Pros, aside from promoting terminal differentiation into EE cell, also blocks cell cycle genes induced by Sc, arresting any mitotic event in the terminal EE cell (Chen *et al.*, 2018).

Although *sc* is expressed in almost all ISCs, its expression levels are not homogeneous and it is weakly expressed in the majority of cells, while small subsets of cells have a higher expression (Chen *et al.*, 2018). This suggests that *sc* levels need to surpass a certain threshold to promote secretory differentiation.

Moreover, it has been reported that Sc drives expression of a member of *E(spl)* complex, which belongs to the class VI of the bHLH family: *E(spl)m8* (Chen *et al.*, 2018). This is a negative regulator of the expression of *sc* and together they form a negative feedback loop to maintain low expression levels of *sc* in ISCs. *sc* expression is also repressed by Notch. Although the mechanism is not clear, knock-down of *Notch* in progenitor cells elevates *sc* expression in all ISCs (Chen *et al.*, 2018; Li *et al.*, 2017a). Therefore, in Notch knocked down guts, secretory differentiation is largely increased (Ohlstein and Spradling, 2006; Ohlstein and Spradling, 2007). Interestingly, it has been reported that when an ISC asymmetrically divide to generate a new ISC and a pre-EE, the expression of *Dl* in the pre-EE induces weak activation of Notch in the ISCs (Guo and Ohlstein, 2015). Moreover, this paper also shows that Pros, which is expressed just before mitosis, is asymmetrically segregated to the pre-EE cell (Guo and Ohlstein, 2015).

**Tramtrack69** (Ttk69) has also been postulated as an inhibitor of the expression of *sc* and *ase* (Wang and Xi, 2015; Wang *et al.*, 2015). Ttk69 is an isoform of *tramtrack*, a C2H2

zinc-finger transcription factor and is expressed at low levels in ISCs and EE cells, medium levels in EBs and high levels in ECs (Wang *et al.*, 2015). By inhibiting the expression of *sc* and *ase*, Ttk69 suppress the formation of EE cells. Loss of function of *ttk* produces an increment of secretory differentiation, indicating that *sc* is being actively repressed to avoid an excess of EE cells. Moreover, Ttk69 is needed in EBs to maintain the absorptive fate, as knocking down the expression of *ttk69* in this transient state induces secretory miss-differentiation and cell proliferation (Wang *et al.*, 2015).

In the peripheral nervous system (PNS), Ttk is repressed by Numb to specify distinct fates after cell division (Guo *et al.*, 1995). It was proposed that Numb represses Notch signalling and Notch activates *ttk* expression (Guo *et al.*, 1996). In the *Drosophila* midgut, overexpression of Numb induces secretory differentiation, while in the absence of *numb* EE cells are not formed (Sallé *et al.*, 2017). Sallé *et al.* showed that in loss of function of *Numb* clones, knocking down the expression of *ttk69* rescued the differentiation of secretory cells, indicating that in this context Numb could be also repressing *ttk69*.

Interestingly, in the gut the expression of *ttk69* is unaffected by knocking down the expression of *Notch*, indicating that in the adult midgut, Ttk69 and Notch act in parallel (Wang *et al.*, 2015). However, a recent report from the same group showed that Notch signalling inhibits the expression of *phyllopod* (*phyl*) which encodes an adaptor protein that facilitates ubiquitination of Ttk69 by the E3 ubiquitin ligase Sina, and its proteolytic degradation (Yin and Xi, 2018). Moreover, Sc induces the expression of *phyl*, promoting a positive feedback loop.

The regulation of *sc* and *ase* is not limited to the control of their expression. *esg* in progenitor cells halts EE cell formation by a different mechanism (Loza-coll *et al.*, 2014). CHIP-seq analysis of the *pros* promoter region showed that both *Esg* and *Sc* share the same binding site, suggesting that *Esg* competes with *Sc* to prevent *pros* expression EE differentiation (Li *et al.*, 2017a).

It was also suggested that EE cells could be inhibiting secretory differentiation in ISCs through Robo2 receptor, which is expressed in progenitor cells (Biteau and Jasper, 2014). Knock down of *Robo2* in progenitor cells increases the number of EE cells, while overexpression of *Robo2* in progenitor cells does not reduce the number of secretory

differentiation (Biteau and Jasper, 2014). However, the mechanism whereby Robo2 is activated is unknown. While EE cells express Slit, a ligand for Robo2, *slit* knock down has no effect on EE differentiation (Biteau and Jasper, 2011; Sallé *et al.*, 2017).

#### B5.2. Secretory differentiation in the mammalian system

The mammalian gut is the organ that contains more hormone producing cells (Rehfeld, 1998). However, this cell population is not the predominant in the intestine by far. This is due to a regulation in the differentiation that promotes that more cells differentiate into the absorptive cells to replenish the lost ECs. The Notch signalling pathway is key for the correct cell fate acquisition (see section B2.1.2). The main target of Notch is *Atoh1*, which is repressed by Her/Hes genes, the E(spl) mammalian homologues (Kim and Shivdasani, 2011). Lineage tracing of *Atoh1*<sup>+</sup> progenitors showed that this cells could differentiate into all the secretory cells (Yang *et al.*, 2001). In this study Yang *et al.* also observed that loss of *Atoh1* in the gut resulted in a failure to generate Paneth, EE or goblet cells (Yang *et al.*, 2001). Moreover, expression of an *Atoh1* transgene resulted in secretory differentiation, suggesting that *Atoh1* is sufficient to induce secretory differentiation (VanDussen and Samuelson, 2010). *Atoh1*<sup>+</sup> secretory progenitors further specialize into EE progenitors or into Paneth cells/Goblet cells progenitors, and this is specification is dependent on the expression of the *Atoh1* downstream target *Gfi1*, which is only expressed in Paneth cells/Goblet cells progenitors. *Gfi1*<sup>-/-</sup> mice could not generate Paneth cells and only few goblet cells, while there was an increased amount of EE cells (Shroyer *et al.*, 2005).

*neurogenin 3* (*ngn3*) is a bHLH expressed in the EE progenitors (*Atoh1*<sup>+</sup> *Gfi*<sup>-</sup>). *Ngn3* is required for the specification into the EE fate, as *ngn3*<sup>-/-</sup> mice lack EE cells, but all other cell types were present normally (Bjerknes and Cheng, 2006; Jenny *et al.*, 2002). *ngn3* expression is regulated by *Atoh1* (Bjerknes and Cheng, 2006).

It is remarkable that in *Drosophila* and mammals bHLH class II are the promoters of secretory differentiation. However, in the *Drosophila* midgut Atonal (*Ato*) is not essential for cell fate (Bardin *et al.*, 2010). However, the *Ngn3* *Drosophila* homolog, target of Poxn (*Tap*), is also necessary for secretory differentiation (Hartenstein *et al.*, 2017). Interestingly, one of the mammalian homologs of Sc, achaete-scute-complex like 2

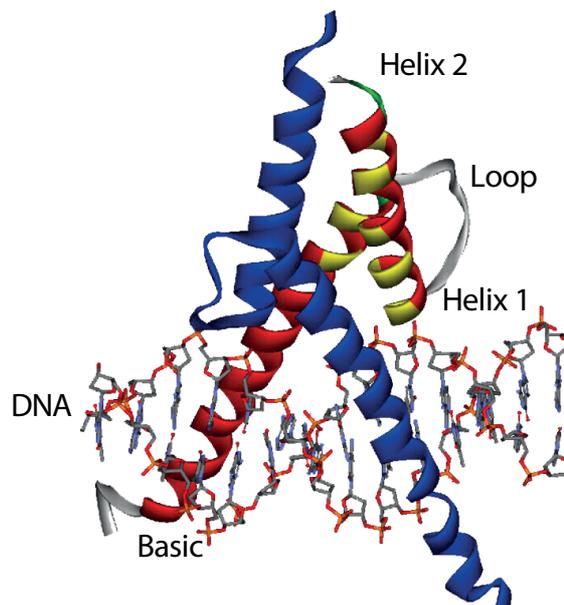
(Ascl2) (Johnson *et al.*, 1990), is important to maintain the stem state but not to induce secretory differentiation (Schuijers *et al.*, 2015; van der Flier *et al.*, 2009). It is also interesting that the secretory Paneth cells support stem cells in mammals (Sato *et al.*, 2011), as some reports show that EE cells could also secrete neuroendocrine hormones to control negatively ISC proliferation (Scopelliti *et al.*, 2014).

## C. bHLH factors

bHLH proteins are of high importance for the correct fate specification in the *Drosophila* and mammalian midgut. Moreover, they are intimately related to the Notch pathway, which is also important in ISC self-renewal and differentiation. In this section we will review the bHLH transcription factor family.

### C1. General remarks

bHLH transcription factors are an evolutionary conserved family. Almost all the members of this family share two main domains: the basic domain, which consist of basic residues that are important for DNA binding, and the helix-loop-helix, necessary for dimerization with other bHLH factors (**fig. I6**) (Murre *et al.*, 1989a).



**Figure I6. bHLH structure**

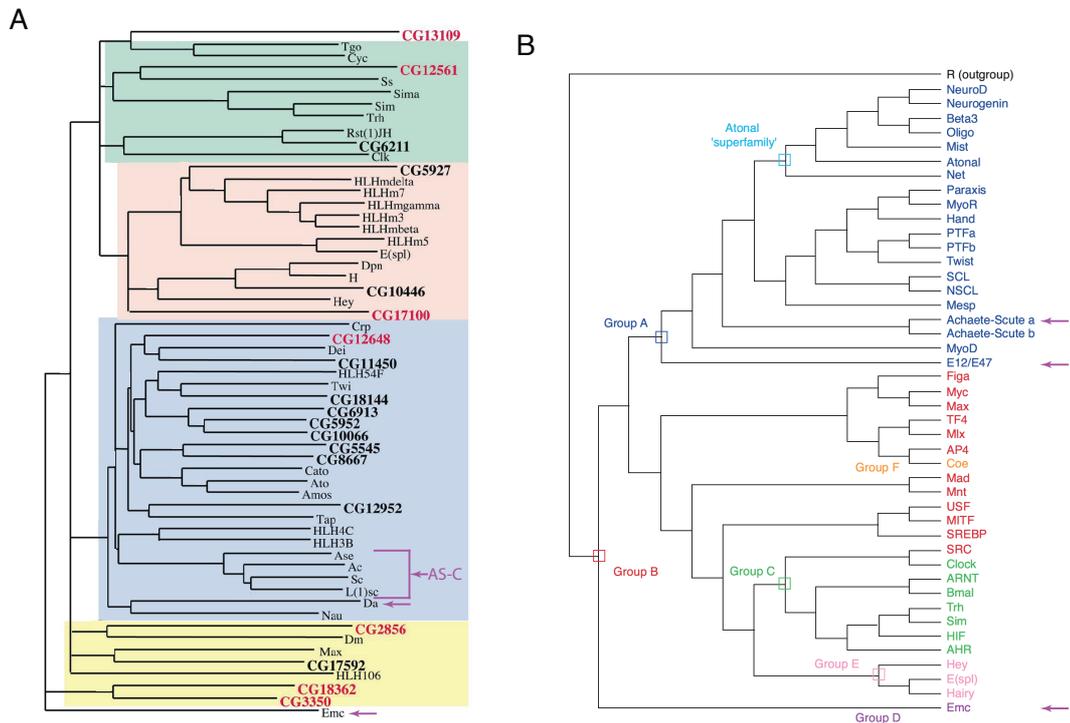
Almost all bHLH proteins (except class V factors) share two structural domains. They contain a dimerization domain formed by two amphipatic  $\alpha$ -helices bound with a loop (Helix-loop-Helix or HLH) and a DNA interaction domain (the basic). Through their HLH domain, bHLH factors form dimers, which then can recognize specific sequences in the DNA to regulate the expression of target genes.

The bHLH family is subdivided in different classes based on the sequence similarity (**fig. I7A**) (Murre *et al.*, 1994):

- Class I bHLH factors, also known as E proteins, are generally ubiquitously expressed in the organisms and form homo- or heterodimers (Murre *et al.*, 1989a).
- Class II bHLH factors have a more restricted expression and heterodimerize with class I members (Murre *et al.*, 1989a).
- Class III bHLH factors contain a leucine zipper after the HLH domain (Zhao *et al.*, 1993).
- Class IV bHLH factors can dimerize with class III members or with themselves (Blackwood and Eisenman, 1991; Blackwood *et al.*, 1992).
- Class V HLH factors lack the basic domain and only contain the HLH dimerization region. Bind with class I and II to inhibit their DNA binding (Banezra *et al.*, 1990; Ellis *et al.*, 1990; Garrell and Modolell, 1990; Van Doren *et al.*, 1991).
- Class VI bHLH factors contain a proline in the basic domain (Klamt *et al.*, 1989; Rushlow *et al.*, 1989).
- Class VII bHLH factors contain a bHLH-PAS domain (Crews, 1998).

Other authors have done a second classification of all bHLH factors depending on their sequences, DNA binding regions, protein regions and additional domains (**fig. I7B**) (Atchley and Fitch, 1997; Jones, 2004; Ledent *et al.*, 2002). With this classification, class I and class II factors form group A which bind CAGCTG or CACGTG E boxes. Class III and class IV form group B which bind CACGTG or CATGTTG binding regions. Class VII factors with the PAS domain are the only members of group C and bind to ACGTG or GCGTG sequences. Class V factors, which do not bind the DNA, form group D. Finally, class VI members form group E and bind preferentially to N boxes with CACGCG or CACGAG sequences (reviewed in Jones, 2004).

Next, we will review class I, II and V as they are the most relevant for this thesis.



**Figure 17. Classification of bHLH factors**

**A.** Phylogenetic tree of *Drosophila* bHLH factors based on their bHLH domain sequence similarity. Da (class I), the *ac/sc* complex (AS-C, class II) and Emc (class V) are indicated (magenta arrows). Modified from Paeyreffite *et al.*, 2001.

**B.** Phylogenetic tree of metazoan bHLH factors based on their domain structure and the similarity of both their aminoacid sequence and of their DNA binding sequence. E12/E47 (da in *Drosophila*, Group A), Achaete-Scute (Group A) and Emc (Group D) are indicated (magenta arrows). Taken from Ledent *et al.*, 2002.

### C1.1. Class I bHLH transcription factors

Da is the only class I factor in *Drosophila*. Da was named after the homozygous deficiencies of the *da* locus (chromosome 2L) in females, which caused that all their progeny were males (Sandler, 1972). *da* expression is ubiquitous and it is expressed almost in all tissues (Cronmiller and Cummings, 1993). Da has three mammalian homologs: E2.2, HEB and E2A. E2A has two spliced products, E47 and E12 (Kamps *et al.*, 1990; Murre *et al.*, 1989b; Walker *et al.*, 1990).

The Da protein contains a basic domain, an HLH domain with two amphipathic helices that mediated dimerization, a repression domain (REP) and two trans-activator domains (TAD): activation domain 1 and loop-helix (LH) (Murre *et al.*, 1989a; Voronova and Baltimore, 1990; Wong *et al.*, 2008; Zarifi *et al.*, 2012a). Through the HLH domain, Da can form homodimers or heterodimers with members of bHLH class II factors (Cabrera and Alonso, 1991).

### C1.2. Class II bHLH transcription factors

There are multiple bHLH class II in *Drosophila* and they typically heterodimerize with Da, as the majority of class II members cannot homodimerize or form heterodimers with other class II factors (Cabrera and Alonso, 1991; Castanon *et al.*, 2001). Twist, a bHLH factor required for mesoderm specification (Leptin, 1991), is an exception to this rule and can form homodimers (Castanon *et al.*, 2001).

The most paradigmatic class II bHLH factors are the AS-C, which have been extensively studied, especially in neurogenesis (Alonso and Cabrera, 1988; Campuzano *et al.*, 1985; Cubas *et al.*, 1991; Garcia-Bellido, 1979; Villares and Cabrera, 1987). Located in the X chromosome, it comprises around 90 kb and contain four genes: *achaete (ac)*, *scute (sc)*, *lethal of scute (l'sc)* and *asense (ase)*, which heterodimerise with Da to promote neural differentiation in *Drosophila*.

### C1.3. Class V HLH factors

Extramacrochaetae (Emc) is the only member in *Drosophila* of the HLH class V. It was identified during a screening to find negative regulators of the AS-C (Botas *et al.*, 1982). Botas *et al.* used an extra copy of the AS-C in combination with X-ray mutated flies, and they identified 20 flies that presented extra macrochaetae (large bristles) and all mapped their mutations in a locus on the third chromosome. Emc mammalian homologs are the inhibitors of DNA binding or inhibitors of differentiation (Id) family (Banezra *et al.*, 1990). It comprehends 4 members: Id1, Id2 (Sun *et al.*, 1991), Id3 (Christy *et al.*, 1991) and Id4 (Riechmann *et al.*, 1994). In different cancers, such as astrocytic cancer or pancreatic cancer, the high expression of Id genes is correlated with poor differentiation of cancer cells (reviewed in Ruzinova and Benezra, 2003; Sikder *et al.*, 2003).

Emc also contains the dimerization HLH domain but lacks the basic domain to bind the DNA (Ellis *et al.*, 1990; Garrell and Modolell, 1990). Thus, Emc does not have transcriptional activity. However, it can dimerize with high affinity with class I and class II bHLH transcription factors, sequestering them and inhibiting their binding to DNA (Cabrera *et al.*, 1994; Martínez *et al.*, 1993; Van Doren *et al.*, 1991).

## **C2. Notch and bHLH factors in the development of the peripheral nervous system**

The *Drosophila* notum is covered by a regular pattern of small bristles or microchaetae (mcs) and large bristles or macrochaetae (MCs), which are mechanosensory organs. Adult flies contain 22 MCs (11 per side with invariable positions) and about 200 mcs (Hartenstein and Posakony, 1989). Each bristle consists of four cells (shaft, glial cell, socket and neuron) that arise from a single cell, the sensory organ precursor cell (SOP) (Hartenstein and Posakony, 1989).

Each SOP arises from a pro-neural cluster (PNC), a group of cells that express AS-C genes, also known as pro-neural genes (Cubas *et al.*, 1991; Romani *et al.*, 1989; Skeath and Carroll, 1991). Initially, all cells in the PNC are equipotent epithelial cells, and for one of them to become an SOP, it must have pro-neural activity above a threshold. When *ac* and *sc* are not expressed, adult flies do not develop any bristles (Moscoso del Prado and Garcia-Bellido, 1984). Pro-neural genes induce their own expression to create a positive feedback loop and accumulate at high levels to overcome the threshold (Culí and Modolell, 1998).

To avoid accumulation of pro-neural genes in all cells of the PNC, *emc* is expressed through the wing disc in a complex pattern, being also expressed in PNCs (Cubas and Modolell, 1992). Therefore, Emc can bind with Da and pro-neural genes to impede that low levels of AS-C genes could auto-activate and form a SOP (Martínez *et al.*, 1993; Van Doren *et al.*, 1992). Hence, inhibiting the expression of *emc* in the notum resulted in the formation of extra MCs. Interestingly, *ac* and *sc* are not required when *emc* is not expressed, and when these genes are missing SOPs still emerge (Troost *et al.*, 2015). As Da can form homodimers that are able to bind to the same E-boxes than Da:Sc or Da:Ac (Cabrera and Alonso, 1991), it is possible that Da:Da is enough to induce neurogenesis.

This hypothesis is supported by the formation of SOP when *da* is over-expressed in the absence of the whole AS-C (Jafar-Nejad *et al.*, 2006; Zarifi *et al.*, 2012a) or when a tethered Da:Da with a flexible polypeptide chain to force Da homodimerization is expressed (Wang and Baker, 2015). Therefore, Emc would also be inhibiting the formation of SOP by Da:Da in the absence of pro-neural genes, and neurogenesis is only possible when the levels of pro-neural genes and Da are sufficiently high to homo- or heterodimerize.

Once the threshold of proneural activity is surpassed, the cell that becomes neural inhibits neural differentiation in the surrounding cells, forcing them to differentiate into epidermal cells. This process, mediated by the Notch pathway, is termed lateral inhibition (Simpson, 1990). Loss of Notch activity induces all cells within the PNC to adopt the SOP fate (Heitzler and Simpson, 1991). Some authors argue that pro-neural genes induce the expression of *Dl* into the SOP to inhibit all the surrounding cells, as they are also expressing pro-neural genes at high levels (Kunisch *et al.*, 1994). When cells around the SOP activate Notch, they express E(spl) proteins, which bind to Da:Sc TADs and inhibit both AS-C self-stimulation and downstream target activation (Culí and Modolell, 1998; Giagtzoglou *et al.*, 2003; Zarifi *et al.*, 2012a).

Importantly, the range of activation of Notch in the PNC seems to be of one cell diameter (Troost *et al.*, 2015), contrary to previous reports where it was proposed that filopodia extended the range of Delta signalling to several cell diameters (Cohen *et al.*, 2010; De Jossineau *et al.*, 2003). This implies that cells in the PNC that are not in direct contact with the SOP are activating Notch through mutual inhibition with other non-SOP cells in the PNC, and the selection of the SOP is mostly due to the relative levels of bHLH and Notch is only a safeguarding mechanism (Troost *et al.*, 2015).

Interestingly, *ase* seems to be expressed only in the SOP and its expression is partly downstream of pro-neural genes (Jarman *et al.*, 1993). Therefore, *ase* is a pan-neural gene necessary for the specification of the SOP (Brand *et al.*, 1993).

### **C3. bHLH proteins in the *Drosophila* developing eye**

During the third larval instar, in the *Drosophila* eye precursor, the eye imaginal disc, appears a fold in the epithelium called the morphogenetic furrow (MF), which starts the

ommatidia differentiation. The MF progresses from the posterior margin inducing differentiation of anterior cells into photoreceptors. This is controlled by the diffusible ligand Hh, expressed in the posterior margin. Hh travels into the anterior compartment, inducing cells to express *hh* to prolong the signalling, as well as *decapentaplegic (dpp)*. In addition, Hh activates expression of the bHLH factor class II *ato*, which in turn promotes differentiation into the photoreceptor R8. The refining of *ato* expression is driven by Notch lateral inhibition (reviewed in Roignant and Treisman, 2010).

*ato* expression is also controlled by the retinal determinant factor Eyeless (Ey) which synergistically interacts with Da:Da homodimers to drive the *ato* transcription (Brown *et al.*, 1996; Tanaka-Matakatsu *et al.*, 2014). In the wing disc, Da can drive its own expression for a positive feedback loop, and at the same time drive *emc* expression, to create a negative feedback loop (Bhattacharya and Baker, 2011). However, Hh and Dpp downregulate *emc* expression in the MF, inhibiting the negative feedback loop and elevating the levels of Da (Lim *et al.*, 2008). Outside the MF, Emc restricts the progression of the MF by negative regulation of Hh signalling (Spratford and Kumar, 2013).

There are conflicting reports about the requirement of Da to dimerize with Ato to induce the differentiation to R8 cells. On one hand, it was shown that cells ectopically expressing *ato* outside of the MF were not able to undergo neural differentiation, and only co-expression of *da* and *ato* resulted in neural differentiation in this region (Bhattacharya and Baker, 2011). On the other hand, expression of *da:ato* forced heterodimers could not induce retinal differentiation ectopically (Tanaka-Matakatsu *et al.*, 2014). This suggests that Da:Ato dimers may be dispensable for neural differentiation in the fly retina, and that the neural differentiation induced by co-expression of *ato* and *da* is not due to their heterodimerisation, but to a combined titration of Emc that frees additional Ato to induce neural differentiation.

Emc in the eye disc has functions beyond the repression of Ato and Da. For instance, it contributes to the specification of the photoreceptor R7 (Bhattacharya and Baker, 2009) and to the early dorso/ventral patterning and planar cell polarity (Spratford and Kumar, 2015a), in both cases acting downstream of the Notch pathway. Interestingly, these Emc functions seem to be Da independent. However, it was shown that Emc was acting

downstream of Notch to promote proliferation in the eye disc by titrating Da (Spratford and Kumar, 2015b). Therefore, in the eye disc, Notch and bHLH factors play a crucial role for the correct development of the tissue.

#### **C4. bHLH proteins in the mammalian gut**

bHLH transcription factors play a role in many differentiation processes in mammals. The first dimerization studies of bHLH proteins were to show that E12 and E47 could form dimers with the class II factor MyoD. These dimers induce muscle specific genes, like the muscle kreative kinase, to promote myogenesis, while Id proteins could inhibit these dimerizations (Murre *et al.*, 1989a; Neuhold and Wold, 1993).

bHLH factors are also important in fate decisions in the mammalian gut (see section B). Indeed, the secretory differentiation is completely dependent of *Atoh1*. Also, *Ascl2* is expressed in Lgr5<sup>+</sup> cells downstream of Wnt pathway (Jubb *et al.*, 2006; van der Flier *et al.*, 2009). *Ascl2* forms an auto-activating loop that is dependent on Wnt signalling, and at the same time, co-operates with  $\beta$ -catenin to drive transcription of stem genes (Schuijers *et al.*, 2015). *Ascl2* dimerizes with E2A and HEB, whose expression is restricted to the crypt (van der Flier and Clevers, 2009). In experiments where *Ascl2* was overexpressed, crypts were hyperplastic, whereas villi were only partly affected, as differentiated cells do not express E proteins and cannot form functional dimers. Moreover, *Ascl2* knock out induced loss of stemness in the crypts, whereas *Ascl2* up-regulation induced invasiveness, tumorigenesis and metastasis in colorectal cancer (Basu *et al.*, 2018)

Id proteins also play a role in the intestine, as Id1 is expressed in the crypts (CBCs, +4 cells and TA cells) and Id2 and Id3 are expressed in more mature cells outside of the crypts (Wice and Gordon, 1998; Zhang *et al.*, 2014). In mice reared in normal conditions, Id1 loss of function had no effects on ISC function. However, when mice were fed with dextran sodium sulfate to produce colitis, *Id1* knockout mice had more severe symptoms, with shorter colons, ulcerations, loss of crypt integrity and inflammation (Zhang *et al.*, 2014). Conversely, expression of *Id1* in the mouse small intestine induces the formation of adenomas. Moreover, it also produced a down-regulation of *Id2* and *Id3m* while E12 and E47 levels were not affected (Wice and Gordon, 1998).

ID proteins are important not also in adult homeostasis but also during development. Nigmatullina *et al.* found that Lgr5<sup>+</sup> cells are formed at E13.5 embryonic stage. They also observed that *Id2* knock out mice could develop Lgr5<sup>+</sup> cells at E9.5, resulting in an augment of Lgr5<sup>+</sup> cells at later stages and the appearance of neoplastic cells. Therefore, *Id2* restricts the specification of Lgr5<sup>+</sup> cells in early stages (Nigmatullina *et al.*, 2017).

In conclusion, bHLH transcription factors control many differentiation programs, including intestinal fates, both in mammals and *Drosophila*. Notch and lateral inhibition programs often involve the regulation of bHLH factors and their transcriptional programs. In the *Drosophila* midgut, Notch-Delta signaling is important for the correct fate acquisition and to exit the stem compartment. Moreover, bHLH transcription factors play a major role in differentiation and stem cell function in mammals. This make them of obvious interest of study in the *Drosophila* intestine. Indeed, Bardin *et al.* (2010) showed that *Da* is important to maintain the progenitor compartment and *Sc* to exit it into the secretory fate. However, much is unknown about how the bHLH network functions in the *Drosophila* midgut. Therefore, in this thesis we will investigate the role of different bHLH factors in depth to find a mechanism responsible for the maintenance of the progenitor state and differentiation. We will mainly focus in *Emc*, *Da* and *Sc*.

It is not known if *emc* is expressed in the *Drosophila* midgut, and as the main inhibitor of class I and II bHLH factors, it could play a major role to impair *Da* and *Sc* function. Moreover, in the formation of wing margins and veins, *emc* expression is regulated by Notch (Baonza *et al.*, 2000), which is the main differentiation pathway in the gut. The role of *Da* is not clear, as it has only been shown that loss of *da* induces ISCs to differentiate. Moreover, apart from the initiation of the secretory differentiation (Bardin *et al.*, 2010), recent findings have shown that *Sc* might play other roles in ISCs (Chen *et al.*, 2018).

Therefore, we will use different genetic tools and confocal microscopy to fill all these gaps and study how these three bHLH factors interact with each other to control ISC fate and maintain the homeostasis of the posterior midgut.

# **Section 2:**

# **Materials and Methods**

## 1. Fly stocks

Most fly stocks were sourced from the *Drosophila* stock centres in the USA (BDSC: Bloomington *Drosophila* Stock Center), Japan (DGRC: *Drosophila* Genetic Resource Center; NIG: National Institute of Genetics), and Austria (VDRC: Vienna *Drosophila* Resource Center); for these, a stock reference number is provided. Strains sourced from the scientific community are referenced to the person who provided it and whenever possible to the publication that described that strain, mutation or transgenic insertion.

Driver lines:

Driver lines are flies that express the yeast protein Gal4 under the control of a native gene promoter. Gal4 is a transcription factor that binds with a specific enhancer to drive gene expression: the upstream activated sequences (UAS). This method allows expression of certain genes or reporters in specific tissues.

***esg<sup>ts</sup>***: *y, w; esg<sup>NP7397</sup>/CyO; tub-GAL80<sup>ts</sup>, UAS-GFP/TM6B* (Jiang and Edgar, 2009).

***esg-F/O***: *y, w; esg-Gal4, UAS-GFP, tub-Gal80<sup>ts</sup>/CyO; UAS-flp, Act5C>CD2>Gal4/TM6C* (Jiang *et al.*, 2009).

***Rab3<sup>ts</sup>***: *w; tub-GAL80<sup>ts</sup>, UAS-GFP /CyO; Rab3-Gal4/TM6B*.

***GBE-Su(H)<sup>ts</sup>***: *y, w; GBE-Su(H)-Gal4/CyO; tub-GAL80<sup>ts</sup>, UAS-GFP/TM6B* (Zeng and Hou, 2015).

***Myo31DF<sup>NP0001</sup>*** (also known as *Myo1A<sup>NP0001</sup>*) (DGRC#112001).

***eyg-Gal4*** (provided by S. Campuzano).

UAS-transgenes:

***UAS-da/CyO*** (S. Campuzano), ***UAS-da:da*** (II and III chromosomes) (Sangbin Park), ***UAS-emc<sub>5.1</sub>*** (Baonza *et al.*, 2000), ***emc<sub>EP3620</sub>*** (this insertion allows UAS-dependent control of endogenous *emc*; (Abdelilah-Seyfried *et al.*, 2000), *y, w; If/CyO; UAS-N<sup>intra</sup>* (A. Martínez-Arias), *w, UAS-crp<sub>s-250</sub>* (BDSC#43494), *w, UAS-Dicer2* (BDSC # 24646), ***UAS-sc<sub>11</sub>/Cyo*** (S. Campuzano), ***UAS-sc<sub>P</sub>/Cyo*** (S. Campuzano), ***UAS-sc<sub>42</sub>/TM6B*** (S. Campuzano), ***UAS-H/CyO*** (A. Bardin), ***UAS-p35*** (S. Campuzano), ***UAS-E(spl)-m8*** (BDSC # 26872), ***UAS-esg*** (Korzelius *et al.*, 2014).

Expression reporter genes:

*emc*CPTI002740 (DGRC #115-317), *crp*CPTI004164/SM6A (H. White-Cooper),  
*esg-lacZ*k00606/CyO (Spradling et al., 1999), *Myo1A-lacZ* (B. Edgar), *da:GFP:FPTB*  
(Expresses GFP, FLAG, PreScission, TEV and BLRP tagged da protein) (BDSC#55836).

Mosaic Analysis with a Repressible Cell Marker (MARCM) stocks:

### **Marker strains**

*y, w, hsflp1.22, tub-Gal4, UAS-GFP; tub-Gal80, FRT40A/CyO, act-GFP* (A. Bardin)

*y, w, hs-flp1.22, tub-Gal4, UAS-GFP; tub-Gal80, FRT80B/TM6B* (S. Campuzano)

*y, w, hs-flp1.22, tub-Gal4, UAS-GFP; tub-Gal80, FRT2A/TM6B* (S. Campuzano)

*w hs-Flp tub-Gal80 FRT19A; tub-Gal4, UAS-GFP/CyO* (S. Yamamoto)

### **Wild type arm strains**

*w;FRT40A* (BDSC#1646)

*w;;FRT80B* (BDSC#1620)

*y,w;;FRT2A* (BDSC#1997)

*y,w,FRT19A* (BDSC#1709)

Mutants:

*w; Df(2L)da<sup>10</sup>, FRT40A/In(2LR)Gla, Bc* (BDSC # 5531), *w;; emc<sup>AP6</sup>, FRT80B/TM6B*  
(BDSC#36544), *w;; emc<sup>AP6</sup> FRT2A/TM6B* (recombined by me using BDSC stocks  
#36544 and BDSC#1997), *w;; emc<sup>l</sup> FRT80B/TM2* (BDSC #5532), *emc<sup>LL02590</sup> FRT2A,*  
*FRT82B/TM6C* (DGRC#140.642), *y, w; crp<sup>K00809</sup> FRT40A/CyO* (DGRC #111.066), *y, w;*  
*crp<sup>KG08234</sup> FRT40A/CyO* (DGGR #111622), *y, w; crp<sup>KG00953</sup> FRT40A/CyO* (DGRC#114.622),  
*Df(1)sc<sup>B57</sup> w FRT19A/FM7g* (Bardin et al., 2010)

RNAi stocks

*y, sc; UAS-da<sup>RNAi</sup><sub>HMS501851</sub>* (BDSC # 38382), *y, v;; UAS-da<sup>RNAi</sup><sub>JF02488</sub>*, (BDSC #29326),  
*UAS-emc<sup>RNAi</sup><sub>1007-2</sub>* (NIG#409.02) (a.k.a. *UAS-emc<sup>RNAi</sup><sub>R2</sub>*), *UAS-emc<sup>RNAi</sup><sub>KK108316</sub>*  
(VDRC#100587), *UAS-emc<sup>RNAi</sup><sub>JF02300</sub>* (BDSC #26738), *UAS-Notch<sup>RNAi</sup>*, *UAS-H<sup>RNAi</sup><sub>JF02624</sub>*  
(BDSC #27315), *UAS-crp<sup>RNAi</sup><sub>KK108184</sub>* (VDRC#100565), *UAS-crp<sup>RNAi</sup><sub>GD13194</sub>* (VDRC#26886),  
*UAS-sc<sup>RNAi</sup><sub>KK100141</sub>* (VDRC#105951), *UAS-esg<sup>RNAi</sup><sub>HMS00025</sub>* (BDSC #34063).

For all knock down experiments, the RNAi transgene was co-expressed with *UAS-Dicer-2* to exacerbate the phenotype, unless explicitly indicated. Dicer-2 (*Dcr-2*) is a ribonuclease of the RNase III family that forms a complex with R2D2, which contains two dsRNA binding domains. This complex is responsible for processing long double-strand RNA and short hairpin RNA into small interference RNA (siRNA) (Liu *et al.*, 2003) and for loading the siRNA into the RNA-induced silencing complex (siRISC) (Liu *et al.*, 2006). Therefore, co-expression of *Dcr-2* facilitates the RNAi processing.

FlipOut Lineage tracing

**For combination with Gal4 drivers:** *UAS-flp*, *Act5C-FRT-CD2-FRT-Gal4/TM6C* (from *esg-FO*)

**For sparse induction:** *y w hs-flp1.22;Act5C-FRT-y<sup>+</sup>-FRT-Gal4*, *UAS-lacZ<sub>20b</sub>* (a.k.a. *Ay-Gal4*, modified from BDSC #4410)

*Drosophila melanogaster* husbandry

Adult flies were raised in standard cornmeal medium, collected daily and maintained in fresh vials with added dry yeast (food replaced every 24-48 h) until 4-7 days old. Flies for temperature sensitive experiments were reared and aged to gut maturity (~7 days) at the permissive temperature (18 °C), and then transferred to 29 °C. For MARCM experiments or *Ay-Gal4*, when flies were 4-7 days old, 1 hour heat shock at 37 °C was induced (unless otherwise indicated), and then kept at 25 °C. In both cases, flies were maintained for 7 days (food replaced every 24-48 h) prior to dissections, unless otherwise indicated.

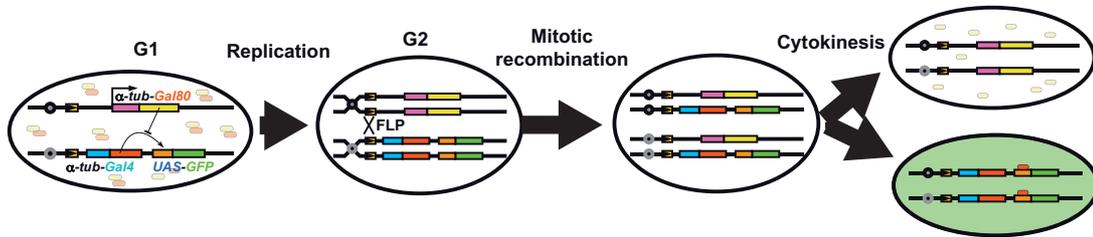
Fly food recipe (for 50L): 337 g of agar, 3.62 kg of maize, 3.75 kg of dextrose, 1.75 kg of yeast, 111 g of Hydroxybenzoic acid methyl ester, 1.3 L of absolute ethanol, 175 ml of propionic acid, 48.5 L of dH<sub>2</sub>O.

## 2. Lineage Tracing experiments

MARCM clones

Mosaic Analysis with a Repressible Cell Marker (MARCM) is a lineage tracing technique based on mitotic recombination (Lee and Luo, 2001). Flies were heat shocked for one hour to induce expression of *FLPase* (*flp*). The FLP can induce mitotic recombination on

FRT sequences in trans to exchange sister chromatids. Therefore, if one of the chromatids was carrying a mutation, there is a 50% chance that one of the daughter cells will inherit this mutation. *tub-Gal80* loss in the daughter cell carrying the mutation allows the expression of *UAS-GFP* as a marker and other desired transgenes (**fig. M1**).

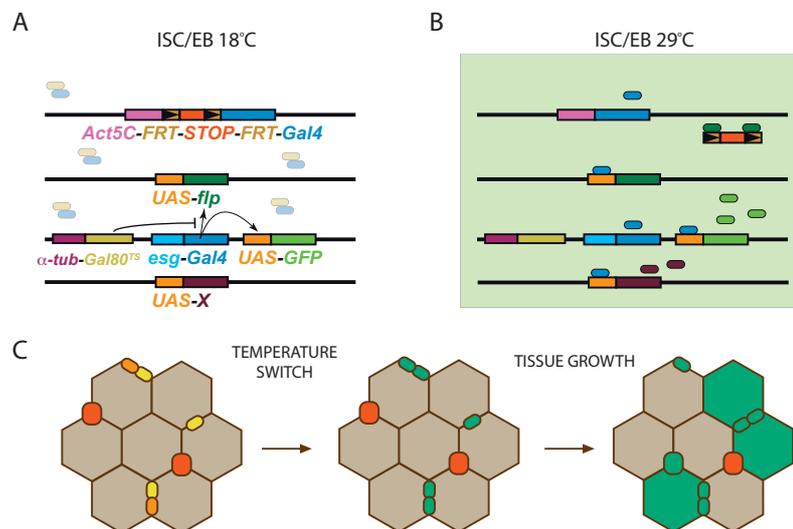


**Figure M1 Clonal labeling for lineage analysis**

All cells have Gal4 repressed by the expression of *Gal80*. The expression of *flp* induces mitotic recombination, inheriting one daughter cell the *tub-Gal80* while the other one does not and is labeled with GFP. If the homologous chromosome carries a mutation, GFP<sup>+</sup> cells will be homozygous for this mutation. Author: Joaquín de Navascués.

## FlipOut

The flipOut cassette contains a constitutive driver (typically a ubiquitous driver such as *Act5C-FRT-STOP-FRT-Gal4*), that will be activated by the expression of a *flp* to remove the STOP cassette. In this thesis we used two different ways to express the *flp*: 1) Using a



**Figure M2 escargot-FlipOut**

**A-B.** Progenitor cells at 18°C express Gal80, which represses Gal4. At 29°C, Gal4 is active and drives expression of the desired gene (X), the *flp* and *GFP* to label progenitor cells. The Flp induces excision of the STOP cassette, activating the expression of *act5C-Gal4*. Therefore, after cell differentiation, cells which do not express *esg* will still be expressing *Gal4*.

**C.** After the temperature switch, all progenitor cells express the desired genes and GFP. When the tissue grows, differentiated cells will sustain gene and GFP expression.

specific driver and a *UAS-flp* (**fig. M2**) or 2) using a *hs-flp* to induce clones. With both mechanisms, all induced cells will be labeled and expressing the correspondent transgenes. While the first mechanism affects a determined pool of cells depending on the driver, the second system only affects random cells. Moreover, when these cells divide or differentiate, they will maintain *Act5C-Gal4* active.

Although the generation of adult flies with the FlipOUT method is slower than MARCM (flies grow at 18 °C instead of 25 °C), fewer fly generations are needed for the final cross. Moreover, with FlipOUT, the number of recombinations needed are reduced. Therefore, the FlipOUT method allows us to do experiments much faster.

### 3. Immunohistofluorescence and imaging

Three different fixation methods were performed during this study:

**Paraformaldehyde fixation:** Flies were dissected for 20 minutes in phosphate-buffered saline (PBS) to collect guts in 4% PFA in ice. Guts were left 2 hours with no agitation at Room Temperature (RT).

**Formaldehyde-Methanol fixation:** Flies are dissected for 8 minutes in ice-cold PBS and the guts are collected in a basket immersed in cold PBS. Then guts are transferred to the interface between formaldehyde 3.7% in PBS and heptane. After 15 minutes, the formaldehyde and heptane are removed and methanol is added. 15 minutes later, guts are re-hydrated by rinsing progressively in 75%, 50% and 25% methanol solutions in PBS with 0.1% Triton-X100 (PBT).

**Heat fixation:** Flies are dissected for 8 minutes in cold PBS and the guts are collected in a basket in cold PBS. The basket with the guts is introduced in heat fixation buffer (0,7 % NaCl + 0,05% Triton in milliQ water) in a double beaker at 95 °C for 5 seconds. The tissue is quickly cooled for 2 minutes in cold PBS.

After fixation, guts were rinsed three times with PBT and blocked with PBT:BSA (2% BSA: bovine serum albumin) for 45 minutes (three 15 minutes washes). Then, guts were incubated with primary antibody overnight with mild agitation at 4 °C in PBT:BSA. The following day guts were washed as above (3x rinses, 3x 15 minutes washes) in PBT,

incubated with the secondary antibody for 2 hours at RT and washed. Finally, guts were equilibrated with mounting medium (4% propyl Gallate in 80:20 Glycerol:PBS) for several hours, and mounted on glass slides.

Confocal stacks were obtained in a Zeiss LSM 710 with an EC Plan-Neofluar 40X oil immersion objective (numerical aperture 1.3). All stack positions were acquired in the posterior midgut. Typically, three positions along the anterior-posterior axis of the posterior midgut were acquired for each organ. In MARCM clone experiments, stacks were acquired from all clones in the posterior midgut.

Protein recognized	Host/type	Dilution*	Source (reference)
GFP	Rabbit/polyclonal	1:10000	Abcam (ab6556)
GFP	Chicken/polyclonal	1:3000	Abcam (ab13970)
$\beta$ -Galactosidase	Rabbit/polyclonal	1:10000	Abcam (ab11132)
$\beta$ -Galactosidase	Chicken/polyclonal	1:2000	Abcam (ab9361)
Armadillo	Mouse/monoclonal	1:50*	DSHB** (N27A1)
Delta	Mouse/monoclonal	1:100	DSHB (C594.9B)
Prospero	Mouse/monoclonal	1:200	DSHB (MR1A)
Prospero	Rabbit	1:1000	Yuh Nung Jan (Vaessin <i>et al.</i> , 1991)
Headcase	Mouse/monoclonal	1:50	DSHB (HDC U33)
Phospho Histone 3	Rabbit/polyclonal	1:200-500	Cell signalling (#9701)
Cropped	Rabbit/polyclonal	1:500	Michael Lehmann (King-Jones <i>et al.</i> , 1999)

\*The indicated dilutions correspond to Formaldehyde-methanol fixation. As Armadillo staining did not work with this fixation method, the dilution indicate is for heat-fixation. For heat fixation the amount of antibody used was doubled.

\*\*DSHB: Developmental Studies Hybridoma Bank

Secondary antibodies conjugated with Alexa fluorophores were from Thermo (1:500).

Host	Species reactivity	Alexa fluorophores (reference)
Donkey	Rabbit	594 (A21207)
	Mouse	594 (A21203)
Goat	Rabbit	488 (A11034), 633 (A21071)
	Chicken	488 (A11042), 633 (A21103)
	Mouse	633 (A21052)

DNA was stained with Hoescht (10 mg/ml from Sigma Aldrich, B2261) (1:5000) and incubated 2 hours with the secondary antibodies.

#### 4. Cell counting and quantification

All the stacks were maximum-intensity projected using Fiji-ImageJ (Schindelin *et al.*, 2012) for each field of view. Using the Cell Counter plugin, each cell type was labelled as follows:

1	ISC GFP <sup>+</sup>	DI <sup>+</sup> NRE <sup>-</sup> Pros <sup>-</sup> GFP <sup>+</sup> / Hdc <sup>+</sup> NRE <sup>-</sup> Pros <sup>-</sup> GFP <sup>+</sup>
2	EB GFP <sup>+</sup>	DI <sup>-</sup> NRE <sup>+</sup> Pros <sup>-</sup> GFP <sup>+</sup> / Hdc <sup>+</sup> NRE <sup>+</sup> Pros <sup>-</sup> GFP <sup>+</sup>
3	EE GFP <sup>-</sup>	DI <sup>-</sup> NRE <sup>-</sup> Pros <sup>+</sup> GFP <sup>-</sup> / Hdc <sup>-</sup> NRE <sup>-</sup> Pros <sup>+</sup> GFP <sup>-</sup>
4	EC GFP <sup>-</sup>	DI <sup>-</sup> NRE <sup>-</sup> Pros <sup>-</sup> GFP <sup>-</sup> / Hdc <sup>-</sup> NRE <sup>-</sup> Pros <sup>+</sup> GFP <sup>-</sup>
5	EE GFP <sup>+</sup>	DI <sup>-</sup> NRE <sup>-</sup> Pros <sup>+</sup> GFP <sup>+</sup> / Hdc <sup>-</sup> NRE <sup>-</sup> Pros <sup>+</sup> GFP <sup>+</sup>
6	EC GFP <sup>+</sup>	DI <sup>-</sup> NRE <sup>-</sup> Pros <sup>-</sup> GFP <sup>+</sup> / Hdc <sup>-</sup> NRE <sup>-</sup> Pros <sup>+</sup> GFP <sup>+</sup>
7	ISC GFP <sup>-</sup>	DI <sup>+</sup> NRE <sup>-</sup> Pros <sup>-</sup> GFP <sup>-</sup> / Hdc <sup>+</sup> NRE <sup>-</sup> Pros <sup>-</sup> GFP <sup>-</sup>
8	EB GFP <sup>-</sup>	DI <sup>-</sup> NRE <sup>+</sup> Pros <sup>-</sup> GFP <sup>-</sup> / Hdc <sup>+</sup> NRE <sup>+</sup> Pros <sup>-</sup> GFP <sup>-</sup>
9	Other	Different combinations from above

Cell Counter generated a table with all cells with their corresponding cell types and their spatial coordinates within the field of view. Percentages of each cell type respective the total number of cells or GFP<sup>+</sup> cells, number of cells per cluster, number of clusters of a determined size and the accumulation of all cells of a specific cell type were obtained with a python script. The Delaunay triangulation method was used to find GFP<sup>+</sup> clusters (Delaunay, 1934) (Script code in **Appendix 1**).

#### 5. Quantification in *UAS-sc* expression

The large proliferation induced by the expression of *UAS-sc* generated large clusters of DI<sup>+</sup> Pros<sup>-</sup> cells or DI<sup>+</sup> Pros<sup>+</sup>. In the majority of field of views these clusters were uncountable due the high density of cells. Therefore, we counted when possible and, in the rest, estimated the proportion of each cell population.

## 6. Quantification of Delta expression

All cells were labelled as explained in section 4 with Fiji ImageJ. Using a python script (Script code in **Appendix 2**), each image was processed separately. Channels were split and the stacks were projected.

### Extraction of Pros<sup>+</sup> cells pixel values

First, a median filter was applied to the Pros/Dl channel to remove small features. A binary mask was created using Otsu thresholding (Otsu, 1979) of the filtered image. This mask captured most of the Pros<sup>+</sup> nuclei (those with higher expression of Pros). To include in the analysis the Pros<sup>+</sup> nuclei missed by the thresholding, we created another mask, produced by creating, for each XY position where a Pros<sup>+</sup> cell was manually determined (see section 4) but not overlapping with the thresholding mask, a disc with a diameter of 3 pixels. We then combined both masks. This was further refined by marker controlled watershed transformation (Meyer and Beucher, 1990) to separate objects that may have been fused by thresholding. Pros expression for each nucleus was determined as the average intensity value of the Pros/Dl channel for each nucleus in the watershed-segmented mask.

### Extraction of Dl<sup>+</sup> cells pixel values in clones

To identify the GFP-labelled MARCM clones, a similar approach was used. First, a mask was generated by thresholding the GFP channel using Li's minimum cross entropy thresholding (Li and Lee, 1993). As this still missed a few cells, the mask was expanded by adding 3-pixel diameter discs for each GFP<sup>+</sup> cell, as manually determined, and the mask consolidated by morphological filling and closing (Diggle and Serra, 1983). Individual Dl<sup>+</sup> cells within the clone were identified by marker-controlled watershed segmentation, using the positions of the cells as manually determined. Dl expression for each cell was determined as the average intensity value of the Pros/Dl channel for each object in the watershed-segmented image.

## Calculating cell intensity

For each cell (Pros<sup>+</sup> or DI<sup>+</sup>), the mean intensity of all pixels of each cell was obtained. As the imaging conditions were set up so that the maximum intensity pixels were just below saturation (in each genetic condition), to take full advantage of the sensitivity range of the detector, a normalization was required. Since the immunoreactivity to the anti-Pros antibody is very robust and homogeneous from sample to sample, the mean cell intensity of Pros or DI was divided by the mean intensity of Pros from the pooled pixels of all Pros<sup>+</sup> cells in that field of view.

## 7. Esg-DamID binding regions

DamID data were obtained from Korzelius et al, 2014 and Loza-Coll *et al.* 2014 (accession number GSE55226). The median values out of the three replicates was plotted. Integrative Genome Viewer software was used for data visualization. The shown area corresponds with the *da* locus.

An Esg binding region was considered to exist if more than 8 and less than 20 consecutive genomic probes had an Esg-Dam/Dam ratio above 2.

## 8. Statistical analysis

Graphs and statistical analyses were generated using Prism 7 GraphPad Software or R3.5.1 using RStudio 1.1.383.

Normality was tested with the Shapiro-Wilk normality test. Parametric or non-parametric tests were used accordingly. Unless otherwise indicated, 'N=' refers to the number of fields of view.

Stacked bars were used for MARCM clones or for cell death phenotypes. The statistical analysis of this section was performed by Joaquín de Navascués. In these experiments we used binomial regression. We used Firth's bias reduced logistic regression for experiments with zero observations in one or more of the genotypes considered ('complete separation' of data (Albert and Anderson, 1984), which we performed using the R package logistf.

## 9. Generation of the *crp*<sup>p121.4</sup> mutant allele

This allele was generated by Paminder Lall during her internship in our group and by Dr. Joaquín de Navascués. The sgRNAs were designed by Aleix Puig.

Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR) technology was used with the aim of inducing a cleavage in two genomic sites of the *crp* locus at the same time. By microinjection of the pCDF4 plasmid (Port *et al.*, 2014) in Cas9-expressing blastocyst embryos, two sgRNAs can be expressed simultaneously, albeit from different promoters:

sgRNA1 (U6:1 promoter): GATTGCAACTAGAGGATTC

sgRNA2 (U6:3 promoter): GCCTTCGTGTCTCGGGAGC

The plasmids were injected in *Act5C-Cas9; FRT40A(isogenic)* (modified from BDSC#54590) embryos. Single emerging adults were crossed with *Gla, Bc/CyO* and *Gla, Bc* individuals from F1 were collected and individually crossed with flies containing the *crp*<sup>CPT100416</sup>/*CyO*, *twi-GFP* mutant allele. *Gla*<sup>+</sup> *Bc*<sup>+</sup> *GFP* flies with defects in bristle patterning were kept as a possible mutant line, as this was reported as a mild *crp* phenotype (Ashburner *et al.*, 1999). A recessive lethal was identified using High Resolution Melting PCR around the cutting regions of the Cas9. A positive mutant was identified, *crp*<sup>p121.4</sup>, which contains a small indel that creates a frameshift in the Crp coding region leading to a truncated protein lacking part of the bHLH domain.

## **Section 3: Results**

## Chapter 1 A: A network of bHLH factors controls self-renewal and bipotential differentiation in the intestine: *Emc*

### 1.A.1 Introduction

The *Drosophila melanogaster* midgut has been shown to be a suitable model to study the transcriptional regulation of adult stem cells. To date, four different transcription factors have been shown to be indispensable to maintain the progenitor state: *Da* (Bardin *et al.*, 2010), *Esg* (Antonello *et al.*, 2015; Korzelius *et al.*, 2014; Loza-coll *et al.*, 2014), *Fkh* (Lan *et al.*, 2018) and *Chn* (Amcheslavsky *et al.*, 2014a).

In addition to *Da*; *Sc* and *Ase*, which are important for the secretory differentiation, are also bHLH transcription factors (Bardin *et al.*, 2010). Therefore, it is likely that *Emc*, the HLH class V and main inhibitor of class I and class II bHLH transcription factors, could also be expressed in the midgut to impair *Da*, *Sc* and *Ase* function. Indeed, over-expression of *emc* in progenitor cells induces ISCs and EBs to differentiate into ECs, similar to when *da* is knocked down in progenitor cells (Lan *et al.*, 2018). *Emc* and its mammalian counterpart have been shown to bind bHLH class I and II to prevent the formation of functional bHLH dimers that can bind the DNA (Cabrera *et al.*, 1994; Cubas *et al.*, 1991; Sun *et al.*, 1991; Van Doren *et al.*, 1991). This titration has been shown to be fundamental in various developmental processes, especially in wing morphogenesis (Baonza and García-Bellido, 1999), correct specification of sensory organs in the wing disc (reviewed in Campuzano 2001) and correct patterning in the eye disc (Bhattacharya and Baker, 2009; Bhattacharya *et al.*, 2017; Spratford and Kumar, 2013; Spratford and Kumar, 2015a; Spratford and Kumar, 2015b). Moreover, the expression of *emc* is induced by *Da*, and the stability of *Emc* depends on its binding to *Da* (Bhattacharya and Baker, 2011), while *Emc* is a negative regulator of *da* (Li and Baker, 2018). Thus, it is likely that *Emc* regulates the function of *Da* in the midgut. This hypothesis is supported by the expression of *Id* members in the mammalian intestine (Nigmatullina *et al.*, 2017; Zhang *et al.*, 2014).

Therefore, in this chapter we will investigate the role of *Emc* in the maintenance of the intestinal homeostasis.

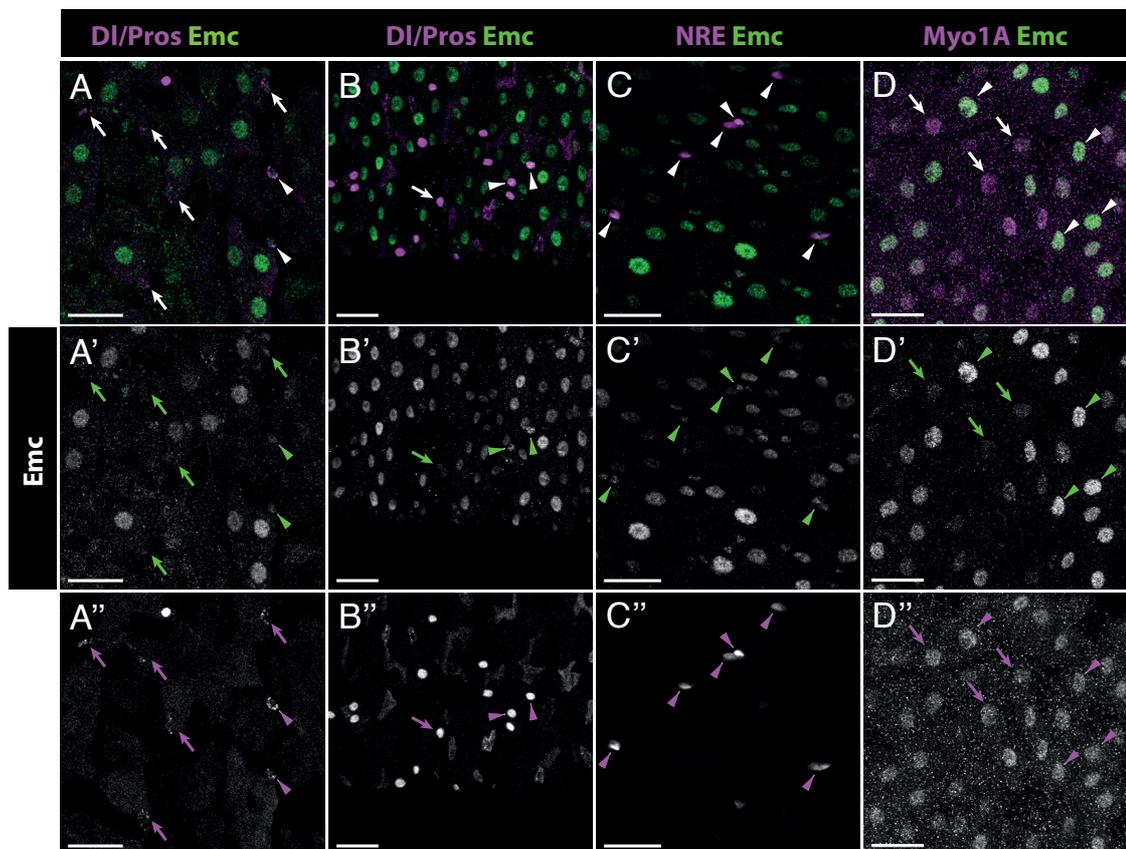
## 1.A.2. Aims

- Describe the expression pattern of *emc* in the adult gut
- Perform the genetic analysis of *emc* function in the adult gut
- Determine whether Notch regulates *emc* expression

## 1.A.3. Results

### 1.A.3.1. Expression of *emc* in the midgut

To describe the expression pattern of *emc*, we used a protein trap (*emc*<sup>CPTI2740</sup>) from the Cambridge Protein Trap Insertion (CPTI) project – a collection of transposon insertions that introduce an artificial exon encoding YFP into endogenous loci (Lowe *et al.*, 2014). This CPTI line is homozygous viable and has a wild type pattern of bristles, indicating that the resulting GFP-tagged Emc protein is functional. In parallel, we used antibodies to label different cell types (**Table 1**) and examine the co-expression with *emc* in those cells. First, we used D1 expression (cytoplasmic vesicles and membrane), which marks specifically ISCs, and the nuclear marker Pros, specific for EE cells. We could observe that *emc* can be expressed in ISCs, but at low levels and in a small subset of cells. Moreover, the expression of *emc* in EE cells is very infrequent and at low levels (**fig. 1.1AB**). Then, we monitored the expression of *GBE-Su(H)-lacZ*, a Notch activity reporter (a.k.a. Notch responsive element, *NRE*) that contains three palindromic binding site in tandem of Grainyhead (GBE) and two Su(H)-binding sites derived from the regulatory region of E(spl)-m8 (Furriols and Bray, 2001). As a Notch activity reporter, *GBE-Su(H)-lacZ* is a EB-specific reporter (Micchelli and Perrimon, 2006). Using this reporter, we could observe that *emc* expression is uniform (medium levels) within all the EBs (**fig. 1.1C**). Next, we could score the EC population using the enhancer trap *Myo1A-lacZ*, where we could find that *emc* was expressed in all ECs, although there was a large degree of expression variability (**fig. 1.1D**). In conclusion, all cell populations in the *Drosophila* midgut have the ability to express *emc*, although in ISCs and EE cells is very sporadic and at low levels. The highest expression of *emc* is found in a subpopulation of ECs.



**Figure 1.1. *emc* is expressed in all cell types at different levels**

**A-B.** *Emc<sup>CPT1</sup>* combined with  $\alpha$ -DI (cytoplasmic and membrane, ISCs) and  $\alpha$ -Pros (nuclear, EE cells). We can observe ISCs DI<sup>+</sup> and *Emc*<sup>+</sup> cells (A-A', arrow heads) and DI<sup>+</sup> *Emc*<sup>-</sup> cells (arrows). EE cells (Pros<sup>+</sup>) can be *Emc*<sup>+</sup> (B-B', arrow heads) or *Emc*<sup>-</sup> (arrows).

**C.** NRE<sup>+</sup> cells (EBs) express *emc* homogeneously (arrow heads), and the small variation of *Emc* signal match the variation of NRE intensity.

**D.** Myo1A<sup>+</sup> cells have a wide range of expression of *emc*, from very highly expressed (arrow heads) to very weak expression (arrows).

Data information: **scale bars, 20 $\mu$ m**

### 1.A.3.2. Loss of *emc* arrests terminal dedifferentiation

Next, we wanted to elucidate the function of *emc* in the *Drosophila* midgut and we generated homozygous clones of a protein null allele (*emc<sup>AP6</sup>*), a recessive lethal allele (*emc<sup>LL02590</sup>*) and a hypomorphic allele (*emc<sup>l</sup>*) using the mosaic analysis with a repressible cell marker (MARCM) technique. MARCM is a lineage tracing technique used to induce homozygotic cells in an heterozygotic background by mitotic recombination (Lee and Luo, 2001). We let the clones grow for 7 days and stained with Dl (ISC marker) and Pros (EE marker). As we could not use any specific marker for EBs and ECs, we identified EBs as Dl<sup>-</sup> Pros<sup>-</sup> diploid cells; and ECs as Dl<sup>-</sup> Pros<sup>-</sup> polyploid cells (**Table 1**). We could observe that *emc<sup>LL02590</sup>* mutant clones contained more Dl<sup>+</sup> cells (61%) with respect to wild type (35%), and less ECs (**fig. 1.2B, compare with 1.2A; quantification in 1.2F**). *emc<sup>l</sup>* clones also presented more ISCs. Interestingly, in *emc<sup>l</sup>* clones, 8% of the cells were EE (11% in wild type), while only 1% were mature ECs compared with the 20% in control conditions, meaning that in the hypomorphic clones only differentiation into the absorptive fate was being affected (**fig. 1.2D, quantification in 1.2F**). *emc<sup>AP6</sup>* mutant clones showed a mild increase of ISCs, but more importantly an increase of diploid Pros<sup>-</sup> Dl<sup>-</sup> cells, presumably EBs or early ECs (**fig. 1.2E, quantification in 1.2F**). These results suggest that *emc* is important for terminal differentiation. Moreover, hypomorphic *emc* conditions allow EE differentiation, but still inhibit the absorptive fate.

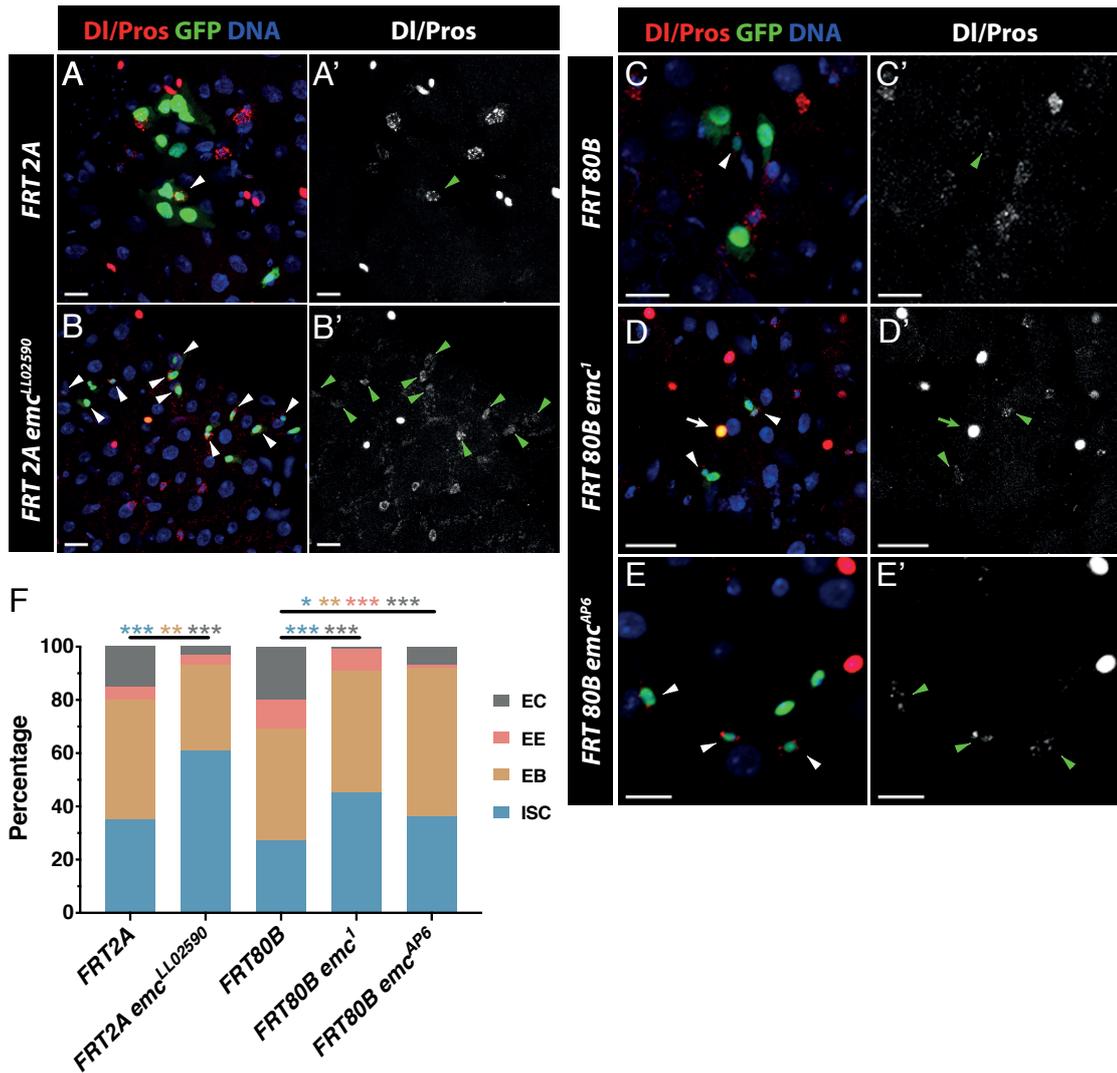
Cell type	Marker
ISC	Dl, Hdc, Esg*
EB	NRE, Hdc, Esg*, Diploidy**
EE cell	Pros
Pre-EE cells	Dl, Esg, Pros***
EC	Poliploidy, Myo1A

**Table 1 Markers used to identify the different cell types**

\* Hdc and Esg are markers for ISCs and EBs (progenitor cells). Esg is also expressed in pre-EE cells.

\*\* Diploidy can be used to identify EBs when there is no marker. In this case, diploid Dl<sup>-</sup>, Pros<sup>-</sup> cells are EBs.

\*\*\* Pre-EE cells are identified when they are expressing Pros and Dl at the same time. Pros intensity is lower than in EE cells.



**Figure 1.2. Expression of *emc* is required for terminal differentiation**

**A, B.** *emc<sup>LL02590</sup>* MARCM clones contain more ISCs. Control clones typically contains only one DI<sup>+</sup> cells (A-A', arrow head) in the clonal area. *emc<sup>LL02590</sup>* clones can contain multiple DI<sup>+</sup> cells (B-B'; arrow heads).

**C-E.** *emc<sup>1</sup>* and *emc<sup>AP6</sup>* MARCM clones has lower differentiation. Control clones contains ECs or EE cells (C-C'). *emc<sup>1</sup>* clones maintain the secretory differentiation (D-D', arrow) but arrest absorptive differentiation, with more DI<sup>+</sup> cells than control (arrow heads). *emc<sup>AP6</sup>* clones contain many DI<sup>+</sup> cells.

**F.** Quantification of the different cell types in the clonal area of the different *emc* alleles and controls (A-E). (N = 194 cells/107 clones for *FRT2A* clones, N = 213 cell/119 clones for *FRT2A emc<sup>LL02590</sup>*, N = 310 cells/119 clones for *FRT80B*, N = 112 cells/98 clones for *FRT80B emc<sup>1</sup>*, N = 208 cells/110 clones for *FRT80B emc<sup>AP6</sup>*)(\*p<0.05, \*\*p<0.01, \*\*\*p<0.001, Binomial regression. Each cell type significance level is coloured accordingly).

Data information: **scale bars, 20µm**

Consequently, we wanted to knock down the expression of *emc* in the whole progenitor compartment. We started screening three available RNAi lines: *UAS-emc<sup>RNAi</sup><sub>KK108316</sub>*, *UAS-emc<sup>RNAi</sup><sub>R2</sub>* And *UAS-emc<sup>RNAi</sup><sub>JF02300</sub>*. We proceeded with the lineage tracing system *esg-flipOut* (*esg-FO*), consisting of the *esg-Gal4* driver combined with the flipOut system (Jiang *et al.*, 2009). This temperature sensitive system contains *actin5C-FRT-STOP-FRT-Gal4*, *UAS-flippase*, *UAS-GFP* and *tub-Gal80<sup>ts</sup>*. Therefore, the expression of the flippase in progenitor cells results in the excision of the STOP cassette and the progeny of these cells will also be expressing the desired transgenes, even in cells that do not express *esg*, as the *actin5C* promoter is ubiquitous (see page 46). We used the expression of *headcase* (*hdc*) to label specifically progenitors cells (Resende *et al.*, 2017). For this system, only *UAS-emc<sup>RNAi</sup><sub>R2</sub>* showed a phenotype, while the other lines displayed wild-type phenotypes (**fig. 1.3A-D**). In *UAS-emc<sup>RNAi</sup><sub>R2</sub>* midguts, we could observe a differentiation arrest in the progenitor clusters, which kept growing without differentiating terminally (**fig. 1.3B**, quantification in **1.3E**). This result supported the results obtained with MARCM clones that *emc* loss blocks cell differentiation.

Surprisingly, the posterior midguts where *emc* was knocked down presented few progenitor clusters. In this experiment we are using a lineage tracing system and if cells were differentiating, they would be labelled. Therefore, there must be another mechanism that could involve cell death or cell extrusion. Strikingly, when we studied the composition of the *emc* knocked down clusters that were not lost, we realised that they presented multiple *Dl*<sup>+</sup> cells, meaning an increased number of ISC (**fig. 1.3F-G**), consistent with the results obtained with *emc* mutant MARCM clones. More important, the *Dl* levels in these cells were surprisingly high. Pros staining in the same channel as *Dl* is consistently brighter, and this can be used as a reference to compare the *Dl* intensity in wild type flies and *emc* knocked down (**compare fig 1.3F' and 1.3G'**). Thus, the increase of *Dl* levels in cells without *emc* suggests that *Emc* is playing a crucial role to control the expression of *Dl*.

However, as *emc* knock down with *UAS-emc<sup>RNAi</sup><sub>KK108316</sub>* and *UAS-emc<sup>RNAi</sup><sub>JF02300</sub>* lines did not show a phenotype, we validated all the *emc* RNAi lines using *eyg-Gal4*. *eyegone* (*eyg*) defines the anterior region of the adult mesothorax, the scutum (Aldaz *et al.*, 2003). Thus, knocking down the expression of *emc* during all larvae and pupal stages should result in

an increased number of bristles. We could observe that flies expressing  $UAS-emc^{RNAi_{R2}}$  had a higher density of bristles compared with the  $UAS-emc^{RNAi_{KK108316}}$  and  $UAS-emc^{RNAi_{JF02300}}$  lines which, in turn, presented more bristles than the wild-type (fig. 1.3H-K). Therefore,  $UAS-emc^{RNAi_{KK108316}}$  and  $UAS-emc^{RNAi_{JF02300}}$  lines present a weaker phenotype than  $UAS-emc^{RNAi_{R2}}$ .

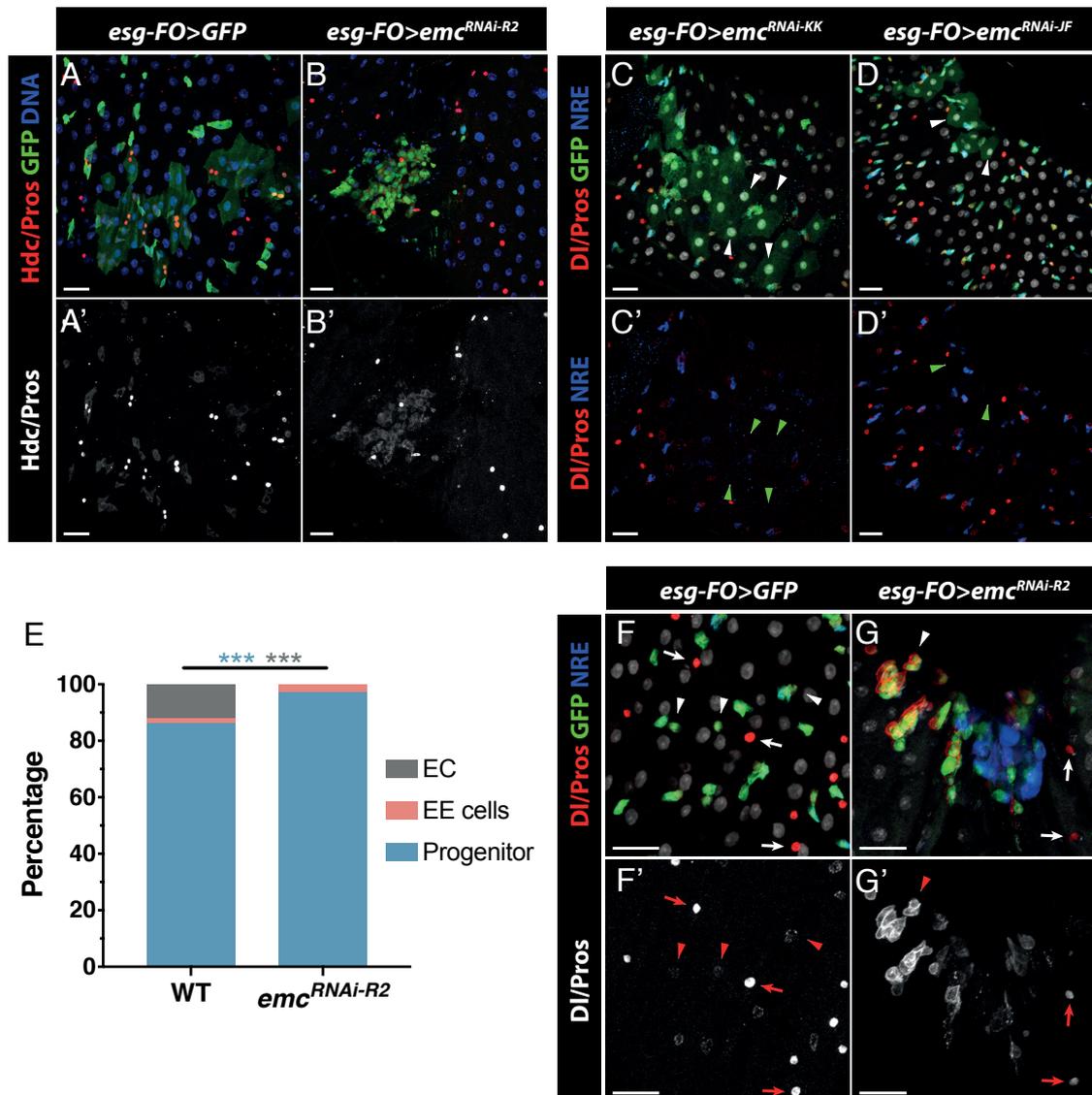
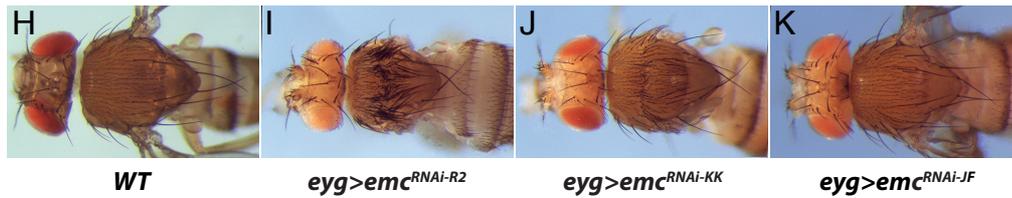


Figure 1.3. *emc* knock down arrests terminal differentiation and induces *Dl* expression (see next page for legend)



**Figure 1.3. *emc* knock down arrests terminal differentiation and induces *Dl* expression**

**A-B.** *esg-FO>emc<sup>RNAi-R2</sup>* halts terminal differentiation even after 7 days of induction. All cells are Hdc<sup>+</sup> (F-F') and there are not polyploid GFP cells as we can find in control flies (E-E'). It is also noticeable that apart of the big cluster, there are not any other GFP<sup>+</sup> or Hdc<sup>+</sup> cell in the right side of the ROI.

**C-D.** Expression of *UAS-emc<sup>RNAi-KK</sup>* (G) and *UAS-emc<sup>RNAi-JF</sup>* (H) in progenitor cells with *esg-FO* allow terminal differentiation, with polyploid *Dl<sup>-</sup> NRE<sup>-</sup> GFP<sup>+</sup>* cells (arrow heads).

**E.** Population distribution in control flies and *esg-FO>emc<sup>RNAi-R2</sup>* showing a differentiation arrest when *UAS-emc<sup>RNAi-R2</sup>* is expressed (WT N = 1005 cells, *UAS-emc<sup>RNAi-R2</sup>* N = 286 cells) (\*\*\*)  $p < 0.001$ , Binomial regression. Each cell type significance level is coloured accordingly).

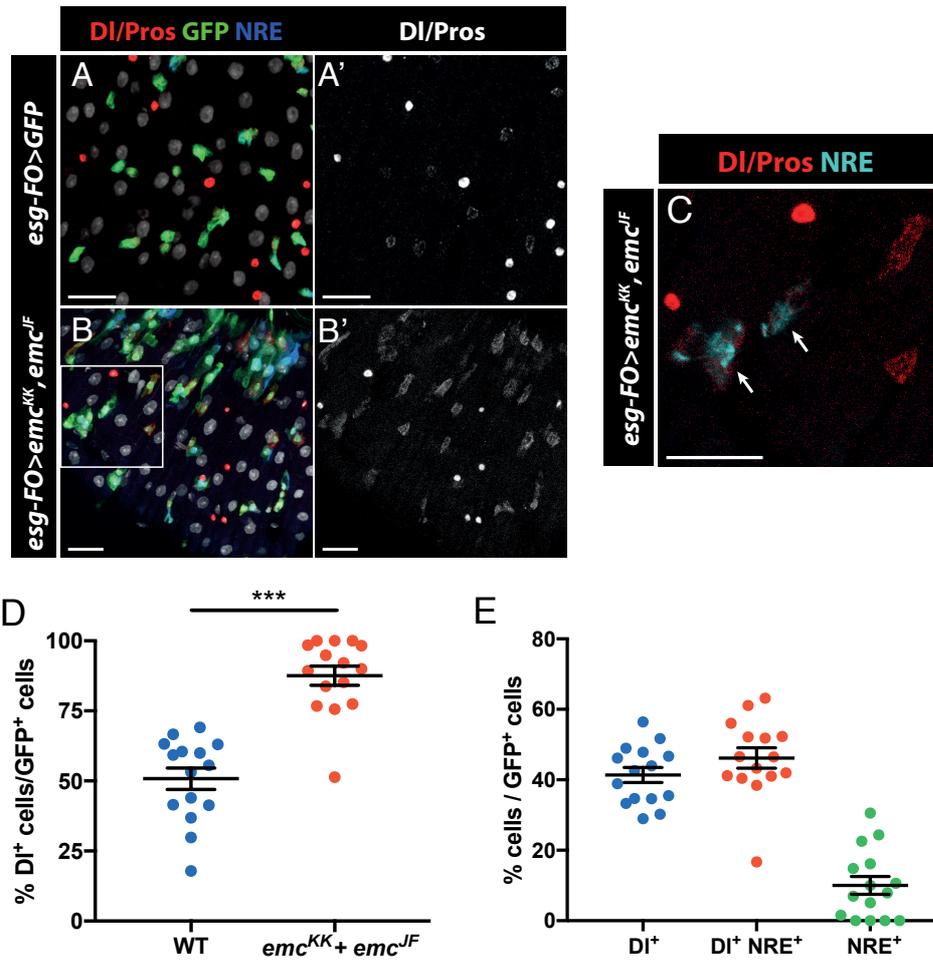
**F-G.** *Dl* signal is increased in *UAS-emc<sup>RNAi-R2</sup>* cells. Prospero staining is reproducible and therefore it is used as a threshold (arrows) for *Dl* intensity (arrow heads) as both were stained with Alexa-594 as a secondary antibody. Control *Dl* signal is lower than Pros, while *UAS-emc<sup>RNAi-R2</sup>* *Dl* is higher.

**H.** WT thorax presents 210-250 microchaetes and 22 macrochaetes (Hartenstein and Posakony, 1989).

**I-K.** Knock down of *emc* with *eyg-Gal4* produces an increase in the number of macrochaetes. It can be observed that the number of macrochaetes using *UAS-emc<sup>R2</sup>* (B) is higher than *UAS-emc<sup>RNAi-KK</sup>* (C) and *UAS-emc<sup>RNAi-JF</sup>* (D).

Data information: **scale bars, 20µm; pictures from panels A-D taken by Joaquín de Navascués**

Since in both approaches we can appreciate that *UAS-emc<sup>RNAi-R2</sup>* is stronger than the other two RNAi lines, we decided to combine *UAS-emc<sup>RNAi-KK108316</sup>* and *UAS-emc<sup>RNAi-JF02300</sup>* to achieve a stronger phenotype. Indeed, the combination of both RNAi lines increased the number of cells expressing *Dl* comparing with the control (**fig. 1.4A-B**, quantification in **fig. 1.4D**). More importantly, more than half of this *Dl<sup>+</sup>* cells were also expressing *GBE-Su(H)-lacZ*, while cells only expressing *GBE-Su(H)-lacZ* represents a 10% of the total GFP population. This result suggests that cells where Notch is activated can still express *Dl* when *Emc* is not present (**fig. 1.4C**, quantification in **fig. 1.4E**). Moreover, we could not observe the disappearance of progenitor cells that we were seeing in *UAS-emc<sup>RNAi-R2</sup>*. Thus, the combination of *UAS-emc<sup>RNAi-KK108316</sup>* and *UAS-emc<sup>RNAi-JF02300</sup>* produced a weaker phenotype than *UAS-emc<sup>RNAi-R2</sup>* but is sufficient to induce the expression of *Dl* in EBs.



**Figure 1.4. Combination of *UAS-emc<sup>RNAi-KK</sup>* and *UAS-emc<sup>RNAi-JF</sup>* induce DI expression**

**A-B.** (A) Comparing control flies and (B) *UAS-emc<sup>RNAi-KK</sup>* construct and *UAS-emc<sup>RNAi-JF</sup>* co-expressed with *esg-FO* shows an increased number of cells expressing DI in the latest.

**C.** Detail from panel B. Co-staining with DI and NRE reveals co-expression of both markers in some cells (arrows), indicating that in this cells Notch is activated while they are expressing DI.

**D.** Quantification of DI<sup>+</sup> cells. Significant increase of DI cells in *UAS-emc<sup>RNAi-KK</sup>, UAS-emc<sup>RNAi-JF</sup>* guts (WT N= 15, *UAS-emc<sup>RNAi-KK</sup>, UAS-emc<sup>RNAi-JF</sup>* N=15) (\*\*\*) (\*\*\*p<0.001, Mann-Whitney test).

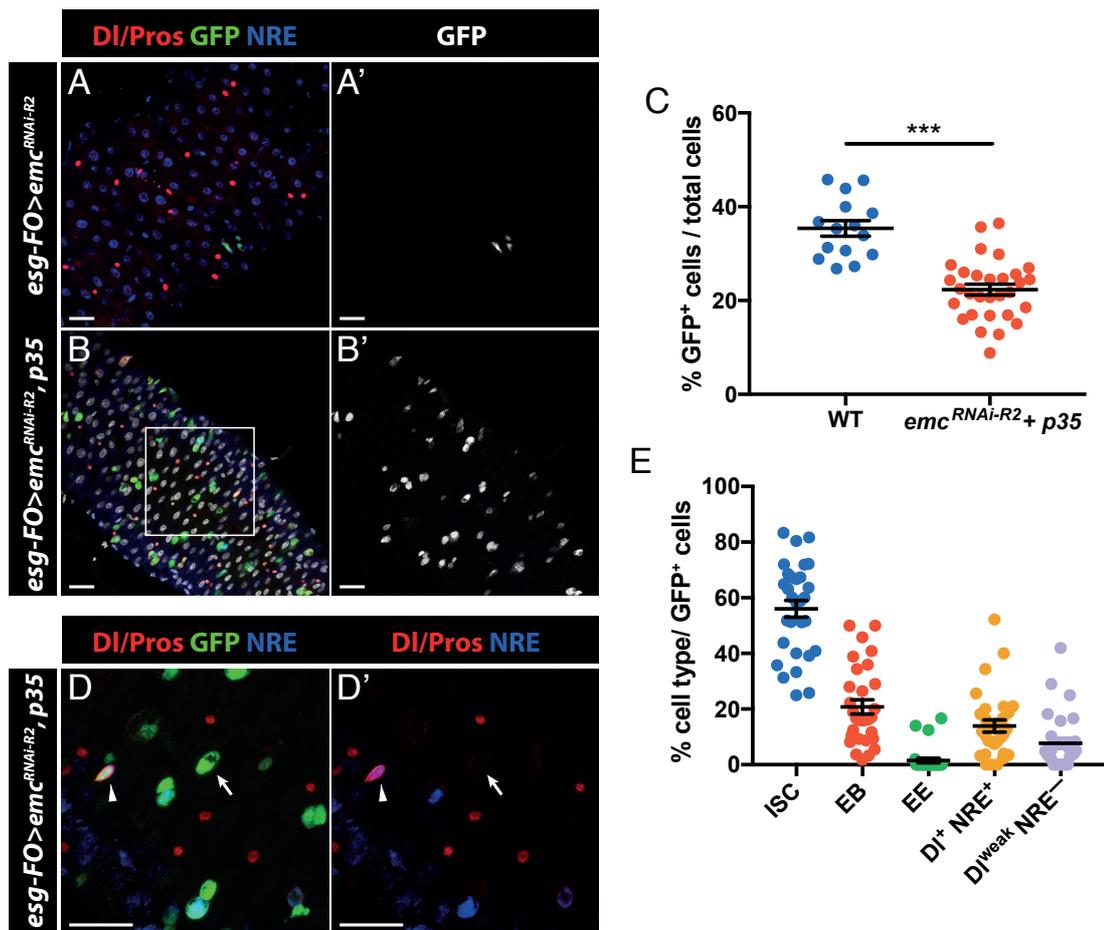
**E.** Quantification of cells expressing DI alone, cells co-expressing DI and NRE and cells expressing NRE only in *esg-FO>UAS-emc<sup>RNAi-KK</sup>, UAS-emc<sup>RNAi-JF</sup>* midguts. DI<sup>+</sup> NRE<sup>+</sup> cells represents a 46.2% of the total population (*UAS-emc<sup>RNAi-KK</sup>, UAS-emc<sup>RNAi-JF</sup>* N=15).

Data information: **scale bars, 20µm**

### 1.A.3.3. Loss of *emc* induces apoptosis

To test if apoptosis was responsible for the disappearance of progenitor cells, we co-expressed *UAS-emc<sup>RNAi</sup><sub>R2</sub>* with *UAS-p35* in all progenitor cells with the *esg-FO* driver. P35 is a caspase inhibitor encoded by the baculovirus *Autographa californica* (Crook *et al.*, 1993), and if cells were dying because of the apoptosis, *p35* expression would inhibit the cell death. The co-expression of *UAS-emc<sup>RNAi</sup><sub>R2</sub>* with *UAS-p35* partly rescued the cell death (**fig. 1.5A,B**), as now there was an evenly distributed progenitor population, although less GFP<sup>+</sup> cells were detected compared with wild-type (**fig. 1.5C**). Hence, the knock down of *emc* in progenitor cells unchains a mechanism that will derive into apoptosis through all the posterior midgut and the cell death can be rescued with P35.

In the guts where we co-expressed *UAS-p35* and *UAS-emc<sup>RNAi</sup><sub>R2</sub>* we also observed that the majority of progenitor cells were DI<sup>+</sup> with increased levels. However, we could also observe a subset of DI<sup>+</sup> cells that were also expressing NRE (**fig. 1.5D**, quantification in **1.5E**). Finally, there were some diploid cells with very weak DI expression. Both of these populations are not present when *p35* is not expressed in *emc* knocked down, and therefore these two populations are dying.



#### 1.A.3.4. Emc is necessary to inhibit dedifferentiation of EBs

We have shown in figure 1.1 that the expression of *emc* in ISC is variable from not expressed to weakly expressed, while in EBs is consistently expressed. As *Dl* is expressed in ISCs but not in EBs, there seems to be a correlation between the expression of *Dl* and the expression of *emc* and it is possible that Emc is blocking the expression of *Dl* in EBs. To test this hypothesis, we used an EB specific driver, *Su(H)-GBE-Gal4*, *UAS-GFP*, *tub-Gal80<sup>ts</sup>* (*Su(H)-GBE-Gal4<sup>TS</sup>*) to knock down *emc*. Whereas in the control guts we normally find single EBs, the guts where *emc* was knocked down in the EBs presented clusters of GFP<sup>+</sup> cells with more than 3 cells (**fig. 1.6A-B**, quantification in **1.6C**). Strikingly, some of these GFP<sup>+</sup> cells were also expressing *Dl*, whereas in control flies we have never observed EBs that were *Dl*<sup>+</sup> (**fig. 1.6D-E**). Moreover, the GFP intensity was not homogenous, and the lower GFP intensity cells presented higher levels of *Dl*. As this was not a FO system, cells with lower levels of GFP are probably cells that are no longer expressing GFP nor *UAS-emc<sup>RNAi</sup><sub>R2</sub>*. These cells that no longer expressed GFP had similar levels of *Dl* than the *emc* knock down mediated by *esg-FO*. Therefore, we surmised that loss of *emc* in EBs results in a loss of committed features and gain of stem characteristics, meaning that these cells could be dedifferentiating into ISCs.

To further test if loss of *emc* in EBs leads to dedifferentiation and acquisition of stem capacities, such as division ability or multipotency, we generated a *Su(H)-GBE-FO*, which contains the *FO* system using the *Su(H)-GBE-Gal4<sup>TS</sup>* instead of *esg-Gal4*. Therefore, we kept the flies at non-permissive temperature for 7 days and we stained with the mitosis marker anti-phospho Histone 3 (PH3) and the EE marker Pros. Whereas in control flies EBs only can be maintained or differentiate into ECs and no longer divide, *emc<sup>RNAi</sup>* EBs generated *Dl*<sup>+</sup> cells that were positive for PH3, and therefore were dividing (**fig. 1.6F**). Moreover, we could observe EE cells labelled with GFP. Thus, loss of *emc* in EBs results in dedifferentiation to an ISC like state. These ISCs can divide and differentiate into EEs.

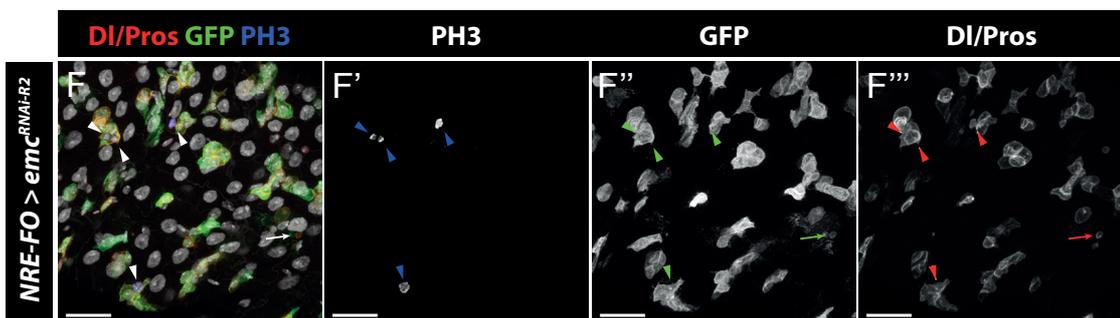
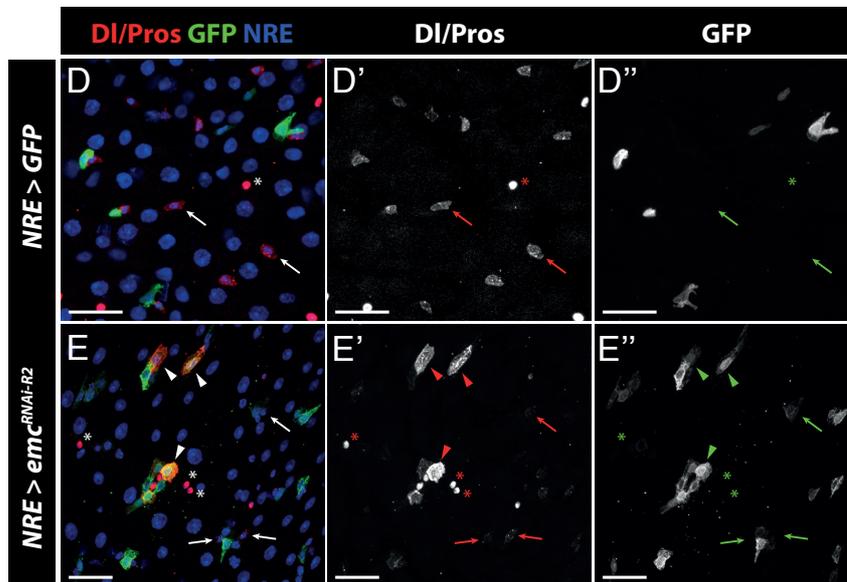
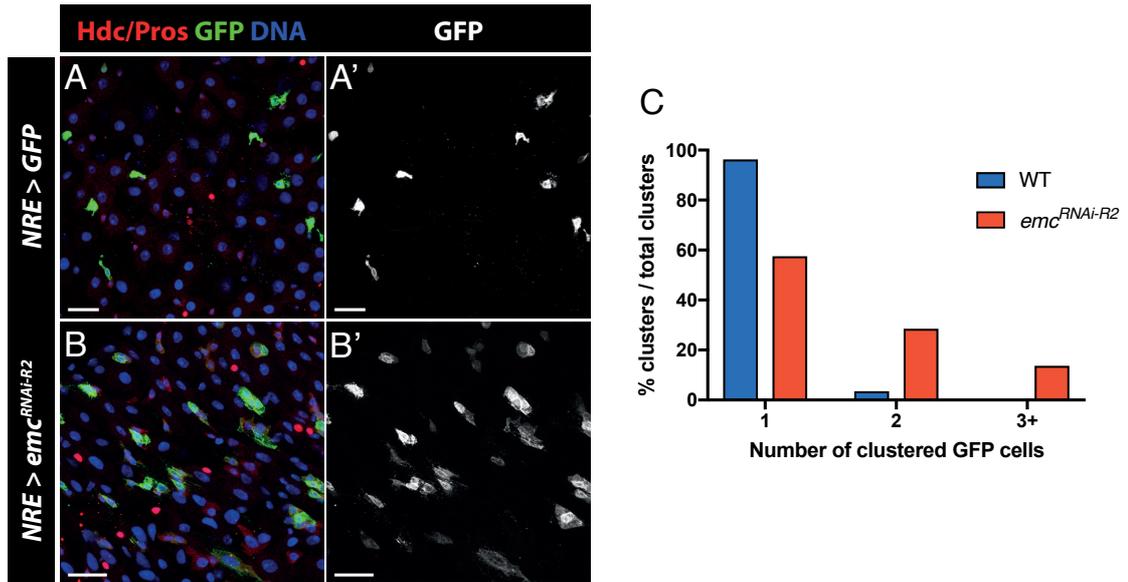


Figure 1.6. Knock down of *emc* in EBs induce de-differentiation to ISC-like cells (see next page for legend)

### Figure 1.6. Knock down of *emc* in EBs induce de-differentiation to ISC-like cells

**A.** Control *NRE>GFP* posterior midguts showed an even distribution through the tissue of single  $Hdc^+$   $GFP^+$  cells, normally next to  $Hdc^+$   $GFP^-$  cells

**B.** Midguts of *NRE>emc<sup>RNAi-R2</sup>* flies. It is observable an accumulation of  $GFP^+$  cells, with a reduction of the number of isolated  $GFP^+$  cells respect the control

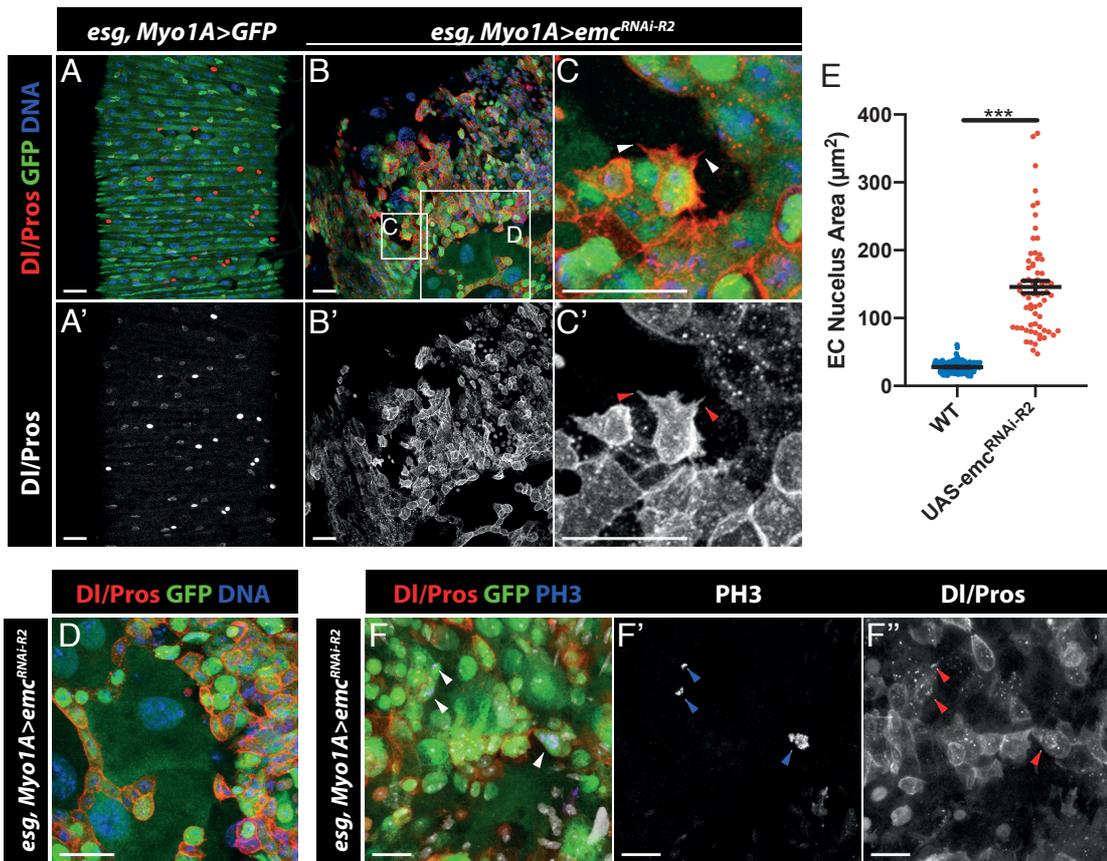
**C.** Quantification of clustered  $GFP^+$  cells. In control guts, the majority of  $GFP^+$  cells do not cluster (96.41%). Knock down of *emc* in EBs produces more clusters of two (28,62%) or more than two (13,79%)  $GFP^+$  cells (WT (blue) number of clusters=279, *UAS-emc<sup>RNAi-R2</sup>* (red) number of clusters=290).

**D-E.** In control flies (D), all  $DI^+$  cells are  $GFP^-$  (arrows, D'-D'') And the intensity of  $DI$  is lower than the intensity of  $Pros$  (asterisk). *NRE>emc<sup>RNAi-R2</sup>* midguts (E) show  $DI^+$   $GFP^+$  cells (E'-E''), with  $DI$  intensity higher than in  $DI^+$   $GFP^-$  cells (arrows) and comparable with  $Pros$  (asterisk).

**F.** Lineage tracing from  $NRE^+$  cells shows  $DI^+$   $GFP^+$  cells that are dividing ( $PH3^+$ , arrow heads). Some  $Pros^+$  cells are  $GFP^+$  (arrows).

Data information: **scale bars: 20 $\mu$ m**

Interestingly, when we knock-down the expression of *emc* only in EBs, we did not observe any sign of cell death as when we express *UAS-emc<sup>RNAi-R2</sup>* in all progenitor cells. In principle, this could suggest that the apoptosis observed in those conditions is non cell-autonomous, i.e. it is the interaction between cells lacking *Emc* that induce apoptosis. Therefore, surviving clones in *esg-FO* experiments could come from ISCs/dedifferentiated EBs that were isolated. As both ISC and EB represent a 18% of the cell population in the midgut (together 36% (Micchelli and Perrimon, 2006)), we wanted to study this hypothetical non-autonomous effect in the whole tissue. Consequently, we combined the *esg-Gal4* with the EC-specific driver *Myo1A-Gal4* to express *UAS-emc<sup>RNAi-R2</sup>* (*esg+Myo1A-Gal4<sup>TS</sup>*). When we performed an immunohistochemistry assay and stained with  $DI$  and  $Pros$ , we could observe that guts were covered with  $DI^+$  cells and there were few ECs, compared with control flies (**fig. 1.7A-B**). Interestingly, some  $DI^+$  cells presented protrusions which are not characteristic of ISC (**fig. 1.7C**) and the cytoplasm and nucleus of ECs were highly enlarged (**fig. 1.7D**, quantification in **1.7E**). Moreover, we wanted to know if the  $DI^+$  cells were proliferative, and we stained for  $PH3$ . This staining showed that these  $DI^+$  cells were proliferative (**fig. 1.7F**). Together, these results show that knocking down the expression of *emc* in the whole tissue induce an ISC over-proliferation and EC loss. The remaining ECs increase in volume.



**Figure 1.7. Knock down of *emc* in a large proportion of tissue induce ISC over-proliferation**

- A.** *esg, Myo1A>GFP* in posterior midguts induce expression in all cells except Pros<sup>+</sup> cells
- B.** *esg, Myo1A>emc<sup>RNAi-R2</sup>* midguts shows large clusters of DI<sup>+</sup> cells that cover the majority of the tissue
- C.** Detail of panel B. It can be observed that DI<sup>+</sup> cells present protrusions (arrow heads) when they are not in contact with other cells.
- D.** Detail of panel B. Surviving ECs present large nuclei and cytoplasm due to several endoreplication cycles.
- E.** Nucleus area is larger in *esg, Myo1A>emc<sup>RNAi-R2</sup>* ECs compared with control ECs nucleus (WT N= 213 ECs, *emc<sup>RNAi-R2</sup>* N=66 ECs)(\*\*\*p<0.001, Mann-Whitney test).
- F.** Knock down of *emc* promotes the proliferation in the tissue of DI<sup>+</sup> cells.

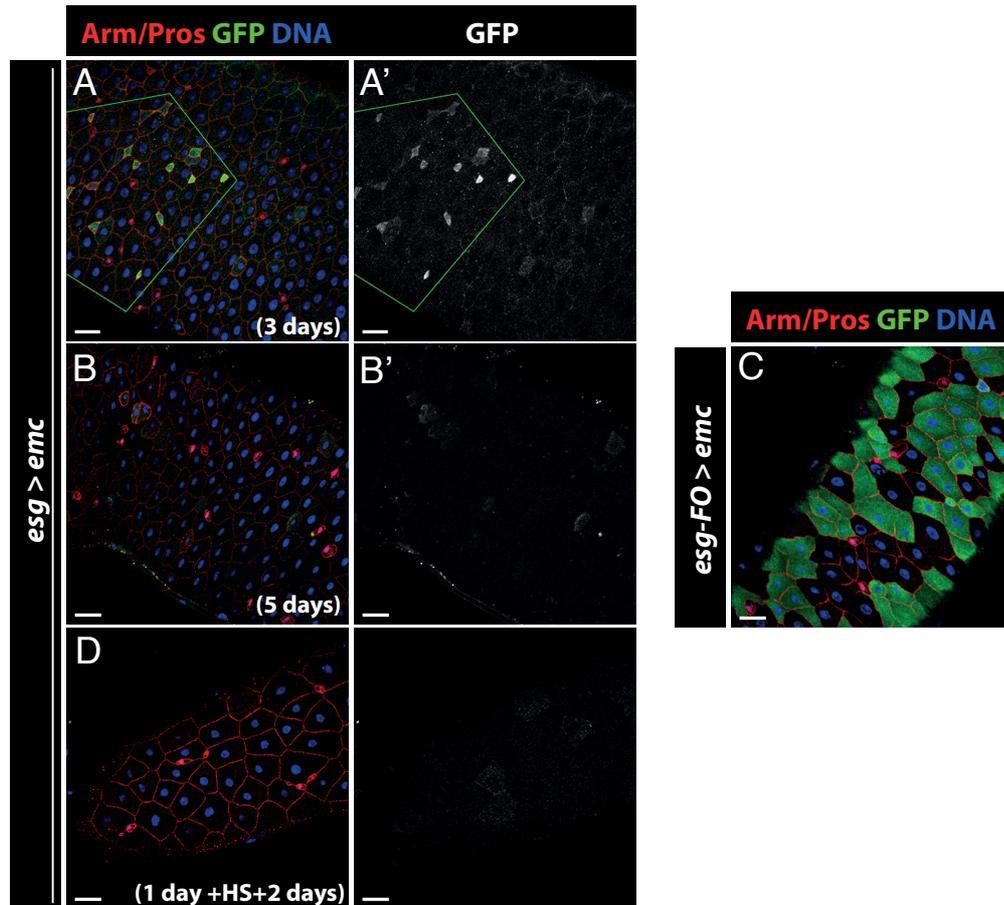
Data information: **scale bars: 20μm**

#### 1.A.3.5. Expression of *emc* selects for absorptive fate

Our previous findings led us to hypothesize that Emc is essential for terminal differentiation in progenitor cells. Therefore, it is plausible that if progenitor cells start expressing high levels of *emc*, they would differentiate. To confirm this, we used the *esg-Gal4<sup>ts</sup>* driver to express *UAS-emc*. At 3 days after expression initiation we observed large areas without progenitor cells (**fig. 1.8A**). Still, other areas showed no difference in progenitor cell content (GFP cells) compared with the controls. After 5 days of overexpression there were not any progenitor cell in the midguts, and the tissue was composed mainly by ECs and some EE cells (**fig. 1.8B**). However, some polyploid cells had a faint GFP expression. This could indicate that *UAS-emc* could be forcing the differentiation of all progenitor cells, although it could be also possible that it is promoting apoptosis and the few survivors could differentiate. To distinguish between these possibilities, we used the *esg-FO* system for lineage tracing to overexpress *UAS-emc* in progenitor cells and their descendant progeny for 7 days (**fig. 1.8C**). The guts still contained EE cells (Pros<sup>+</sup>) and ECs (polyploid), however, now we could observe that a big proportion of these ECs were GFP<sup>+</sup>, meaning that the progenitor pool was lost due to differentiation instead of apoptosis. It was also notable that none of the EE cells were GFP<sup>+</sup>, indicating that *emc* overexpression induces only absorptive differentiation.

#### 1.A.3.6. *emc* acts downstream of Notch activation

These results highlight the importance of Emc to promote cell differentiation and ensuring that EBs do not dedifferentiate. Interestingly, we could observe that after 3 days of overexpression, the differentiation was only localized in some parts of the tissue. This suggests that the responsiveness to *UAS-emc* is not synchronous. This asynchronous kinetics of differentiation of progenitor cells has also been observed as well when NICD is overexpressed (de Navascués, unpublished data); NICD drives the terminal differentiation of ISCs and EBs into ECs (Ohlstein and Spradling, 2007). Moreover, challenging the flies with heat stress can accelerate NICD kinetics of differentiation (de Navascués, unpublished data), possibly through promoting tissue turnover. To confirm that heat stress can accelerate progenitor differentiation I heat-shocked *UAS-emc, esg<sup>TS</sup>* flies at 37 °C for 1 hour and then let them recover at 29 °C for 3 days (**fig. 1.8D**).



**Figure 1.8. *emc* expression results in terminal absorptive differentiation**

**A-B.** Expression of *UAS-emc* using the *esg-Gal4* for 3 (A) and 5 days (B). It can be observed that GFP cells disappear from the tissue. The disappearance of GFP<sup>+</sup> cells does not progress evenly in the whole tissue, as we can see large surfaces with GFP<sup>-</sup> cells next to areas with GFP<sup>+</sup> cells (limit delineated in green in A'). After 5 days, all progenitor cells have differentiated (B').

**C.** Lineage tracing for the expression of *UAS-emc* in progenitor cells for 7 days shows that all cells have differentiated into ECs, being the absorptive differentiated selected.

**D.** Expression of *UAS-emc* using the *esg-Gal4* for 3 days, with a heat shock (one hour, 37°C) after day one. All progenitor cells have differentiated

Data information: **scale bars, 20µm**

Remarkably, all cells in the guts were differentiated, suggesting that the delay depends on the dynamics of the tissue rather than in the expression of *UAS-emc*.

The similarity between NICD and *emc* overexpression phenotypes, suggested that *emc* could be a target downstream of Notch activation. This hypothesis is supported by previous studies which report that Notch drives *emc* expression during wing vein

formation (Baonza *et al.*, 2000) and eye disc development (Spratford and Kumar, 2015b). Therefore, we wanted to analyze the expression of *emc* in NICD overexpression clones, and to that end we used the *emc*<sup>CPT12740</sup> line in combination with *Actin5C-FRT-γ<sup>+</sup>-FRT-Gal4*, *UAS-lacZ* (a.k.a. *Ay-Gal4*) and *hs-flippase*. This system allowed us express *UAS-NICD* in the clones (traced with the *UAS-lacZ*). In contrast with the MARCM technique, where the clones are formed by mitotic recombination and therefore in dividing ISC, *Ay-Gal4* recombination can occur in any cell type. The clones were induced at 37 °C for 15 minutes, and the flies were kept for 3 days at 25 °C prior to dissection to ensure the expression of Notch downstream targets (**fig. 1.9A**). We could not appreciate any difference with the GFP intensity inside and outside the clones, considering the differential expression in each cell type. Thus, it seemed that there was not an increased expression of *emc*. However, the expression of *emc* in WT midguts is not constant within each cell type.

As the expression of *emc* downstream of Notch activity was not clear, we decided to further investigate with an epistatic analysis. We used *esg-FO* to knock down the expression of *Notch* (*UAS-N<sup>RNAi</sup>*) and overexpress *UAS-emc* at the same time. While the *UAS-N<sup>RNAi</sup>* typical phenotype results in an over-proliferation of stem cells, formation of clusters of ISC and EE cells and blocking the absorptive differentiation (**fig. 1.9B**), with *UAS-emc* co-expression cells differentiate into EC, similar as when we express *UAS-emc* alone (**fig. 1.9C**). This result suggests that *emc* function could be downstream of Notch. However, it also could be that the knock down process is too slow, and when it starts to have an effect, *emc* has already forced the differentiation into ECs irreversibly. To avoid this potential problem with the expression dynamics, we decided to generate a Notch loss of function phenotype by overexpressing *H*. The overexpression of *UAS-H* has been shown to mimic the *UAS-N<sup>RNAi</sup>* phenotype in the *Drosophila* gut (Bardin *et al.*, 2010). Furthermore, the absence of H produce the loss of stemness, similar to the expression of *UAS-NICD* (Bardin *et al.*, 2010), although while the expression of *UAS-NICD* drives differentiation in the vast majority of the tissue after 7 days, the knockdown of *H* takes more time to make the transition from EB to EC (Bardin *et al.*, 2010). Therefore, we proceeded with the epistatic analysis and the co-expression of *UAS-H* and *UAS-emc*. Control flies expressing only *UAS-H* could form clusters of ISC, with few and small EE clusters. However, when we co-express *UAS-H* and *UAS-emc*, we arrested completely the

secretory differentiation, ISCs did not proliferate and we increased the absorptive differentiation, with a 36% of ECs compared with the 1% in *UAS-H* condition (fig. 1.9D-E, quantification in 1.9H). Complementary, the double knock-down with *UAS-H<sup>RNAi</sup><sub>HMS01182</sub>* and *UAS-emc<sup>RNAi</sup><sub>R2</sub>* showed 32% of DI<sup>+</sup> cells after 14 days, with high DI levels, whereas in *UAS-H<sup>RNAi</sup><sub>HMS01182</sub>* flies all ISC had differentiated (fig. 1.9F-G, quantification in 1.9I). Noticeable, guts with the double knock-down had an important loss of progenitor cells (fig. 1.9H<sup>''</sup>). Taken together, these results suggest that Emc activity is epistatic over Notch and Emc is necessary for the correct differentiation towards ECs downstream of N activity.

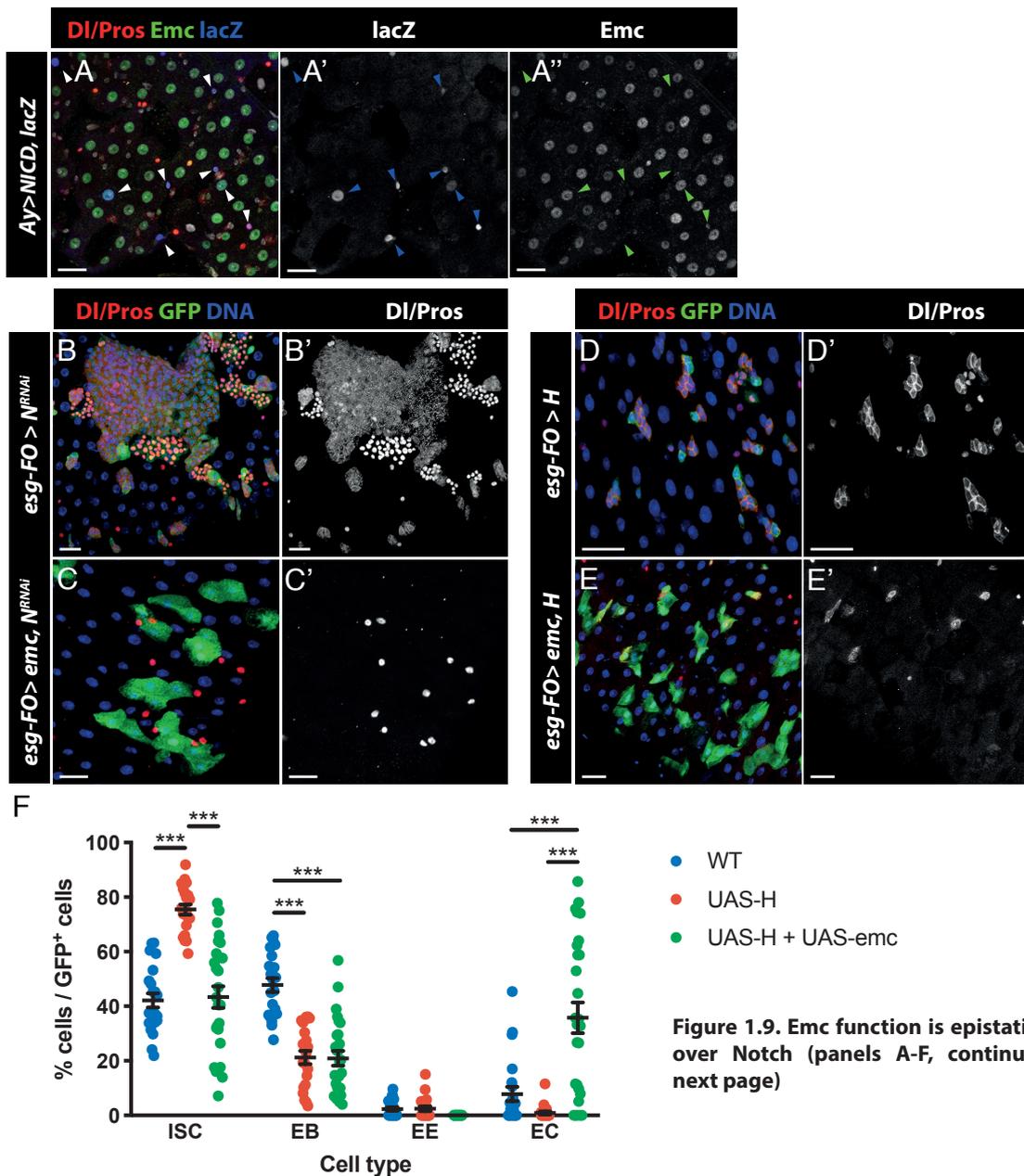
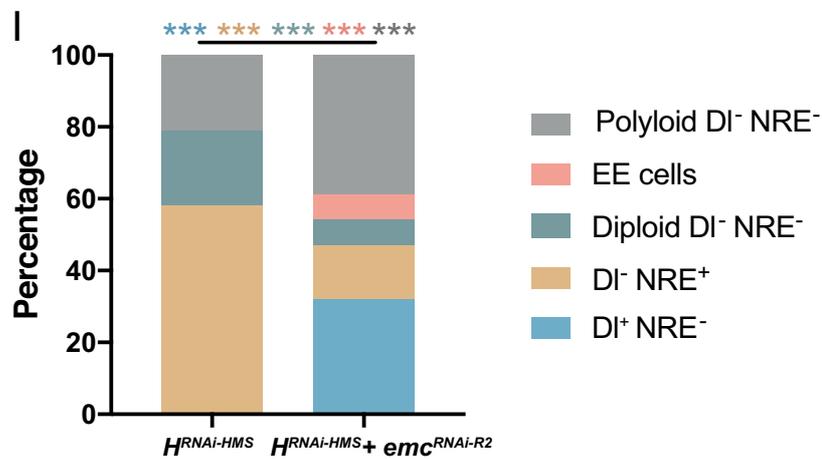
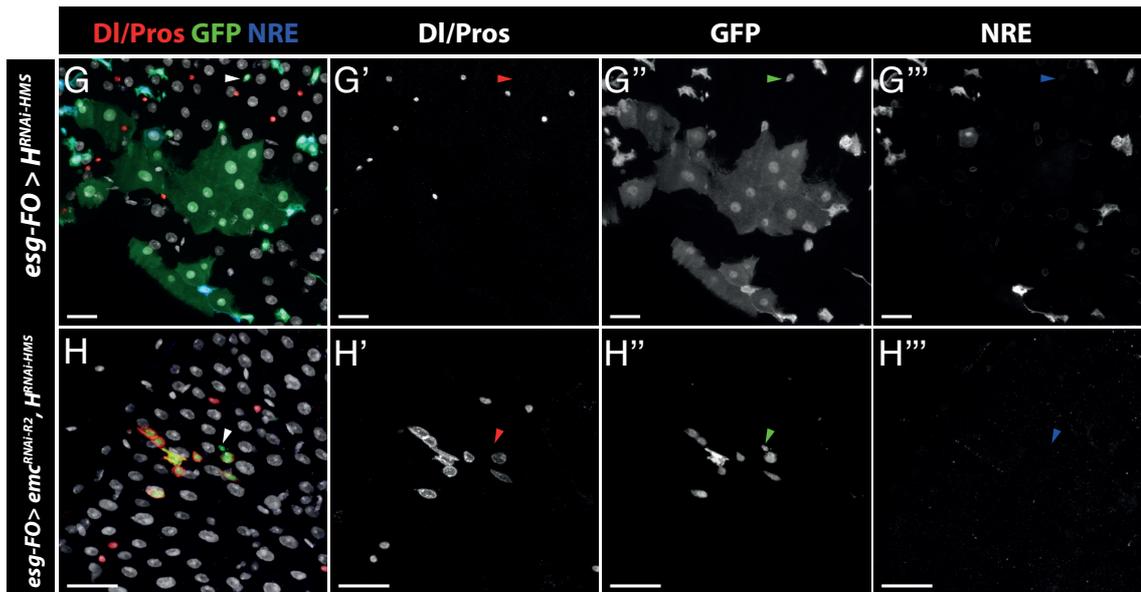


Figure 1.9. Emc function is epistatic over Notch (panels A-F, continue next page)



**Figure 1.9. Emc function is epistatic with Notch (panels A-G, continue next page)**

- A.** Overexpression of *UAS-NICD* in clones (clones highlighted with arrow heads) shows different *emc* expression levels, with some cells not expressing *emc*.
- B.** Knock down of *Notch* in progenitor cells produce over-proliferation of ISCs and formation of large ISC clusters and EE cell cluster, inhibiting absorptive differentiation.
- C.** Expression of *UAS-emc* in *Notch* knocked down guts promotes terminal differentiation.
- D.** Expression of *UAS-H* inhibits Notch transcriptional activity, mimicking *UAS-N<sup>RNAi</sup>* phenotype. *DI<sup>+</sup>* cells form proliferating clusters and without absorptive differentiation.
- E.** Progenitor cells expressing *UAS-emc* and *UAS-H* are able to differentiate into ECs, with some remaining *DI<sup>-</sup>* cells.
- F.** The percentages of the different cell populations in *UAS-H* and *UAS-H, UAS-emc* expression in progenitor cells. (*UAS-H* N=21; *UAS-H, UAS-emc* N=27) (\*\*p<0.01, \*\*\*p<0.001, Multiple T test).
- G.** Knock down of *H* for 14 days slowly promotes absorptive differentiation. All cells are *DI<sup>-</sup>*. Moreover, some cells are *DI<sup>-</sup> NRE<sup>-</sup>* and contain a small nuclei (arrow head).
- H.** Double knock down of *emc* and *H* in progenitor cells present *DI<sup>+</sup>* cells and also the unidentified *DI<sup>-</sup> NRE<sup>-</sup>* population (arrow head).
- I.** Quantification of cell population in over-activation of Notch with and without *emc* expression. Knock down of *emc* restores *DI<sup>+</sup>* and *Pros<sup>+</sup>* cell populations (*UAS-H<sup>HMS</sup>* N=15 guts, 696 cells; *UAS-H<sup>RNAi-HMS</sup>, UAS-emc<sup>RNAi-R2</sup>* N=20 guts, 90 cells) (\*\*p<0.001, Binomial regression. Each cell type significance level is coloured accordingly).

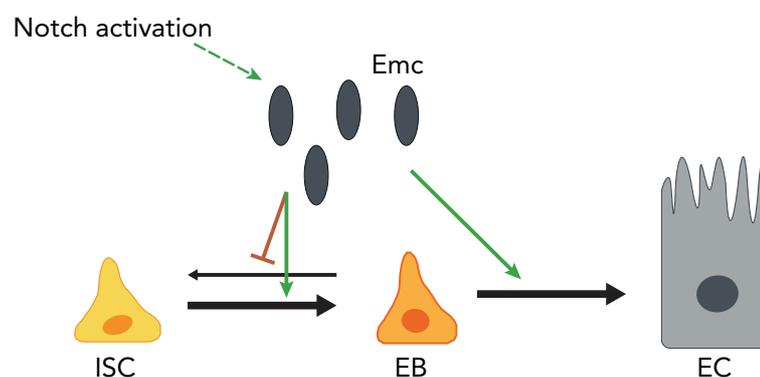
Data information: scale bars, 20µm

#### 1.A.4. Conclusions

We have shown that *emc* is expressed in the *Drosophila* posterior midgut in all cell types. However, the expression levels vary among and within cell types, having the highest levels in ECs, and the most constant expression in EBs (**fig 1.1**). Moreover, Emc is necessary for terminal differentiation (**fig 1.2** and **fig 1.3**) and a high expression of *emc* induces absorptive differentiation (**fig 1.8**). We have also observed that the differentiation function is downstream of Notch signalling, which activity also induces absorptive differentiation (Micchelli and Perrimon, 2006). However, *emc* expression is not controlled by Notch (**fig 1.9**).

Aside from inducing absorptive differentiation, we have discovered that Emc is also necessary to maintain EBs committed, and loss of *emc* in these cells induce de-differentiation into ISCs and expression of *Dl* (**fig 1.6**)( **fig 1.10**).

Finally, loss of *emc* in progenitor cells induces apoptosis that can be partly rescued with the expression of *p35* (**fig 1.5**). However, when *emc* is knocked down in progenitor cells and ECs at the same time, ISCs over-proliferate and ECs increase their volume (**fig 1.7**).



**Figure 1.10. Emc promotes absorptive differentiation and blocks de-differentiation**

Emc function is downstream of Notch, although *emc* expression levels do not vary when Notch is activated. High levels of Emc induce terminal absorptive differentiation. Moreover, loss of *emc* induces EB to de-differentiate, indicating that Emc is crucial to halt de-differentiation.

## Chapter 1B: A network of bHLH factors controls self-renewal and bipotential differentiation in the intestine: Da and Sc

### 1.B.1. Introduction

Little is known about how Da controls the homeostasis of the tissue, as it has only been observed that loss of *da* results in absorptive differentiation (Bardin *et al.*, 2010). To drive transcriptional regulation, Da can form protein complexes of different composition:

1. Da/Class II dimers: Da typically dimerize with bHLH class II for its transcriptional function. The most studied partnership is the one formed by Da and members of the AS-C for the formation of bristles in the notum (Culí and Modolell, 1998; Murre *et al.*, 1989a; Troost *et al.*, 2015). In the midgut, *sc* and *ase* have been shown to be important for the secretory fate (Bardin *et al.*, 2010; Li *et al.*, 2017a). *Sc* is also necessary in ISC, although its function is not fully elucidated and its expression is tightly regulated (Chen *et al.*, 2018; Wang *et al.*, 2015).
2. Da/Zn finger complexes: Da is able to physically interact with the C2H2 Zn finger transcription factor *senseless* (Jafar-Nejad *et al.*, 2006). Hence, Da could maintain the progenitor state by binding to other Zn finger proteins from the C2H2 family, such as *Chn*.
3. Da/Da homodimers: bHLH class I members can also homodimerize, although the dimer binding is weaker than the heterodimers (Cabrera and Alonso, 1991; Murre *et al.*, 1989a). This homodimeric form can bind DNA and drive transcription, as it has been shown that it induces the expression of *ato* in the eye disc (Tanaka-Matakatsu *et al.*, 2014) and *expanded (ex)* in the eye and wing disc (Wang and Baker, 2015). Therefore, Da could also form homodimers in the midgut to prevent terminal differentiation.

It has also been shown that the FoxA factor *Fkh* and Da share gene targets to maintain the progenitor state (Lan *et al.*, 2018). However, Lan *et al.* propose that *Fkh* would serve as an initiator of the transcription by opening the chromatin and facilitate the accessibility of other transcription factors. Therefore, it is likely that Da and *Fkh* do not dimerise.

Therefore, in this chapter we will explore the role of Da controlling the homeostasis and with which partners does it dimerise to exert its different roles.

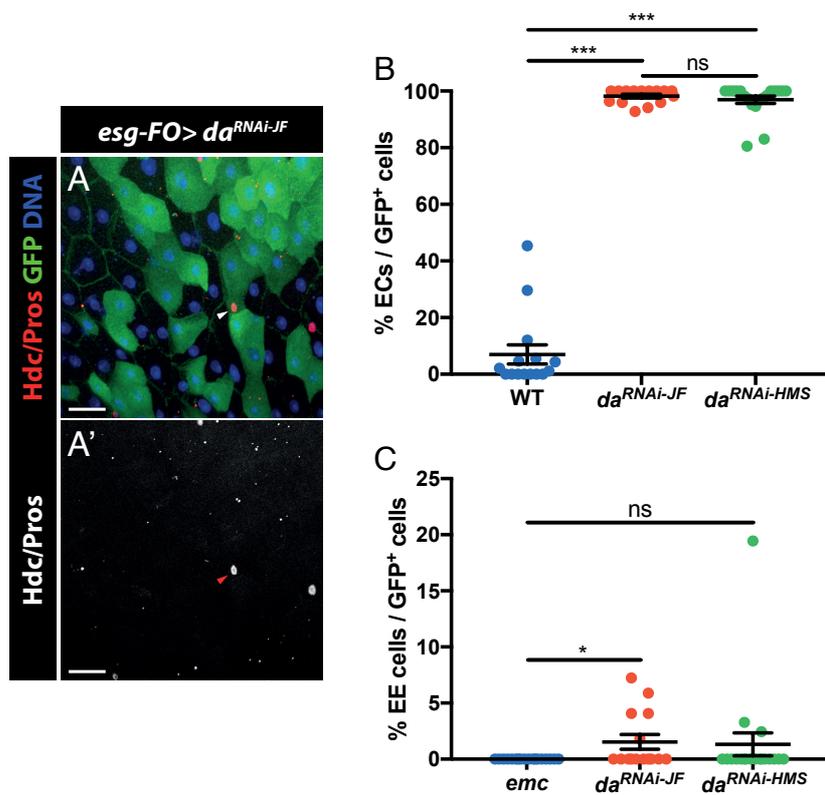
### 1.B.2. Aims

- Determine the function of Da during adult intestinal homeostasis
- Study the functional relationship between Emc and Da in the gut
- 
- Explore the relationship between Sc and Emc during secretory differentiation

### 1.B.3. Results

#### 1.B.3.1. Emc opposes Da function

In 2010, Bardin *et al.* showed that when they induced  $da^{10}$  clones in adult midguts, the stemness was lost, the clones did not grow and cells became ECs. Although we have not induced overexpression clones of *UAS-emc*, the *da* loss of function phenotype is comparable with the overexpression of *UAS-emc*. It is highly possible that high levels of Emc binds with all Da and therefore inhibits its function. We then wanted to know the effect of knocking down the expression of *da* in all progenitor cells using the *esg-FO* technique and two different RNAi: *UAS-da<sup>RNAi</sup><sub>JF02488</sub>* (**fig. 1.11A**, quantification in **1.11B**) and *UAS-da<sup>RNAi</sup><sub>HMS01851</sub>* (data not shown, quantification in **1.11B**). We obtained the same phenotype with both *da* RNAi lines, and is similar to *UAS-emc* overexpression phenotype, although some cells were differentiating into EE cells (**fig. 1.11C**). We also induced formation of  $da^{10}$  MARCM clones, and observed that all *da* mutant cells differentiated, as previously described (Bardin *et al.*, 2010). The majority of cells also differentiated into ECs, and a small proportion of them into EE cells (data not shown). This confirms that Da is required to maintain the stem cells. Considering the reduced level of EE differentiation associated with *da* loss, it is also possible that Da is needed for secretory differentiation.

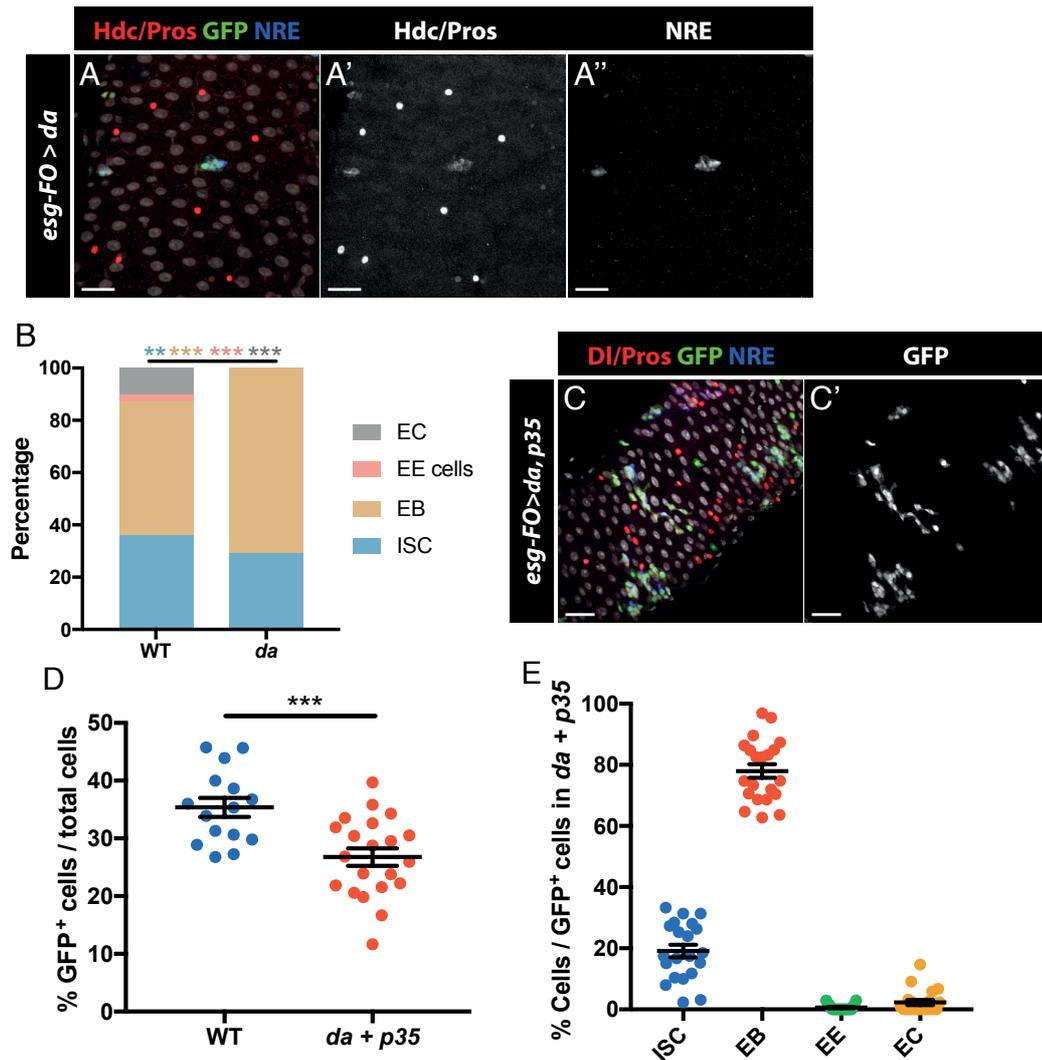


**Figure 1.11. Loss of *da* leads to loss of stemness**

- A.** Knock down of *da* with the *esg-FO* system induce terminal differentiation and loss of stemness (arrow head indicates Pros<sup>+</sup> GFP<sup>+</sup> cells).
- B.** Knock down of *da* promotes mainly absorptive differentiation (WT N = 15, *UAS-da<sup>RNAi-JF</sup>* N = 15, *UAS-da<sup>RNAi-HMS</sup>* N = 19) (\*\*\*)  $p < 0.001$ , Kruskal-Wallis/Dunn test).
- C.** Cells without *da* still can differentiate into EE cells in contrast with over-expression of *UAS-emc* (*UAS-emc* N = 18, *UAS-da<sup>RNAi-JF</sup>* N = 15, *UAS-da<sup>RNAi-HMS</sup>* N = 19) (\* $p < 0.05$ , \*\*\* $p < 0.001$ , Kruskal-Wallis/Dunn test).

Data information: **scale bars, 20 $\mu$ m**

Next, we were interested in whether *da* is necessary to maintain the stem pool and find which is the dimerization partner. To investigate these two points, we started expressing *UAS-da* using the *esg-FO* system (**fig. 1.12A**, quantification in **1.12B**). We found that there were few surviving clusters in the gut, and those were composed mostly of progenitor cells. Moreover, there was an increased number of cells expressing *GBE-Su(H)-lacZ*. We then wanted to know if the levels of Dl were elevated, and we could observe that surviving clones express few Dl<sup>+</sup> cells, whose expression was not elevated (**fig. 1.12C**). Therefore, *UAS-da* strikingly mimics the phenotype of *UAS-emc<sup>RNAi</sup><sub>R2</sub>* in progenitor cells, although it seems that Da alone is not responsible for the regulation of *Dl*, and Emc must be inhibiting another bHLH to control *Dl* expression. As the overexpression of *UAS-da* also promotes loss of progenitor cells, we wanted to rescue apoptosis. We co-expressed *UAS-da* and *UAS-p35* using *esg-FO* (**fig. 1.12D**, quantification in **1.12E**). These guts presented cells that were either Dl<sup>+</sup> or *GBE-Su(H)-lacZ*<sup>+</sup>, but never both markers together, as we saw when we rescued cell death in *emc* knocked down progenitor cells. Interestingly, the number of Hdc<sup>+</sup> NRE<sup>-</sup> was reduced from a 33% in *UAS-da* only expression to a 19 % with *UAS-p35* co-expression, while the number of Hdc<sup>+</sup> NRE<sup>+</sup> represents a 78% of the GFP<sup>+</sup> population when cell death is rescued. Together, these results show that Emc inhibits a different target than Da to impair the expression of Dl and promote dedifferentiation.



**Figure 1.12. Expression of *da* prevents differentiation**

**A.** *esg-FO > da* expression induce disappearance of the majority of progenitor cells. The surviving clusters contain only progenitor cells (A').

**B.** Quantification of the cell population in WT and *da* over-expressing midguts shows a complete arrest of differentiation (WT N = 987 cells, *UAS-da* N = 270 cells) (\*\* $p < 0.01$ , \*\*\* $p < 0.001$ , Binomial regression. Each cell type significance level is coloured accordingly).

**C.** Co-expression of *UAS-da* and *UAS-p35* in progenitor cells rescue the disappearance of the majority of progenitor cells.

**D.** Co-expression of *UAS-da* and *UAS-p35* does not rescue completely the cell loss (WT N = 15, *UAS-da* N = 21) (\*\*\* $p < 0.001$ , Unpaired T test).

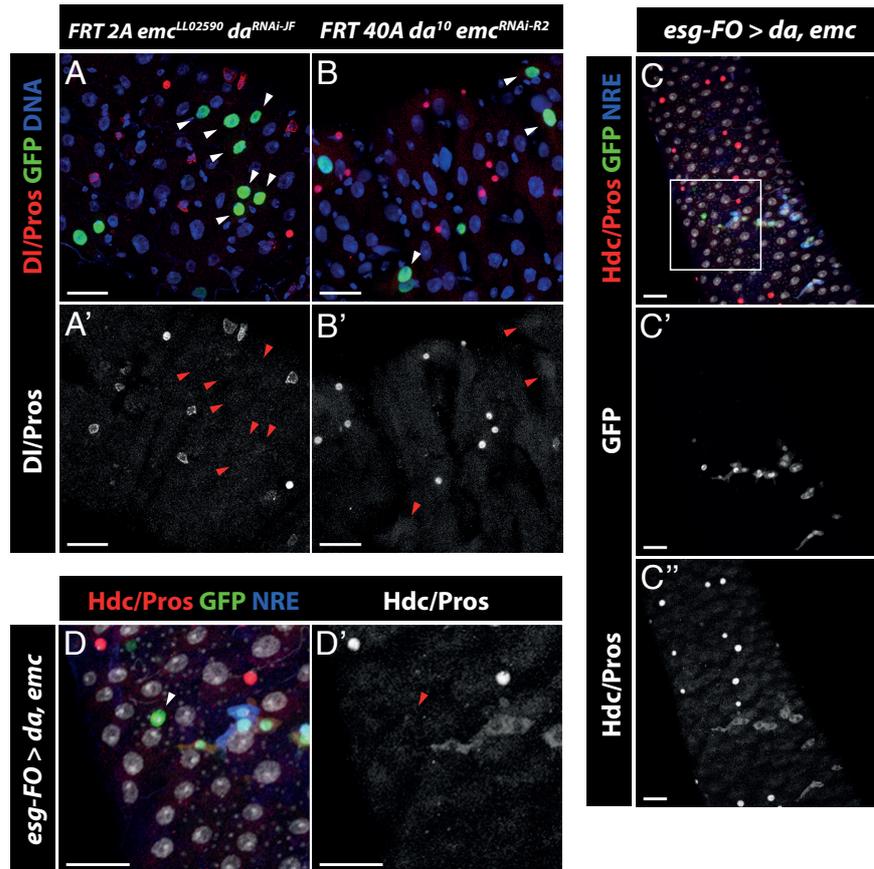
**E.** Population distribution in *esg-FO > da, p35* showing that when cell death is impaired, only few cells differentiate, and the majority remain as EBs (*UAS-da* N = 21).

Data information: **scale bars, 20 $\mu$ m**

Our previous results showed that Da and Emc have opposite effects and most likely this is because the main function of Emc is sequestering Da (and other bHLH factors) and inhibiting its transcriptional activity. This model predicts that in the absence of Da, Emc would be dispensable. For this experiment we wanted to avoid the widespread cell death, and therefore we induced MARCM clones of *emc<sup>LL</sup>* and *UAS-da<sup>RNAi</sup><sub>JF02488</sub>* (**fig. 1.13A**). As expected, all cells differentiated into ECs. To further confirm this effect, we induced *da<sup>10</sup>*, *UAS-emc<sup>RNAi</sup><sub>R2</sub>* MARCM clones (**fig. 1.13B**). Both results showed exactly the same phenotype, all cells differentiated into ECs, and there was not any progenitor cell left in the clones. Thus, when Da is not present, Emc function is not relevant.

Conversely, in excess of Da and Emc, we would expect an intermediate phenotype where some cells would stay as progenitors and some cells differentiate. To confirm this, we used the *esg-FO* system to overexpress *UAS-da* and *UAS-emc* and we found that there were few surviving clusters, and those were mainly progenitor (*Hdc<sup>+</sup>*) cells, although there were few *Hdc<sup>-</sup>* *GFP<sup>+</sup>* cells that had small nuclei (**fig. 1.13C-D**). This phenotype could indicate that *UAS-da* transgene is stronger than *UAS-emc*, and therefore we are producing a *UAS-da* phenotype, with some escapers that would start absorptive differentiation an escape from apoptosis but cannot undergo through endoreplication.

Altogether, our data indicates that Emc would have three main functions: inhibit Da function to drive cell differentiation, inhibit EB dedifferentiation and control the expression of *Dl*, although the two last activities could be using the same mechanism.



**Figure 1.13. Da function is regulated by Emc**

**A-B.** ISCs differentiate into ECs (arrow heads) when both *da* and *emc* expression are impaired. We used MARCM clones to induce knock down of *da* in *emc*<sup>LL02590</sup> mutant clones (A) or to induce knock down of *emc* in *da*<sup>10</sup> mutant clones (B).

**C.** Co-expression of *UAS-da* and *UAS-emc* maintains the majority surviving clusters as progenitors. Guts show a low survival of the GFP<sup>+</sup> population.

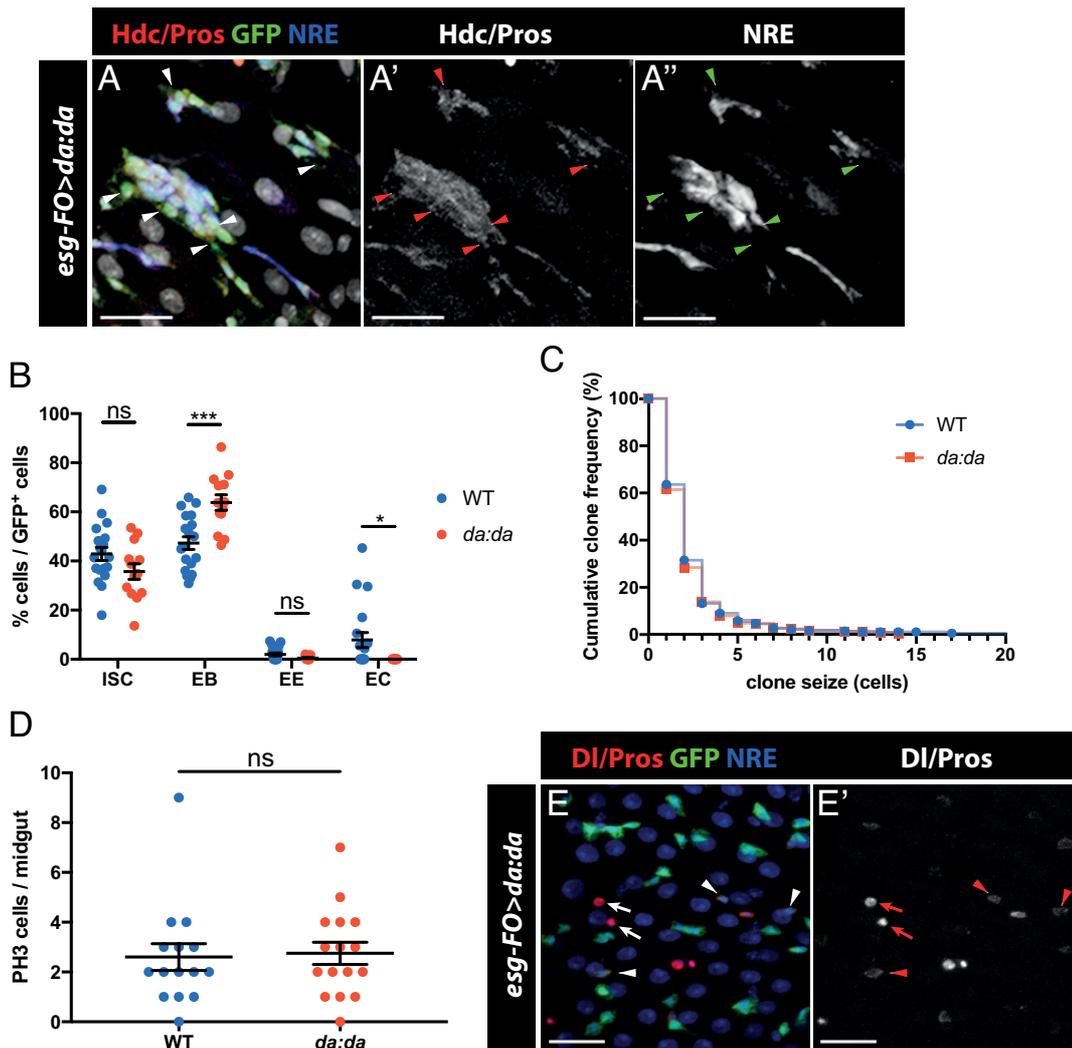
**D.** Detail of panel C, it can be observed that some GFP<sup>+</sup> cells are diploid Hdc<sup>-</sup> NRE<sup>-</sup> cells.

Data information: **scale bars, 20μm**

### 1.B.3.2. Da homodimerizes to maintain the progenitor state

Thereafter, we wanted to determine the nature of the dimer that Da was participating in to maintain the progenitor state. In the next chapter of this thesis we will explore a bHLH candidate called Cropped. Da dimerises typically with bHLH class II to bind specific E boxes DNA sequences (CANNTG). Biochemistry assays showed that Da can form homodimers that weakly binds DNA (Murre *et al.*, 1989a; Murre *et al.*, 1991), similarly as the formation of E12 homodimers in mammals (Sun and Baltimore, 1991). In 2014, Tanaka-Matakatsu *et al.* found that Da homodimers induce retinal neuron differentiation (Tanaka-Matakatsu *et al.*, 2014).

We wanted to explore the possibility that Da:Da dimers have a function in the gut maintaining the progenitor state. For these means we expressed a *UAS-da:da*: two *da* sequences linked by a flexible peptide, whose proximity forces the dimerization (this approach was described in mammals (Neuhold and Wold, 1993) and later it was used in *Drosophila* to create tethered *twist* homodimers and *twist:da* heterodimers (Castanon *et al.*, 2001)). This *UAS-da:da* construct has previously been used (Tanaka-Matakatsu *et al.*, 2014; Wang and Baker, 2015). We expressed *UAS-da:da* with the *esg-FO* system (**fig. 1.14A**, quantification in **1.14B**). As a result, we could see that progenitor cells were forming clusters without differentiating. About two thirds of the cells were NRE<sup>+</sup> and the other 1/3 were DI<sup>+</sup>. In addition, in big clusters with more than one ISC, we could never find two ISC in direct contact, and there were always EBs between them. But more importantly, these midguts did not present any sign of cell death, clusters were comparably distributed like the control, although when these clusters grew, instead of differentiating, they were accumulating progenitor cells. However, comparing the size of clusters, there was not any sign of incremented proliferation compared with wild-type (**fig. 1.14C**), and the size of the clusters were comparable. To verify that there were not more mitotic cells, we stained with PH3, and there were no significant differences with control flies (**fig. 1.14D**). As expected, we could not observe any increment in DI levels, similar than *UAS-da* over-expression (**fig. 1.134**). Together, these results show that Da homodimerize in order to maintain the progenitor state, although the homodimer is not responsible of the cell death, and there must be another molecule that partners with Da and is responsible of it.



**Figure 1.14. Da:Da prevents terminal differentiation**

**A.** *esg-FO* guts expressing a tethered *UAS-da:da* forms clusters of progenitor cells. Cells inside the cluster are *Hdc*<sup>+</sup> (A') and the majority are *NRE*<sup>+</sup> (A''), while there are few *NRE*<sup>-</sup> (arrow heads).

**B.** Quantification of the different cell populations in WT and *esg-FO>da:da* intestines. EB population increase significantly, in detriment of terminally differentiated ECs (WT N = 19 guts, *UAS-da:da* N = 13 guts) (\**p* < 0,05 \*\*\**p* < 0,001, multiple t-test) (WT data from figure 1.3A).

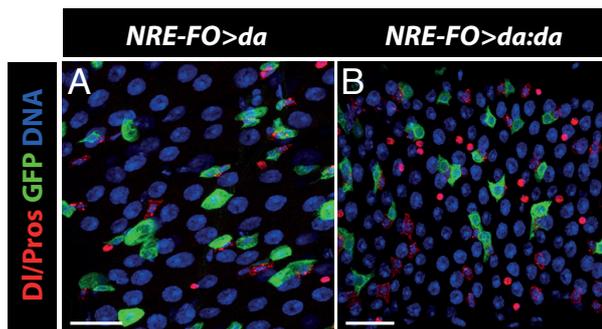
**C.** The cumulative clone frequency shows a comparable size distributions of the clusters (WT N = 552 clones, *UAS-da:da* N = 288 clones) (not significant *p* = 0.4187, Log-rank (Mantel-Cox) test) (WT data from figure 1.3A).

**D.** The total number of cell divisions in WT and *esg-FO>da:da* guts have no significant difference (WT N = 15 guts, *UAS-da:da* N = 16 guts) (not significant *p* = 0.6255, Mann-Whitney test).

**E.** *DI* levels in *esg-FO>da:da* posterior midguts (arrow heads) are lower than the signal of *Pros* (arrows).

Data information: **scale bars, 20μm**

Notably, the accumulation of EBs and the normal levels of DI in *UAS-da* and *UAS-da:da* clusters indicate that they are unable to promote de-differentiation. To test this, we expressed both transgenes with the *GBE-Su(H)-FO* system (**fig. 1.15A-B**). As expected, all the GFP cells were EBs, which were accumulating around the ISCs in both cases. Interestingly, when we overexpressed *UAS-da* only in EBs, these cells do not undergo into apoptotic processes, reinforcing the hypothesis that cell death is not induced in single cells. These results also suggest that Emc controls DI independently of Da:Da.



**Figure 1.15. Excess of Da in EBs produce EB accumulation**

**A.** In *NRE>da* midguts, all GFP<sup>+</sup> cells are DI<sup>-</sup>. GFP<sup>+</sup> cells do not differentiate. Cell death is not observable in this condition.

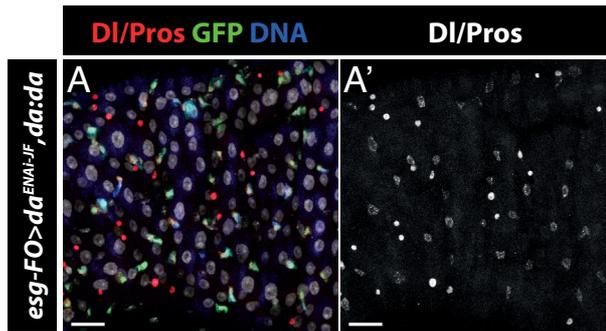
**B.** *NRE>da:da* EBs do not de-differentiate nor differentiate.

Data information: **scale bars: 20µm**

Next, we wanted to know the specific function of the Da:Da dimer, and we decided to use two different approaches:

- Remove all the endogenous *da* and express the forced dimer in all progenitor cells.
- Compare the differential expression of downstream genes when we express the monomer alone or the tethered dimer.

For the first approach we used the *esg-FO* system to express *UAS-da:da* and knock down the expression of *da*. To do this, we had to find a *UAS-da* RNAi line that does not target our *UAS-da:da* transgene. Our transgene does not contain the *da* 5'UTR, which is the *UAS-da<sup>RNAi</sup><sub>JF02488</sub>* target sequence. Therefore, we could remove *da* monomers that could bind other potential partners, while the Da:Da is present. We determined that the homodimer alone is enough to maintain the progenitor state, accumulating EBs and inhibiting their differentiation (**fig. 1.16A**).



**Figure 1.16. Da homodimer suffice to maintain the progenitor state**

**A.** *UAS-da<sup>RNAi-JF</sup> UAS-da:da* expression arrest terminal differentiation and accumulates progenitor cells, with only one  $DI^+$  cell per cluster (A')

Data information: **scale bars: 20μm**

### 1.B.3.3. Notch controls the function of Da homodimers

Our observations support a model in which Da:Da dimers are antagonized by Emc, whose function would be downstream of Notch activation. Therefore, in the presence of forced Da:Da expression, Notch activation should not enforce terminal absorptive differentiation. To test this, we co-expressed *UAS-da:da* and *UAS-H<sup>RNAi</sup><sub>HMS01182</sub>* in all progenitor cells with *esg-FO* for 14 days. Here, it is possible that there is a delay with the added kinetics of RNAi biogenesis, although, for our experiment, it would be advantageous that *UAS-da:da* is expressed before *H* is knocked down, as then we can assess if Notch activation can force differentiation in the presence of Da:Da without worrying that *H* knockdown would lead to irreversible differentiation before Da:Da was in place. Indeed, Notch activation could not induce terminal differentiation when Da:Da is present, as all  $GFP^+$  cells were progenitor cells ( $Hdc^+$ ) (**fig. 1.17A**, quantification in **1.17B**). Interestingly,  $Hdc^+$  cells expressed different levels of *GBE-Su(H)-lacZ* + (**fig. 1.17A**), from high to no expression in a graded manner, but none of them expressed DI (**fig. 1.17C**). In the absence of DI expression (and therefore NICD production) and the loss of H-mediated repression, the observed lacZ expression is possibly due to the basal enhancer activity of the GBE sequence in the *GBE-Su(H)-lacZ* transgene.

The presence of only committed cells when Notch is activated and *UAS-da:da* expressed was expected, as we have previously observed that with the expression of *UAS-da:da* alone, all  $GFP^+$  cells are  $DI^-$  due to Notch activation (**compare 1,17A with fig. 1.14A**). This points out the importance of Emc to inhibit the formation of Da homodimers in the EB to allow absorptive differentiation. This suggests that when cells do not have Da, Notch might not be required to drive differentiation. We demonstrated this hypothesis by knocking down the expression of *da* and expressing *H* to block Notch activity (**fig.**

**1.17D**, quantification in **1.17E**). We observed that the majority of cells adopted the EC fate even when Notch signalling was impaired. We could also see terminal absorptive differentiation with the double knock down of *da* and *Notch* using the *esg-FO* lineage tracing system (**fig. 1.17F**).

Finally, the data indicates that loss of *Notch* and expression of *UAS-da:da* have two main differences: (1) while Da:Da overexpressing cells can differentiate into EBs, but maintain the progenitor state, *Notch* mutant clones are composed either by ISCs or by EE cells and (2) *Notch* mutant cells have proliferation rates that are largely increased, while we demonstrated that *UAS-da:da* over-expression did not affect cell divisions. Therefore, we expressed *UAS-da:da* and knocked down the expression of *Notch* with the *esg-FO* to see if in these conditions cell proliferation would still be increased as when *N* is knocked down alone, or on the contrary, the expression of *UAS-da:da* would reduce the proliferation (**fig. 1.17G**). After 7 days, we could observe clusters of progenitor cells, but there was no sign of over-proliferation. We could find some *GBE-Su(H)-lacZ*<sup>+</sup> cells, but never more than one per cluster and the majority of cells were ISCs. The presence of *GBE-Su(H)-lacZ*<sup>+</sup> cells could be due to a delay of the effect of *UAS-Notch*<sup>RNAi</sup> or perdurance of the  $\beta$ -galactosidase. We could also find single Pros<sup>+</sup> GFP<sup>+</sup> cells, but there was no sign of the typical EE clusters of *Notch* loss of function. Together, these results indicate that Notch is epistatic to Da and Da:Da dimers prevent cell differentiation and proliferation (**fig. 1.17G**).

**Figure 1.17. *da* function is regulated by Notch activity (next page)**

- A.** Expression of *UAS-da:da* and knock down of *H* using the *esg-FO* system produce an arrest of terminal absorptive differentiation. Except some Pros<sup>+</sup> escapers, all cells are Hdc<sup>+</sup> and express different levels of *NRE*.
- B.** Quantification of EBs in *esg-FO>da:da, H<sup>RNAi-HMS</sup>* guts. As all Hdc<sup>+</sup> cells are DI<sup>-</sup>, we consider them as EBs that, due a prolonged time without having Notch activated by DI, Notch reporter expression has ceased. (*UAS-H<sup>RNAi-HMS</sup>* N = 15, *UAS-da:da, UAS-H<sup>RNAi-HMS</sup>* = 15) (\*\*\*)p<0,001, Mann-Whitney test).
- C.** All GFP<sup>+</sup> cells are DI<sup>-</sup> in *esg-FO>da:da, H<sup>RNAi-HMS</sup>* guts.
- D.** Expression of *UAS-da<sup>RNAi-JF</sup>, UAS-H* in progenitor cells with *esg-FO* promotes terminal differentiation, with polyplod DI<sup>-</sup> NRE<sup>-</sup> GFP<sup>+</sup> cells.
- E.** Quantification of ECs in *esg-FO>da<sup>RNAi-JF</sup>, H* guts. (*UAS-H* N = 21, *UAS-da<sup>RNAi-JF</sup>, UAS-H* = 21) (\*\*\*)p<0,001, Mann-Whitney test).
- F.** Double knock down of *da* and *Notch* produce a terminal absorptive differentiation in all progenitor cells.
- G.** Expression of the tethered dimer *UAS-da:da* and knock down of *Notch* with *esg-FO* arrest absorptive differentiation and maintains the majority cells as DI<sup>+</sup> (arrow heads). Cells can still differentiate into Pros<sup>+</sup> cells (arrows).

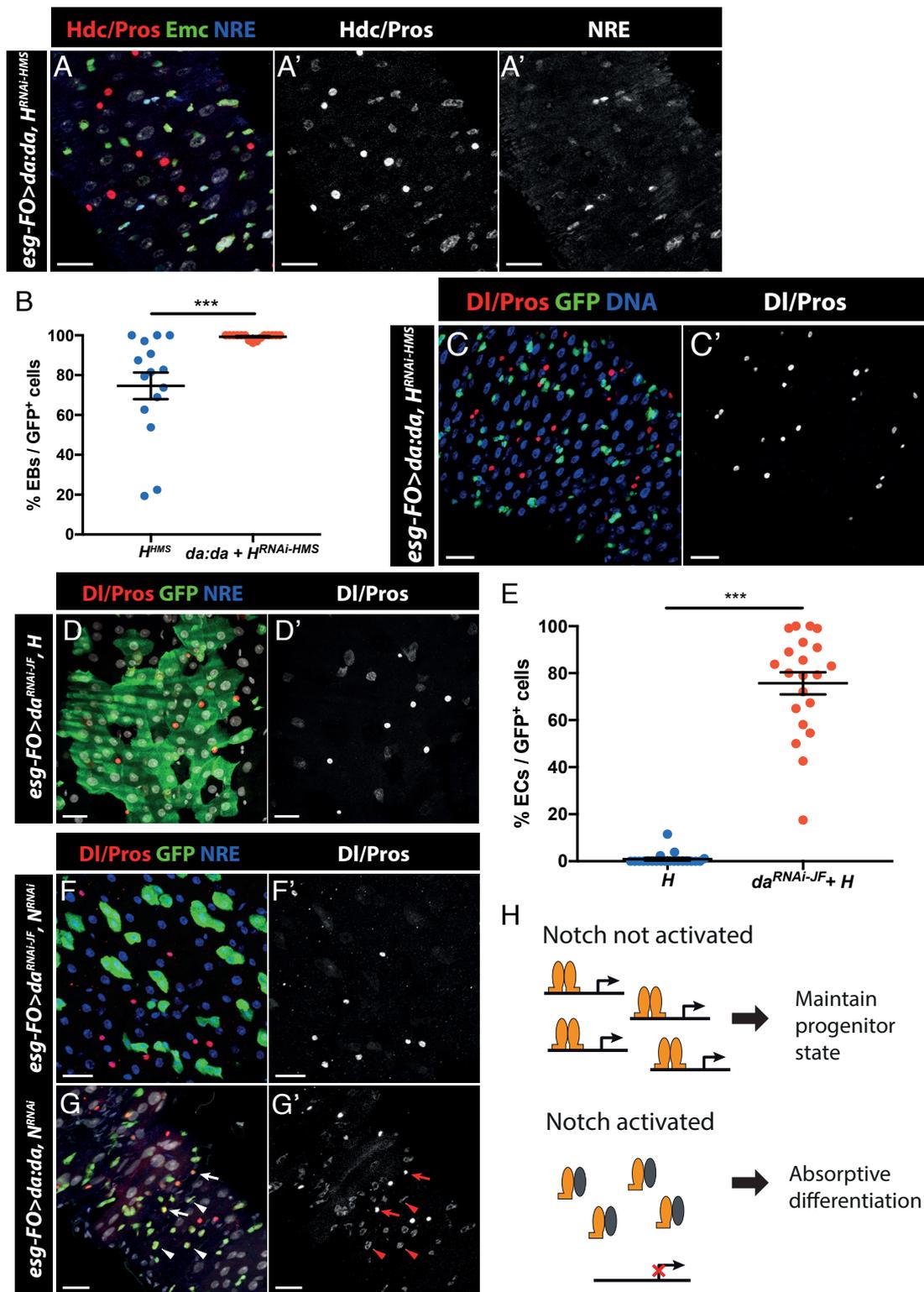


Figure 1.17. *da* function is regulated by Notch activity (continued)

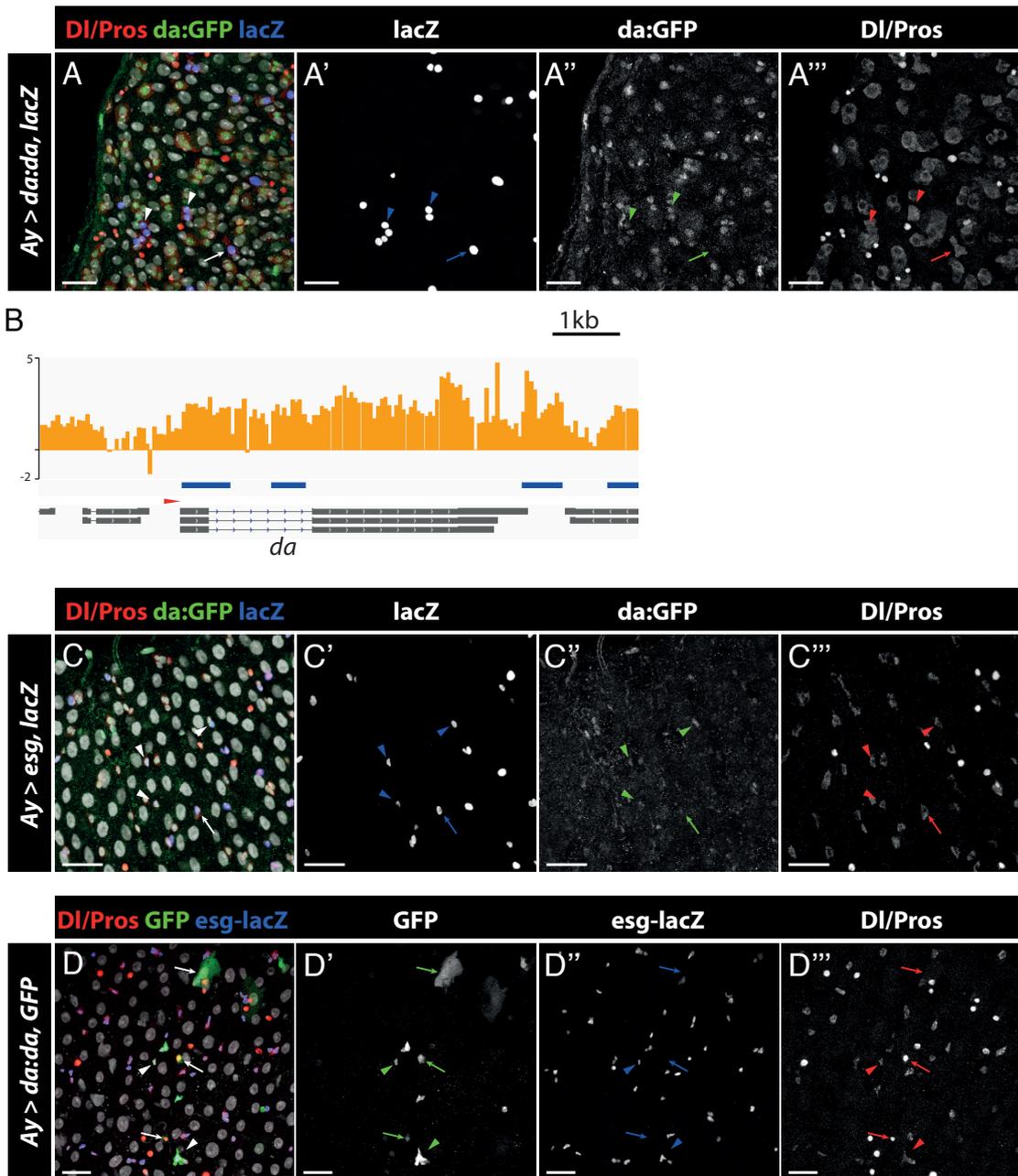
**H.** When Notch is not activated, Da (orange) homodimerizes to drive the expression of genes that maintain the progenitor state. In Notch activated cells, Emc (grey) dimerizes with Da, inhibiting its transcriptional activity.

Data information: scale bars, 20µm

#### 1.B.3.4. Da homodimers do not auto-regulate and maintain stemness in parallel to Esg

Next, we wondered if Da:Da could be autoregulating *da* expression. Previous studies indicate that although in many tissues the expression of *da* is ubiquitous, it has a complex transcriptional regulation. *da* is autoregulated both positively and negatively in the somatic ovary (Smith and Cronmiller, 2001) and in the eye-antennal disc (Bhattacharya and Baker, 2011). We addressed this question by using the *Ay-Gal4*, producing clones that express *UAS-da:da* and comparing the intensity of Da:GFP inside and outside the clones (**fig. 1.18A**). There were no differences in the GFP intensity within the tissue. Thus, it seems that Da:Da does not auto-regulate *da* expression.

Esg is a transcription factor expressed only in ISCs and EBs, and it has been shown to be important for maintaining the progenitor state. Loss of *esg* results in differentiation (either secretory or absorptive), while overexpressing *esg* in MARCM clones produces clusters of progenitor cells with accumulation of EBs (Korzelius *et al.*, 2014; Loza-coll *et al.*, 2014). Importantly, guts over-expressing *esg* in progenitor cells have a lower mitotic rate after injury. All these features are similar with the over-expression of *UAS-da:da*. Moreover, although there is no consensus in this regard, some reports indicate that *esg* is epistatic to Notch (Li *et al.* 2017). Thus, *esg* could be either upstream or downstream of *da*. Available DamID data for Esg (Korzelius *et al.*, 2014; Loza-coll *et al.*, 2014) showed binding of Esg in the *da* promoter sequence (**fig. 1.18B**). However, RNAseq data (Korzelius *et al.*, 2014) of knocked down and over-expressed *esg* in progenitor cells showed no significant up-regulation or down-regulation of *da*. Thus, *da* seems not to be under Esg control. Moreover, we induced clones expressing *UAS-esg* with *Ay-Gal4* to check the variability on the Da:GFP reporter (**fig. 1.18C**). In accordance with the RNAseq data, the expression of *da* was not altered by Esg.



**Figure 1.18. Da and Esg work in parallel**

**A.** Overexpression of *UAS-da:da* in clones shows that *da* expression is variable inside the clones even in the same cell type. Some *DI*<sup>+</sup> cells do not express *da* (arrows), while some have expression (arrow heads).

**B.** Esg DamID profile surrounding the *da* loci. Bars represent the median log<sub>2</sub> (intensity ratio) between Esg:Dam and Dam-control profiles. Blue lines represent escargot binding sites. The red arrow head indicates *da* polarity. In gray there are the different genes and transcripts. (Raw data acquired from Korzelius et al, 2014/Loza-Coll et al, 2014).

**C.** Esg does not regulate *da* expression, as come clones express *da* (arrow heads) and some do not (arrows).

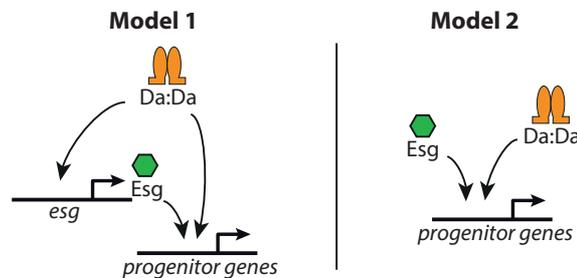
**D.** Da:Da homodimers do not regulate *esg* expression. All progenitor cells inside and outside the clones has the same *esg* expression (arrow heads represent *DI*<sup>+</sup> cells in clones), while differentiated cells in clones do not express *esg* (arrows).

Data information: **scale bars, 20μm**

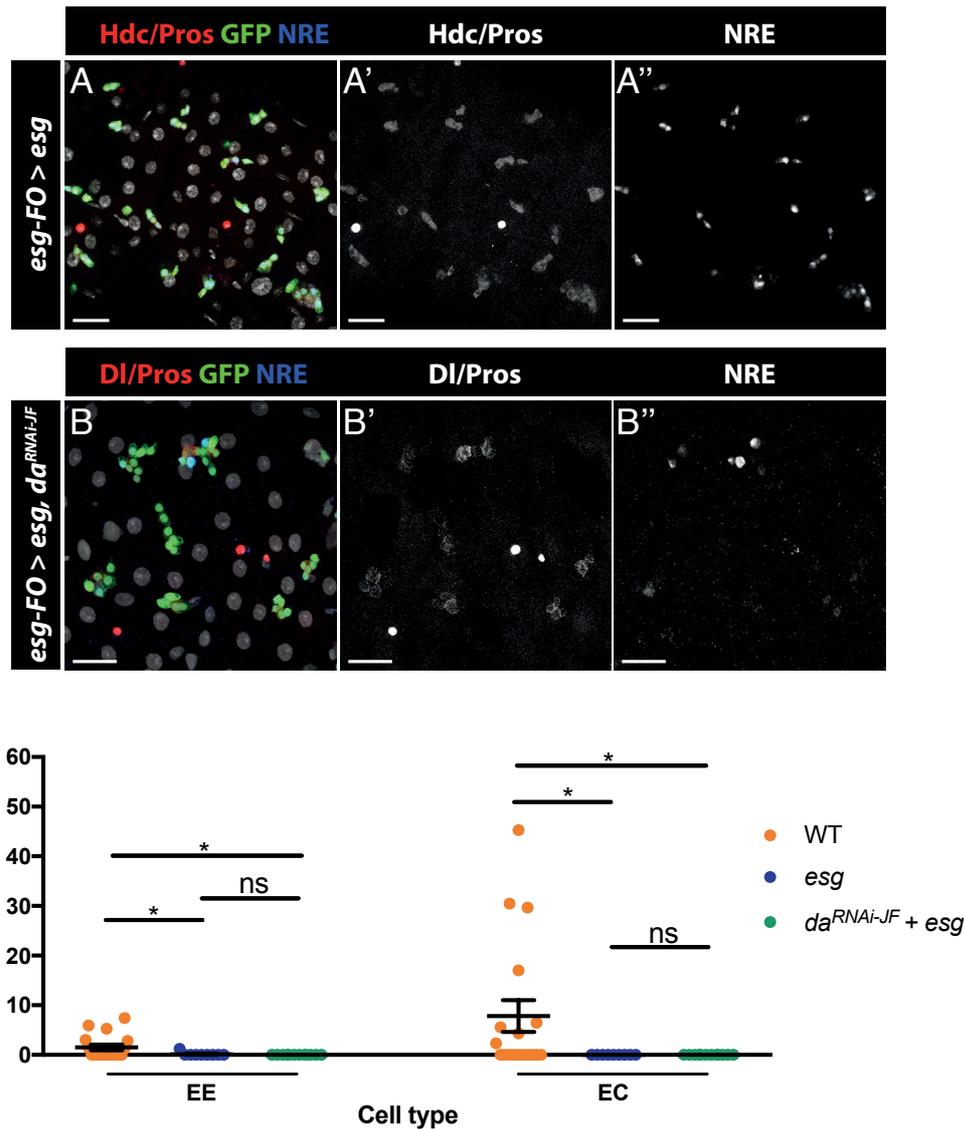
To test if *esg* activity could be controlled by Da:Da, we knocked down expression of *da* and expressed *UAS-esg* with *esg-FO* (**fig. 1.19A-B**, quantification in **1.19C**). All cells could be maintained in the progenitor state even in the absence of *da*, which could suggest that *esg* role is downstream of Da. If that were the case, when *esg* is knocked down (*UAS-esg<sup>RNAi</sup><sub>HMS00025</sub>*) and we express *UAS-da:da*, we would expect that all cells differentiate into ECs and EE cells (**fig. 1.20A,B**, quantification in **1.20C**). Interestingly, in these conditions, progenitor cells did not differentiate into ECs and, although present, secretory differentiation was significantly reduced compared to *esg* knock down flies. This is interesting, as previous reports have shown Esg as an inhibitor of the secretory differentiation, and knock down of *esg* induces an increase of EE cells, but we can see that this increment is reduced when *UAS-da:da* is expressed (**fig 1.20C**).

These results suggest two different hypotheses:

1. Da:Da has different downstream targets that regulate the progenitor state and one of them is *esg*. When *esg* is not present, the other targets are sufficient to inhibit the absorptive differentiation.
2. Da:Da and Esg work in parallel to maintain the progenitor state.



To address this question, we used the *Ay-Gal4* system to induce clones expressing *UAS-da:da* and compare the levels of the *esg-lacZ* (**fig. 1.18D**). The resulting clones did not show any increase on the levels of *esg*. Therefore, the data suggests that Da:Da and Esg are not epistatic and work in parallel (**Model 2**).



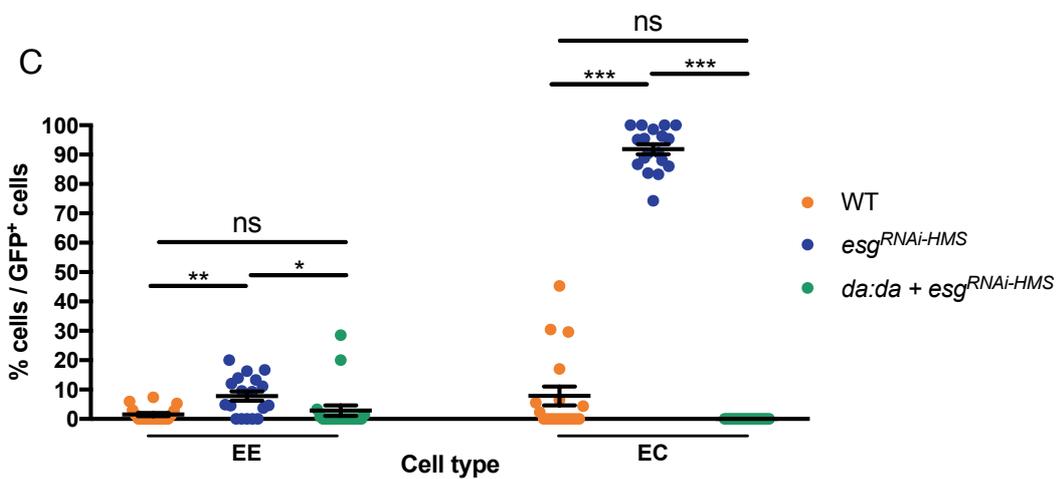
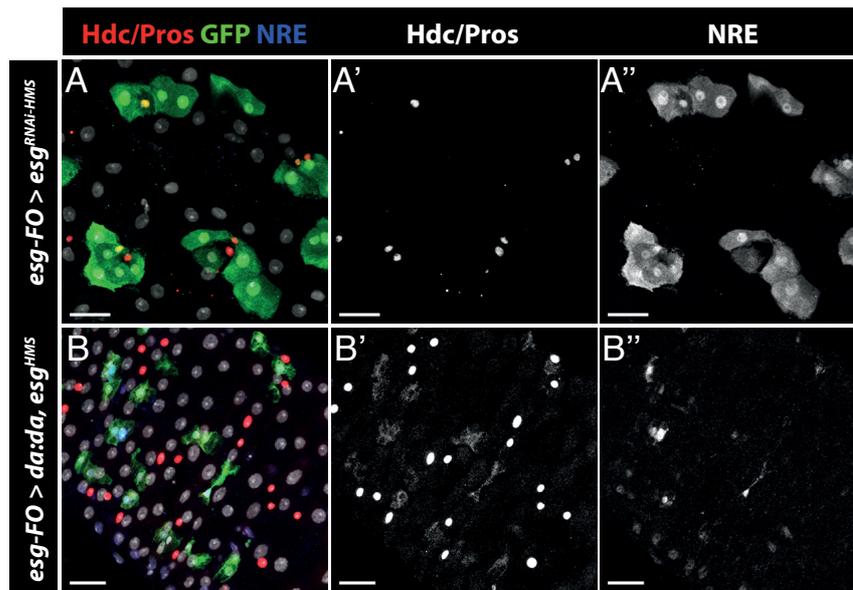
**Figure 1.19. Over-expression of *esg* maintains the undifferentiated state in the absence of *da***

**A.** Over-expression of *UAS-esg* arrests terminal differentiation.

**B.** Expression of *UAS-esg, UAS-da<sup>RNAi-JF</sup>* in progenitor cells arrests terminal differentiation.

**C.** Quantification of EE cells and ECs in *esg>UAS-esg* and *esg>UAS-esg, UAS-da<sup>RNAi-JF</sup>* conditions, compared with wild type midguts. In both conditions differentiation is arrested completely (WT N=21; *UAS-esg* N=9; *UAS-esg, UAS-da<sup>RNAi-JF</sup>* N=27) (\* $p < 0.05$ , Kruskal-Wallis/Dunn test).

Data information: **scale bars, 20 $\mu$ m**



**Figure 1.20. Da:Da maintains the progenitor state when *esg* is lost**

**A.** *esg* knock down results in terminal differentiation into the absorptive and secretory fates.

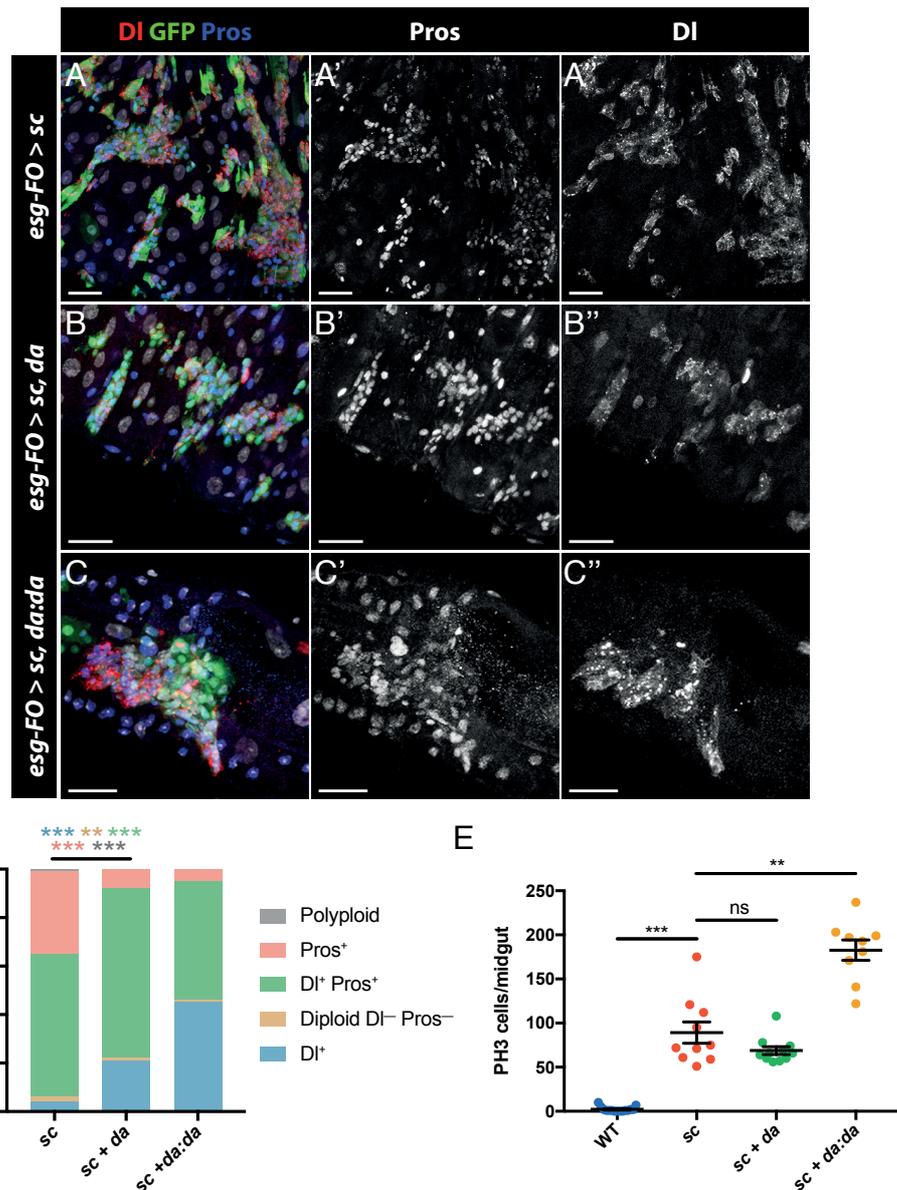
**B.** Knock down of *esg* and expression of *UAS-da:da* using *esg-FO* arrest absorptive differentiation and secretory differentiation.

**F.** Quantification of EE cells and ECs, comparing WT flies with knock down of *esg* alone in progenitor cells and with knock down of *esg* and over-expression of *da:da*. (WT N=21; *UAS-esg<sup>RNAi-HMS</sup>* N=18; *UAS-da:da, UAS-esg<sup>RNAi-HMS</sup>* N=18) (\*p<0.05, \*\*p<0.01, \*\*\*p<0.001, Kruskal-Wallis/Dunn test).

Data information: scale bars, 20 μm

1.B.3.5. Sc forms heterodimers with Da and functionally opposes Da:Da dimers to promote secretory differentiation

The bHLH class II Sc has been described as the main initiator of secretory differentiation in *Drosophila* (Bardin *et al.*, 2010) by promoting the expression of the transcription factor *pros* (Wang *et al.*, 2015). Sc binding into a certain *pros* enhancer is opposed by *Esg* (Li *et al.*, 2017a). Our results show that in the absence of *esg*, if Da:Da dimers are present, secretory differentiation is still reduced when compared to the *esg* knock down (**fig. 1.20C**). To bind DNA, all members of the AS-C form heterodimers with Da and therefore, more Sc:Da dimers lead to a decrease of Da:Da dimers. It is possible that for secretory differentiation, not only Sc:Da heterodimers are needed, but also the titration of the Da:Da homodimers. Therefore, we proceeded to create three conditions to explore the different scenarios: *sc* expressed alone to titrate as much Da monomer as technically possible, *sc* and *da* co-expressed to create a situation where neither heterodimers nor homodimers are scarce, and finally *sc* and tethered *da:da* to ensure that the homodimer is present at high levels and cannot be titrated. We used the *esg-FO* system to induce the expression for three days. We could observe that the expression of *UAS-sc* induced an over-proliferation, producing big clusters (**fig. 1.21A**, quantification in **1.21D**). Surprisingly, 63% of the cells in these clusters were  $Dl^+$  cells. However, only a 4% were expressing *Dl* alone and the other 59% were expressing *Dl* and *Pros*. It has been described that before EE cells are terminally differentiated, they differentiate into a very short-lived transient state that express *esg*, *Dl* and *pros* (Biteau and Jasper, 2014; Chen *et al.*, 2018; Zeng and Hou, 2015). Importantly, these transient cells, called pre-EE cells, can proliferate once before terminally differentiating (Zeng and Hou, 2015). Interestingly, we increased the pre-EE ( $Dl^+ Pros^+$ ) population when we co-expressed *UAS-sc* and *UAS-da*, while the terminally differentiated EE cells only represented an 8% compared with the 38% when only *UAS-sc* is expressed (**fig. 1.21B, D**). This indicated that the co-expression of *UAS-sc* and *UAS-da* is important to retain this transient state. Moreover, this co-expression produced clusters of  $Dl^+ Pros^-$  cells, which represent a 21% of the  $GFP^+$  cells. This phenotype was even more drastic when we co-express *UAS-sc* with *UAS-da:da*, with a 45% of  $Dl^+ Pros^-$  and 49% of  $Dl^+ Pros^+$ , with only 5% of  $Dl^- Pros^+$  (**fig. 1.21C, D**). Therefore, Sc and Da are antagonistic in allowing the terminal differentiation of EE cells, but they cooperate to produce  $Dl^+$  (which Da or Da:Da do not increase).



**Figure 1.21. Sc induces proliferation and secretory differentiation**

**A-C.** Expression of *UAS-sc* alone (A), with *UAS-da* (B) or with *UAS-da:da* (C) using *esg-FO* for 3 days. Expression of *sc* increase the number of  $DI^+$  cells,  $Pros^+$  cells and the appearance of double positive  $DI^+ Pros^+$  cells.

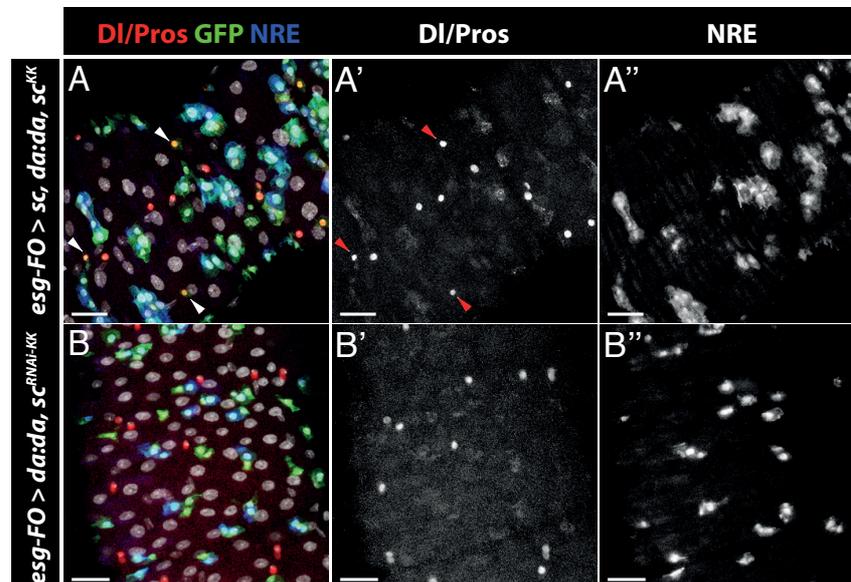
**D.** Estimated population distribution of A-C (see materials and methods) (*UAS-sc* N = 1241 cells; *UAS-sc, UAS-da* N = 2131 cells) (\*\* $p < 0.01$ , \*\*\* $p < 0.001$ , Binomial regression. Each cell type significance level is coloured accordingly).

**E.** Quantification of mitotic events in the whole posterior midgut of WT and guts of A-C (WT N=15 guts; *UAS-sc* N=10 guts; *UAS-sc, UAS-da* N=11 guts; *UAS-sc, UAS-da:da* N=9 guts) (\*\* $p < 0.01$ , \*\*\* $p < 0.001$ , Kruskal-Wallis/Dunn test).

Then, we wanted to check the proliferation in the three conditions, as these big clusters indicated an important raise of the mitotic index. We could observe by PH3 staining that there were significantly more cell divisions in all three conditions (fig. 1.21E). Interestingly, even if the number of mitosis per gut are greatly augmented in *UAS-sc*

over-expression condition, co-expression of *UAS-sc* with *UAS-da:da* double the number of mitosis, while co-expression of *UAS-sc* and *UAS-da* does not change significantly the number of divisions respect *UAS-sc* alone. Thus, Sc is important for cell division not only in pre-EE, but also in ISCs, and presence of Da:Da exacerbates the number of mitoses.

Notably, we obtained these strong phenotypes only after three days of expression. As *sc* is expressed weakly in ISCs (Chen *et al.*, 2018), we wondered if we could obtain more ISCs if we express *UAS-da:da* and weakly *UAS-sc*. To this end, we co-expressed *UAS-da:da* and *UAS-sc*, and at the same time *UAS-sc<sup>RNAi</sup><sub>JF02104</sub>* (**fig. 1.22A**). After three days, we observed that the secretory differentiation and proliferation were reduced. Moreover, we obtained clusters that were accumulating EBs, similarly to *UAS-da:da* expression. Finally, when we expressed *UAS-da:da* and *UAS-sc<sup>RNAi</sup><sub>JF02104</sub>* without expressing *UAS-sc*, we also obtained accumulation of EBs, without secretory differentiation, forming smaller clusters (**fig. 1.22B**). The data indicates that mild *sc* over-expression induces little secretory differentiation, but not an increase of ISCs as cells can progress to EBs, and when *sc* expression is knocked down, ISCs lose the capacity to differentiate into EE cells.



**Figure 1.22. Intermediate expression of *sc* cannot prevent EB differentiation**

**A.** Expression of *UAS-sc* and *UAS-da:da* and knock down of *sc* expression, reducing the high expression of *sc*. Cells can still differentiate into EE cells (arrow heads), but they differentiate terminally without proliferating. ISCs cannot form DI<sup>+</sup> clusters, and progenitor cells are accumulated as EBs.

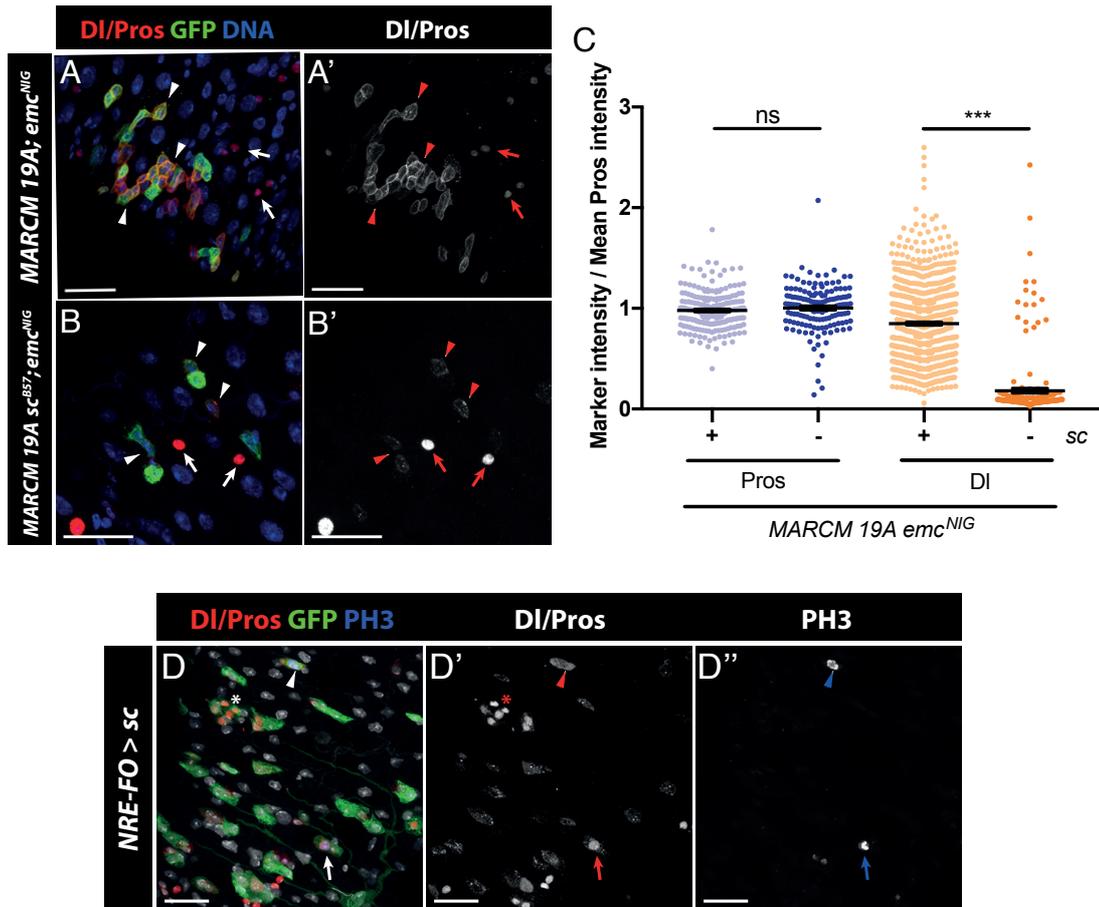
**B.** Expression of *UAS-da:da* with *UAS-sc<sup>RNAi-KK</sup>* halt all terminal differentiation. Progenitor clusters contain DI<sup>+</sup> cells or NRE<sup>+</sup> cells.

Data information: **scale bars, 20µm**

#### 1.B.3.6. Emc downregulates Dl by antagonising Sc

Importantly, our data showed that when we expressed *UAS-sc* an important number of cells expressed *Dl*. Interestingly, it has been recently observed that *Notch* loss of function induces an up-regulation of *sc* in the ISC clusters (Li *et al.*, 2017a). Therefore, it is possible that *sc* can induce the expression of *Dl* in ISCs. If that is the case, *emc* could control this expression and that would explain the up regulation of *Dl* when we knock down the expression of *emc*. To check this hypothesis, we knocked down the expression of *emc* in MARCM clones that were wild type for *sc* or were *Df(1)sc<sup>B57</sup>* mutant clones (deficiency of the whole *AS-C*, comprising the bHLH genes *sc*, *ac*, *l(1)sc* and *ase*; overexpression of *ac* and *ase* has also been shown to induce EE differentiation (Bardin *et al.*, 2010; Guo and Ohlstein, 2015) ) (**fig. 1.23A-C**). The control flies where only *emc* was knocked down contained clones with high levels of *Dl*. However, we could reduce the up-regulation of *Dl* with a null allele of *sc*. Therefore, Sc regulates the expression of *Dl* and Emc binds Sc to regulate this activity.

We have observed that Emc inhibits the function of Sc to regulate the expression of *Dl*, and this suggests that different levels of *sc* can trigger different effects. When *sc* is expressed at high levels, it cannot be titrated by Emc and there is secretory differentiation. However, lower expression of *sc* is not enough to promote differentiation but can induce the expression of *Dl*. We wondered then if apart from *Dl*, Sc could regulate other genes that maintained the stemness, and therefore, when over-expressed in EBs, *sc* could induce de-differentiation. We used the *GBE-Su(H)-FO* system to express *UAS-sc* (**fig. 1.23D**). Interestingly, we found many *Dl*<sup>+</sup>, *Pros*<sup>+</sup> and *Dl*<sup>+</sup> *Pros*<sup>+</sup> cells that were labelled with GFP. This result indicates that EBs can de-differentiate when *sc* is expressed at high levels and, because the de-differentiated ISCs will maintain the *UAS-sc* expression, they can differentiate into pre-EE cells and finally into EE cells. The large amount of GFP<sup>+</sup> cells suggests that after de-differentiation cells gain the capacity to proliferate. We wanted to confirm this by staining with PH3 and we could observe that GFP cells were now proliferating and ISCs and pre-EE cells were both proliferating (**fig. 1.23D**). Together, these results suggest that *sc* has a dual function: on one side it promotes the secretory differentiation, and on the other it is important for the expression of stem genes.



## 1 B.4. Conclusions

In this chapter we have observed that over-expressing *da* in progenitor cells, terminal differentiation was arrested, although apoptosis was also induced (**fig. 1.12**). Apoptosis could be rescued co-expressing *p35* (**fig. 1.12**). Interestingly, we have found that Da homodimerizes to maintain the progenitor state (**fig. 1.14**). However, although we have stated that Emc regulates negatively Da (**fig. 1.13**), Da:Da over-expression did not increase *Dl* levels (**fig. 1.14**).

We have also confirmed that Notch function regulates negatively Da (**fig. 1.17**), most likely through Emc.

Esg, a transcription factor responsible to maintain the progenitor state, acts in parallel with Da:Da, neither upstream nor downstream (**fig. 1.18**). Over-expression of one of the two transcription factors are sufficient to maintain the progenitor state, even in the absence of the other (**fig. 1.19** and **fig. 1.20**).

We have confirmed that Sc induces secretory differentiation, although we have observed that sustaining the expression of *sc*, cells remain as pre-EE cells and keep proliferating. Moreover, when we co-expressed *sc* with *da* or *da:da*, the number of terminally differentiated EE cells was reduced, while the number of ISCs increased (**fig. 1.21**).

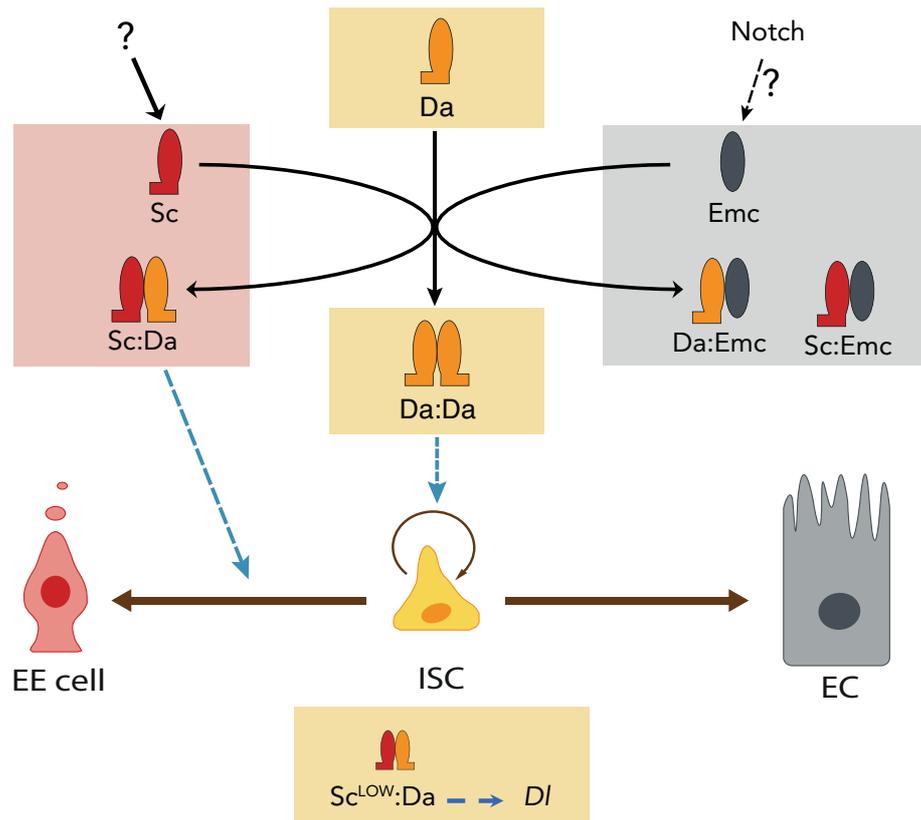
Finally, we have found that Sc is responsible to upregulate *Dl* levels when *emc* expression is knocked down (**fig. 1.23**). Furthermore, *sc* expression in EBs induces de-differentiation into ISCs, which can then differentiate into secretory cells (**fig. 1.23**).

## 1 A&B.5. Discussion

Ours results indicate that a regulatory network of bHLH transcription factors controls the proliferation and differentiation of ISCs. In previous studies, it has been reported that *da* is necessary to maintain the progenitor state in the posterior gut (Bardin *et al.*, 2010). We have found that Da is also sufficient to inhibit differentiation, as when we expressed *da*, cells were accumulated in clusters as progenitor cells. More importantly, the accumulation was also observed when we expressed a covalent *da:da* homodimeric form, indicating that Da homodimerize to inhibit terminal differentiation. The HLH Emc has been studied extensively as an inhibitor of the transcriptional activity of Da and bHLH class II (Cubas *et al.*, 1991; Van Doren *et al.*, 1991), and we confirmed that *emc* is expressed in the *Drosophila* midgut and antagonise Da. We showed that its function is found downstream of Notch activity for correct acquisition of the absorptive fate. Moreover, we also found that Emc is important in the EB to inhibit de-differentiation. Emc performs this repression by inhibiting Sc, which has a double function: on one hand starts the secretory differentiation, as reported previously (Bardin *et al.*, 2010; Li *et al.*, 2017a; Zeng and Hou, 2015), and on the other *sc* induces expression of stem genes like Dl. Altogether, Da, Emc and Sc constitute a network that regulates the stem cell fate (**fig. 1.24**).

### 1.5.1. Imbalance of the bHLH equilibrium controls proliferation and differentiation

All our experiments have been performed by inducing high expression of a certain bHLH gene or by removing it completely, in specific cell types. However, we and others have observed that *emc* (**fig. 1.1**), *da* (Bardin *et al.*, 2010) and *sc* (Chen *et al.*, 2018) are all expressed in ISCs and there has to be a constant functional tension between them. Li and Baker described that this tension is important for the stability of the bHLH factors, as the Emc half-life increases ten times when it has dimerized with Da compared with the monomeric Emc (Li and Baker, 2018). They also showed that when dimerization with Emc, Da degradation was increased. Therefore, Da is important for Emc stability. In addition, it has been proposed that prior to neurogenesis of the adult bristles, the onset of neural precursor selection occurs when AS-C genes are expressed in a context where there is a prior balance between Emc and Da, tipping this balance towards formation of



**Figure 1.24. A network of bHLH factors controls self-renewal and bipotent differentiation in the *Drosophila* midgut** (see main text for details)

Da:ClassII heterodimers – but in the absence of Emc and AS-C, DaDa homodimers can initiate bristle development too (Troost *et al.*, 2015). This breaks the equilibrium between Da and Emc, and now Da will be able to dimerise with proneural genes and/or homodimerize. Altogether, this suggests that bHLH factors are constantly interacting in the different cells of the gut, and depending of the levels of each bHLH factors, different binding combinations will be formed and cells will respond differentiating (Emc:Da/Sc for absorptive differentiation and Sc:Da for secretory differentiation), proliferating (Da:Da and Sc:Da) or remaining undifferentiated (Da:Da).

Importantly, the predominance of one of these dimeric combinations depends on three variables: 1) bHLH levels (altered by expression and degradation rates), 2) dimerization affinity and 3) DNA binding. Biochemical assays showed that Da:Da homodimer binds very weakly to the DNA compared to heterodimers of Da and members of the AS-C. More important, AS-C proteins have more affinity for Da than Da for itself (Cabrera and Alonso, 1991). Therefore Da, in the presence of Sc (at similar expression levels) will form

heterodimers rather than homodimers, and when Da:Da dimers are formed, they weakly bind the DNA. In this study, they also observed that although members of the AS-C can dimerise with themselves, they cannot bind the DNA. In addition, in the presence of Emc, AS-C factors and Da preferentially form inactive dimers with Emc (Cabrera *et al.*, 1994). Thus, when present, Emc binds to Da and Sc inhibiting functional heterodimers and homodimers. Therefore, ISCs need to have enough Da as part of it is inhibited by Emc, another part binds with Sc and probably to another bHLH class II and still have to homodimerize to maintain the progenitor state even with a weak DNA binding. Hence, any variation of the levels of Emc or Sc is likely to shift the equilibrium towards less Da:Da and more Da:Partner.

#### 1.5.2. *emc* expression is controlled by different signals

*emc* can be expressed in all cell types, although its most clear function would be in EBs, where its expression is homogeneously high. It has been previously reported that *emc* expression depends on Notch activity in the formation of the wing margins and veins (Baonza *et al.*, 2000) and in the eye disc to promote growth by inhibiting Da (Spratford and Kumar, 2015b). Moreover, it has been observed two Su(H)/Mam binding sites close to the *emc* transcriptional start site and null mutations of either *Su(H)* or *mam* results in loss of *emc* expression in the eye disc (Spratford and Kumar, 2015a). However, we could not see an increased signal of *emc*<sup>CPT1002740</sup> when we express NICD (**fig. 1.9A**). Interestingly, it has already been reported that in the eye disc, Notch activation affects *emc* RNA levels, but not protein levels (Bhattacharya *et al.*, 2017).

Notably, *emc* expression in ISCs, EE cells and ECs vary from cell to cell, although in these cells Notch is not activated. We have not identified how *emc* is regulated in these cells, although a possibility would be that Da itself drives transcription of *emc* in the other cells, as expression of *da* seems to be ubiquitous (Bardin *et al.*, 2010), creating a negative feedback loop. This kind of regulation has already been described in the larval eye disc (Bhattacharya and Baker, 2011) and in mammals, where the Da homolog E47 can activate the expression of members of the Id family (Jordà *et al.*, 2007; Schwartz *et al.*, 2006). Interestingly, Jordà *et al.* (2007) reported that the Zn finger transcription factor Snail, founding member of the family of proteins which Esg belongs to, can positively regulate

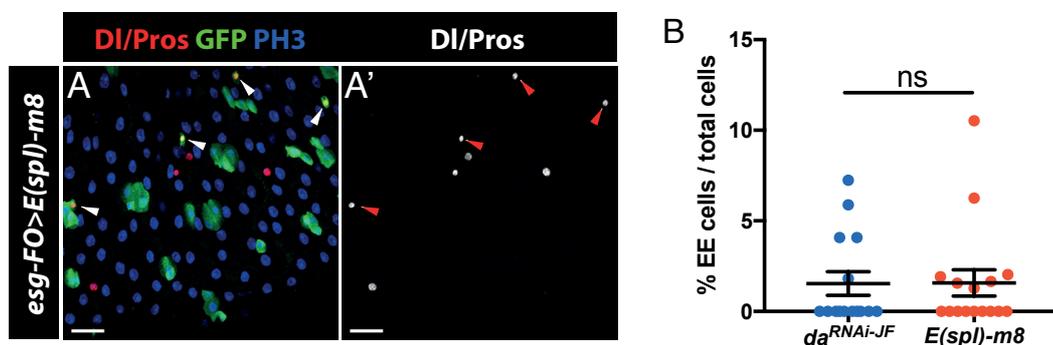
the expression of the *emc* homolog Id1 in immortalized canine epithelial kidney cell lines (Jordà *et al.*, 2007). By contrast, RNAseq data for ISCs and EBs where *esg* levels had been manipulated showed no significant differences in the expression of *emc* (Korzelius *et al.*, 2014). Therefore, we are planning to test whether Da can regulate *emc* using RNAseq analysis of tissue where *da* and *da:da* have been overexpressed in ISCs and EBs. In any case, the expression pattern of *emc* strongly suggests that one should consider more candidate regulators beyond Notch and Da.

Recent work suggests that our analysis of Emc function in the fly gut may be relevant in mammalian intestinal biology. Id family proteins are also expressed in the gut. Id1 expression is restricted to the CBC, the +4 cells and TA cells (Zhang *et al.*, 2014), while Id2 and Id3 expression increases when cells leave the crypt (Wice and Gordon, 1998). During development, Id2 is important for the proper specification of CBCs (Nigmatullina *et al.*, 2017). In addition, Id1 is necessary for regeneration when the intestine is damaged (Zhang *et al.*, 2014), and forced expression of Id1 induce adenomas (Wice and Gordon, 1998). These functions of Id factors contrast with the role of Emc in *Drosophila*, as Emc promotes differentiation (see **figure 1.8**). However, the Notch pathway in *Drosophila* promotes absorptive differentiation, while in mammals maintains the stemness. Therefore, it is conceivable that Emc and Id factors produce opposite effects in the *Drosophila* and mammalian intestines, respectively.

### 1.5.3. Da:Da regulates the progenitor state

The expression of *UAS-da:da* and *UAS-esg* in progenitor cells have very similar effects (see **figures 1.14A and 1.19A**). Therefore, we studied the possibility that *esg* could be upstream or downstream of *da* transcription and found that they likely work in parallel (see **figure 1.18**). So far, Esg has been reported as inhibitor of Da, as in the eye imaginal disc Esg promotes Da degradation and negatively regulate *da* expression (Yang *et al.*, 2010). Moreover, in cultured cells Esg can bind to the same E2 boxes than Da:Sc, inhibiting their transcriptional activity (Fuse *et al.*, 1994). However, there is no context known where Esg acts as *da* activator. Thus, it is not strange that they do not control each other's expression.

Importantly, while it is safe to assume that the tethered Da:Da form is insensitive to repression by Emc, it may still be susceptible of other forms of negative regulation. Da has two TADs, AD1 and LH, of which AD1 can bind to Orange domain-containing E(spl) proteins, recruiting them to E boxes to repress transcription (Zarifi *et al.*, 2012b). Therefore, the homodimeric Da:Da could be susceptible of repression by E(spl) proteins. Interestingly, we have observed that overexpression in progenitor cells of E(spl)-m8, which contains an Orange domain (Zarifi *et al.*, 2012b), drives terminal differentiation into both ECs (predominantly) and EE cells (fig. S1.1A). The level of secretory differentiation was indistinguishable from that resulting from the overexpression of monomeric Da (fig. S1.1B). However, we have noted that expression of *UAS-da:da* while simultaneously activating the Notch pathway by knocking down *H*, (a situation where transcription of *E(spl)* genes should be increased), does not promote terminal differentiation; instead, cells remain as diploid  $DI^- NRE^+$  (see fig. 1.17A,C). Therefore, our data suggests that E(spl) is not able to inhibit the transcriptional activity of Da:Da, and we propose that regulation of Da depends mostly on Emc.



**Figure S1.1. E(spl)-m8 induces differentiation**

**A.** *esg-FO>E(spl)-m8* in posterior midguts induces terminal differentiation, mainly to the absorptive fate, although cells could also differentiate into EE cells (arrow heads).

**B.** Quantification of EE cells in *esg-FO>UAS-da<sup>RNAi-JF</sup>* and *esg-FO>UAS-E(spl)-m8* shows no significant differences in secretory differentiation. (*UAS-da<sup>RNAi-JF</sup>* N=15; *UAS-E(spl)-m8* N=16) (not significant p=0.7958, Mann-Whitney test).

Data information: **scale bars: 20µm**

#### 1.5.4. Sc triggers commitment to secretory fate, but delays terminal EE differentiation

Recently it was reported that EE cells do not differentiate from EBs, but from a distinct progenitor which was called pre-EE (Zeng and Hou, 2015). It has been observed that pre-

EE cells express lower levels of Pros (Chen *et al.*, 2018). Apart from Pros, pre-EE cells also express Dl and Esg (Biteau and Jasper, 2014; Zeng and Hou, 2015) and it has been argued that Esg competes with Sc for the same promoter region of *pros* (Li *et al.*, 2017a). Therefore, it is possible that in pre-EE cells, Esg/Sc competition reduces the expression of *pros* until cells terminally differentiate and *esg* is not expressed. However, in our FO experiments of *sc* over-expression, Dl+ Pros+ cells persist in their *sc* expression, but Dl+ Pros+ cells still express weaker Pros levels than fully differentiated EE cells. Therefore, the competition model might be more complicated than a simple titration or competition for binding with Esg.

Moreover, it is likely that the regulation of *pros* expression is important to ensure that one round of mitosis can occur in pre-EE cells (Chen *et al.*, 2018; Zeng *et al.*, 2010). This is supported by the Pros function to inhibit cell-cycle and induce terminal differentiation during neurogenesis (Choksi *et al.*, 2006; Lai and Doe, 2014; Li and Vaessin, 2000). While Sc induces *pros* expression, it also controls its activity and promotes expression of cell cycle genes (Chen *et al.*, 2018). Therefore, there must be other mechanisms involved in the regulation of *pros* that will allow terminal differentiation.

Interestingly, while EBs are long-lived progenitors that can remain undifferentiated for many days (Antonello *et al.*, 2015), pre-EE cells are very short lived progenitors (Zeng and Hou, 2015). However, we found that uninterrupted expression of *sc* in pre-EE cells prolongs this intermediate state, suggesting that Sc could be trapping cells into the secretory progenitor pool. This effect is increased when *da* is co-expressed with *sc*. Thus, Sc:Da dimers seem to be important to maintain this state where cells can still divide. Therefore, to terminally differentiate pre-EE cells into EE cells, Sc:Da transcriptional regulation needs to be stopped. We propose two different mechanisms for terminal secretory differentiation: 1) Emc, present in a subset of pre-EE cells binds to Sc (and Da) and inhibits transcription of Sc targets, allowing pre-EE cells to progress with the differentiation; 2) Sc induces expression of *ase* and *pros*, generating pre-EE cells, which can also divide once more, and Pros induces terminal differentiation by inhibiting the expression of *sc* and *ase*. *Ase* also promotes secretory differentiation (Bardin *et al.*, 2010). *ase* has already been shown to be expressed downstream of the other AS-C factors in the PNS after SOP specification (Domínguez and Campuzano, 1993; Jarman *et al.*, 1993). A

similar model of what we propose can be found in neuroblasts (neural stem cells), which express *ase* downstream of other proneural genes (Brand *et al.*, 1993), and Da:Ase dimers initiate differentiation by activating *pros*. Neuroblasts divide asymmetrically and produce a ganglion mother cell (GMC). *Pros* is inherited by the GMC, where it represses the expression of *ase* and initiates differentiation (Choksi *et al.*, 2006; Yasugi *et al.*, 2014).

It is still unclear how ISCs initiate EE formation. While some authors showed that there are three ways for EE differentiation (symmetric differentiation into two EE cells, asymmetric division into one ISC and one EE cell and direct differentiation) (Zeng and Hou, 2015) others stated that ISCs first go through an asymmetric division producing a daughter ISC and a pre-EE cell, and then the pre-EE cell goes through one round of division to form two EE cells (Chen *et al.*, 2018). Our results indicate that the duration of the pre-EE state and whether it divides or not could depend on how long Da:Sc dimers are active.

#### 1.5.5. Sc regulate Dl expression in ISCs and pre-EE cells

Our data indicate that Sc can regulate the expression of Dl, as when *emc* is impaired, loss of the AS-C reverts the elevated levels of *Dl* to normal (see **fig. 1.23C**). It has been described that there is an enrichment of Da:Sc binding motifs in the cis-regulatory region of *Dl* (Dutta *et al.*, 2015). Therefore, we suggest that the expression of Dl is partially regulated by low levels of *sc* present in ISCs. Low expression of *sc* in ISCs has been recently reported (Chen *et al.*, 2018). However, their finding contrast with ours, as they report that the expression of *sc* in ISCs is heterogenous and cells with higher levels of *sc* have weaker expression of *Dl* and will ultimately differentiate (Chen *et al.*, 2018). It could be possible that ISCs that are expressing higher levels of *sc* are already starting the differentiation into pre-EE cells and *Dl* expression is weaker. The expression of Dl will be maintained through the pre-EE state, but at lower levels. In fact, the regulation of *Dl* by Sc is also observed in *Drosophila* embryos (Kunisch *et al.*, 1994) and in other organisms. The *sc* homolog *Ascl1*, aside from promoting secretory differentiation in zebra fish (Flasse *et al.*, 2013), is also responsible of *Dl* expression in the intestinal epithelium (Roach *et al.*, 2013). Moreover, *Ascl1* is also required for normal Delta-like expression and lateral inhibition in mice and chicks during retinal development (Nelson and Reh, 2008; Nelson *et al.*, 2009).

Nonetheless, it could be possible that another member of the AS-C would be responsible of the regulation of *Dl*, as in our experiment we used a dull deficiency of the whole AS-C. In the future, we would like to check whether *sc* is the only gene responsible of this regulation.

We also observed that ectopic expression of *sc* in EBs promotes de-differentiation, and the newly formed ISC-like cells are able to proliferate and differentiate. Although *sc* expression in EBs was not reported by others (Chen *et al.*, 2018; Li *et al.*, 2017a), our data indicates that there could be that some *sc* expression and *Emc* expression in EBs is necessary to avoid the reversion into ISCs. Actually, in mammals, *Ascl2* maintains the stemness of the LGR5<sup>+</sup> population (van der Flier *et al.*, 2009). Although in mammals *Ascl2* has no function in secretory differentiation, this indicates that *Sc* could be driving transcription of stem genes other than *Dl* and cell proliferation genes (Chen *et al.*, 2018), which would explain why EBs de-differentiate when *Sc* is expressed. However, in *Drosophila* *Sc* is not necessary to maintain the stem pool like *Ascl2* in mammals, as knock down of *sc* does not promote absorptive differentiation or loss of *Dl* expression (Bardin *et al.*, 2010).

In *C. elegans*, a bHLH combinatorial code plays a role in the gonads to select sexual dimorphism of regulatory cells (Sallee *et al.*, 2017). All cells express the *da* homolog HLH-2 in combination with different bHLH factors depending on the cell fate. Ectopic expression of a bHLH gene expressed in a different regulatory cell or elimination of its own bHLH trigger a cell reprogramming mechanism.

#### 1.5.6. De-regulation of *da* induces cell death

*da* expression has been shown to be ubiquitous (Bardin *et al.*, 2010) which contrast with the expression of the mammalian *da* homologs *E22* and *Heb*, whose expression is restricted to the intestinal crypt (van der Flier *et al.*, 2009). We wondered if this ubiquitous expression is due to a positive feedback loop that maintains a constant expression of *da* throughout the tissue, as it has been observed previously that *da* has a positive feedback loop (Bhattacharya and Baker, 2011; Smith and Cronmiller, 2001). We checked this possibility by expressing *UAS-da:da*, and monitoring endogenous expression with *da:GFP* but autoregulation did not seem to occur. However, we cannot

rule out the possibility that another bHLH binds to Da to drive *da* transcription. This would also imply that Emc would be capable of regulating *da* transcription.

As we have observed, the regulation of *da* is very important, as low levels cannot maintain the stemness and produces cell differentiation, while high levels cause apoptosis. In the same way, excess of Emc induce differentiation and low levels produce cell death. This Emc regulation of Da to avoid a detrimental effect have also been observed in the eye and wing discs, where clones with high levels of *da* or lack of *emc* do not survive (Bhattacharya and Baker, 2011). However, in imaginal discs Da:Da drives transcription of *ex*, which is a regulator of the Hippo pathway. Therefore, the Hippo pathway is hyperactivated and antagonizes Yki, a transcriptional activator, to block growth and proliferation and promote cell death (Wang and Baker, 2015). However, our data shows that cell death occur when monomeric *da* is overexpressed, but not when we express the covalent *da-da* homodimeric form, indicating that Da might bind a different partner to induce cell death. As *da* is a ubiquitous bHLH factor (Caudy *et al.*, 1988), the extent of its function is normally determined by the expression of the binding partner. Based on our results, we hypothesize that the Da binding partner is already expressed in sufficient amounts to induce apoptosis levels, but it is tightly restricted by Emc either by direct binding, or by titrating the Da monomers. Therefore, when we overexpress *UAS-da*, the high levels of Da can bind this unknown binding partner (UBP).

However, it is intriguing that cell death seems to occur only when *Esg*<sup>+</sup> cells are over-expressing *UAS-da* (or *UAS-emc*<sup>RNAi<sub>R2</sub></sup>), as the expression in EBs only did not showed apoptosis. This suggests that the mechanism of cell death involves cellular interactions. It is possible that cell death occurs only if the cell expressing *UAS-da* is in contact with another cell expressing *UAS-da*. There might be a communication mechanism that detects that two cells in contact have high levels of the Da:UBP complex and triggers a cell death mechanism.

One possibility could be that Da:UBP dimers only promote cell death in ISCs. Some studies showed that apoptotic cells produce *eiger* (*egr*), the *Drosophila* tumor necrosis factor (TNF) homolog and promote the phenomenon called apoptosis-induced apoptosis (AiA) (Pérez-Garijo *et al.*, 2013). *Egr* is a secretable molecule that can induce cell death

to neighbouring cells through its unique receptor Grindelwald, activating the JNK pathway (Andersen *et al.*, 2015). With this model, dying ISCs would induce cell death to neighbouring EBs. Therefore, further investigation must be done in this direction, starting by checking if expression of *UAS-da* only in ISCs produce cell death or if dying cells are activating the JNK pathway.

It could also be possible that a pair of ISC and EB form a niche where cells support each other. Therefore, when one cell receives an insult that would induce cell death, it could signal to the other cell, which would send rescue signals to maintain the homeostasis of the tissue. This is observed after radioimmunotherapy in a mechanism called rescue effect (Chen *et al.*, 2011). Tumorous cells that had been treated with ionizing radiation were partnered with untreated cells. They observed that irradiated cells could signal to the next cell by two different mechanisms: cell to cell contact, with gap junctions and secretable signals through the medium. These signals induced the bystander cell to respond by sending rescue signals that repairs the irradiation damage in the treated cells, and inhibit the apoptosis (Chen *et al.*, 2011; Lam *et al.*, 2015). This mechanism have also been observed when a radiated zebrafish embryo and an untreated zebrafish embryo were together in the same medium (Choi *et al.*, 2012). We propose that this survival feedback loop could also be possible in the niche as a safeguard mechanism, and therefore, only when both cells are over-expressing *da*, cells die.

#### 1.5.7. Loss of *emc* in all the tissue as a potential tumor model

Interestingly, knock down of *emc* expression using *Myo1A*, *esg-Gal4* produced an outgrowth of ISC-like cells and a reduction of the number of ECs through the posterior midgut. This is a surprising result, considering that the equivalent knock-down of *emc* with *esg-Gal4* only leads to the cell death of most ISCs and EBs. The comparison of these phenotypes suggests that cell death induced by loss of Emc or excess of monomeric Da requires cellular interactions (see section 1.4.6).

The replacement of ECs with overproliferating ISCs observed upon depletion of *emc* in *esg*<sup>+</sup> and *Myo1A*<sup>+</sup> cells is reminiscent of the response to high cellular density in other tissues. In human colon cells, canine cells and zebrafish epithelia, to maintain the homeostasis of overcrowded tissues, there is a compensatory effect with apoptotic cells

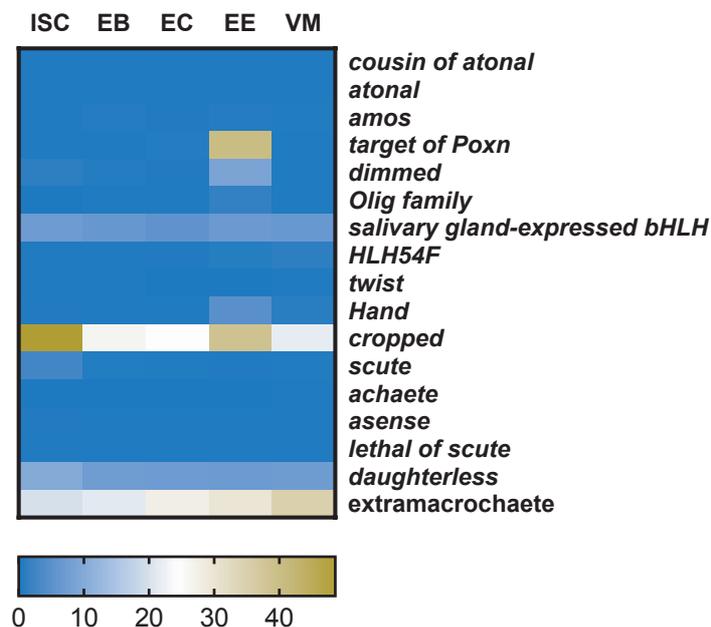
that are extruded from the tissue (Eisenhoffer *et al.*, 2012). Indeed, this effect has been described in the *Drosophila* midgut when ISCs over-proliferate due to *Notch* loss of function (Patel *et al.*, 2015). ECs that surround large *Notch*-depleted ISC clusters lose the attachment to the basal membrane and start activating JNK pathway and *yki* expression, which in turn, send mitogenic signals that favours cluster growth (Patel *et al.*, 2015). In our experiments, we can observe that there are only few surviving ECs with very large nucleus and cytoplasm, probably due to endoreplication. ECs can grow and endoreplicate as a compensatory mechanism to maintain the integrity of the tissue after the extrusion of a large proportion of EC population (Jiang *et al.*, 2009). Moreover, Yki promotes cell endoreplication by regulating the expression of *cyclin E* (Edgar and Orr-Weaver, 2001; Li *et al.*, 2017b; Shu and Deng, 2017), suggesting that when we knock-down the expression of *emc* in all the gut except EE cells, ECs might be expressing *yki* that produce these enlarged cells.

Another important feature in *emc* knocked down ISC-like cells is that they present protrusions, while ISCs in wild type flies do not. In cancer, metastasis depends on the acquisition of a motile and invasive phenotype from cancer cells in the primary tumour. Metastasis is a process in which cells from the tumour can escape the tissue and spread through the organism to form secondary tumours in distant tissues (Fidler, 2003). Therefore, cancer cells present protrusions called invadopodia that permits the metastasis (Yamaguchi *et al.*, 2005). Invadopodia are actin-rich structures that use Rho GTPases to integrate its adhesion with the extracellular membrane and degrade it in response to invasive signals (Condeelis *et al.*, 2005; Schmitz *et al.*, 2000; Yamaguchi *et al.*, 2005). It would be interesting to study if protrusions in ISC-like cells are invadopodia and therefore, if these cells could migrate to other tissues. In a recent study in the *Drosophila* hindgut, it was observed that simultaneous alteration of *ras*, *pten*, *p53*, *apc* could cause tumours with dissemination into the abdominal cavity (Bangsi *et al.*, 2016). It would be interesting to see if knocking down the expression of *emc* in the posterior midgut could induce metastasis and induce distant tumours.

## Chapter 2 Crp opposes proliferation and differentiation in ISCs

### 2.1. Introduction

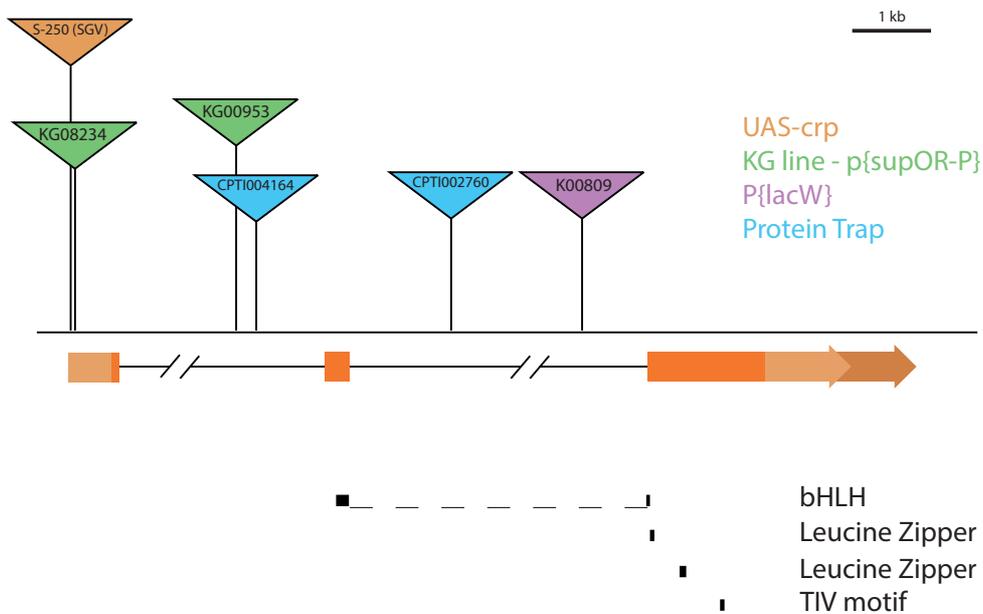
Work of others as well as our own shows that bHLH factors of class I and II are important regulators of intestinal homeostasis. This led us to consider the role of *cropped* (*crp*), a class II-like bHLH factor with comparatively high expression levels in the *Drosophila* gut (fig. 2.1). Crp has not been clearly classified in any bHLH group. In the Ledent *et al.* bHLH classification, they propose that Crp belongs to the group B, which contains bHLH factors with a leucine zipper domain immediately after the HLH motif in the C terminal end (Ledent *et al.*, 2002). However, it has been suggested that the bHLH/Zip group has polyphyletic origins, and *crp* bHLH sequence has a higher similarity with the bHLH group A (Atchley and Fitch, 1997). In this group we can find *da*, *sc* and *emc*.



**Figure 2.1. *crp* is highly expressed in the *Drosophila* midgut compared with other bHLH**

Expression levels of all bHLH class I and class II in the different cell types of the adult midgut (data from flygutseq). It can be observed that *crp* has the highest expression, specially in ISCs.

Interestingly, Crp is the only bHLH factor that contains a second leucine zipper domain and a third domain with unidentified function, called TIV motif (**fig. 2.2**) (King-Jones *et al.*, 1999). The *crp* bHLH/Zip sequence is similar to the human *Activator protein-4* (AP-4) (69%) (King-Jones *et al.*, 1999). The TIV motif is similarly conserved (65%), while the second leucine Zipper is less conserved (35%).



**Figure 2.2. *crp* locus**

The position of different insertions in *crp* mutants are shown in inverted triangles. The nature of the insertions is colour coded. The dark orange boxes indicate the coding region. *crp* has two alternative 3'UTR (arrows). The different domains are indicated on the lower part of the figure, with the bHLH domain at the end of the second exon and beginning of the third, the first leucine zipper immediately after, and the second leucine zipper and TIV motif also in the third exon.

Additional information: **scale bar, 1 kb in exons and UTR**

In *Drosophila*, *crp* is ubiquitously expressed, similarly to Da, and owes the name to its function in tracheal terminal branching (Wong *et al.* 2015). The over-expression of *crp* causes an increase in organ size, affecting salivary glands and central nervous system during larval stages (Wong *et al.*, 2015). In this study they observed that proliferation was not affected, and cells were growing by going through endoreplication.

In human cells AP-4 is implicated in proliferation, and is a downstream target of *c-Myc* (reviewed in Jung and Hermeking, 2009; Jung *et al.*, 2008). Importantly, de-regulation of AP-4 has been associated with many cancers, such as lung cancer (Hu *et al.*, 2016) or prostate cancer (Chen *et al.*, 2017), colorectal cancer (Cao *et al.*, 2009; Jackstadt *et al.*, 2013; Ma *et al.*, 2018; Xi *et al.*, 2017), where high expression of AP-4 is a marker for poor prognosis (Xinghua *et al.*, 2012; Yang *et al.*, 2018). Although its function is not fully elucidated, AP-4 promotes cell cycle progression, proliferation, epithelial-mesenchymal transition and inhibits apoptosis. Moreover, it is also involved in cisplatin resistance (Wang *et al.*, 2018).

Therefore, although Crp does not affect proliferation in the tissues that have been studied so far, the functions of hAP4 suggest that Crp might promote stemness in the adult midgut. In fact, it is known that E12 and AP4 cannot form heterodimers through the HLH domain (Hu *et al.*, 1990), but in *Drosophila* Da and Crp have been described to bind to the same E-boxes, although forming distinct complexes (King-Jones *et al.*, 1999), which suggests that Crp could regulate the transcription of some Da targets.

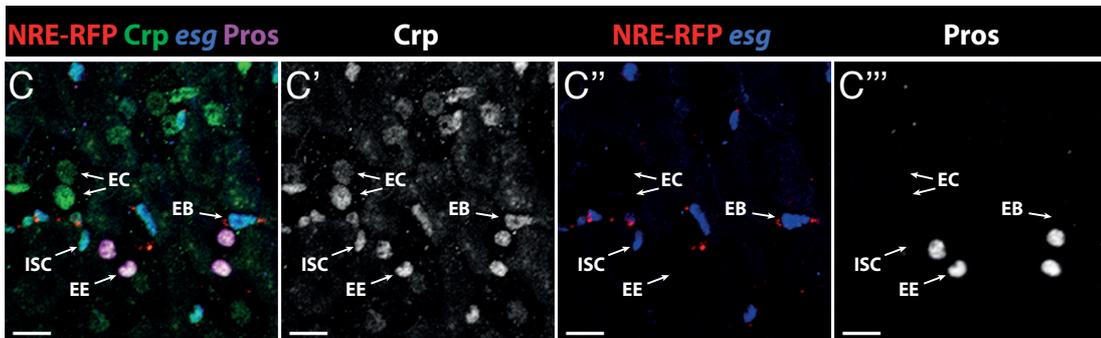
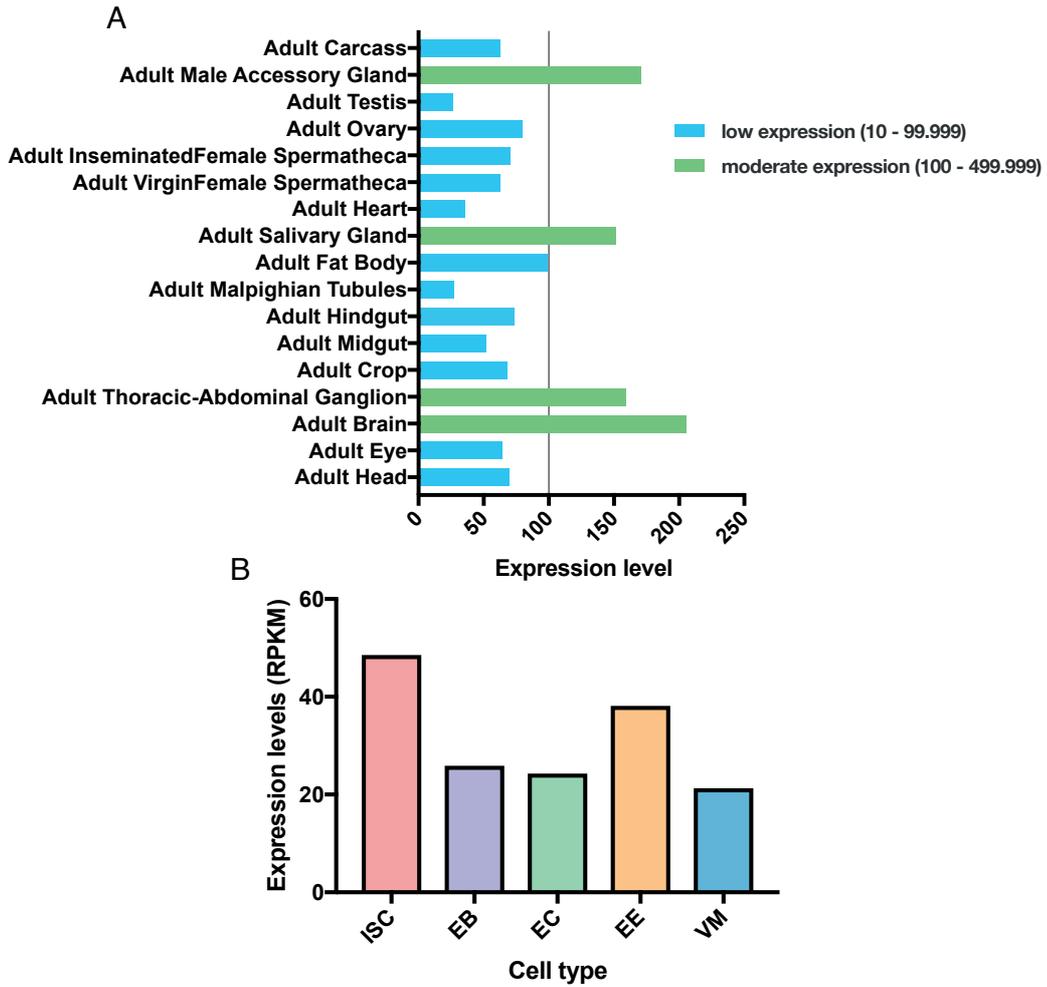
## 2.2. Aims

We wanted to determine the expression pattern of Crp and its function in the different cell types of the gut.

## 2.3. Results

### 2.3.1. Crp is expressed in all cell types of the adult posterior midgut

High-throughput approaches identify *crp* as a gene expressed in the adult *Drosophila* intestine of *Drosophila* (flyatlas.org, flygutseq.buchonlab.com; Dutta *et al.* 2015) (**fig. 2.3. A-B**). However, to obtain more information about the cell type specificity, we turned to immunohistofluorescence, and to this aim we used a CPTI protein trap line (*crp*<sup>CPTI00416</sup>) combined with *esg-lacZ*, *GBE-Su(H)-RFP*, anti-Pros staining to identify the different cell types of the gut. We observed that Crp-YFP is expressed in all cell types. However, its expression was high in ISCs, EBs and EE cells, and lower in ECs (**fig. 2.3. C**), which shows that protein abundance does not entirely correlate with mRNA, at least in EBs (**fig 2.3. C, compare with B**).



**Figure 2.3. *crp* is expressed in the *Drosophila* midgut**

**A.** Expression levels of *crp* in the different tissues of an adult fly (data from flyatlas). It can be observed that the midgut express *crp*.

**B.** Expression levels of *crp* in the different cell types of the adult midgut (data from flygutseq). *crp* is highly expressed in all cell types, specially in ISCs and EE cells.

**C.** *esg-lacZ/NRE-RFP; emc<sup>CPT1002740</sup>* flies were stained for  $\beta$ -gal and Pros. ISC (*esg*<sup>+</sup> RFP<sup>-</sup>), EB (*esg*<sup>+</sup> RFP<sup>+</sup>), EE cell (Pros<sup>+</sup>), EC (Not labelled specifically, polyploid). *crp<sup>CPT1</sup>* (green) is being expressed in all cells, although in ECs at lower levels.

Data information: **scale bars, 20 $\mu$ m**

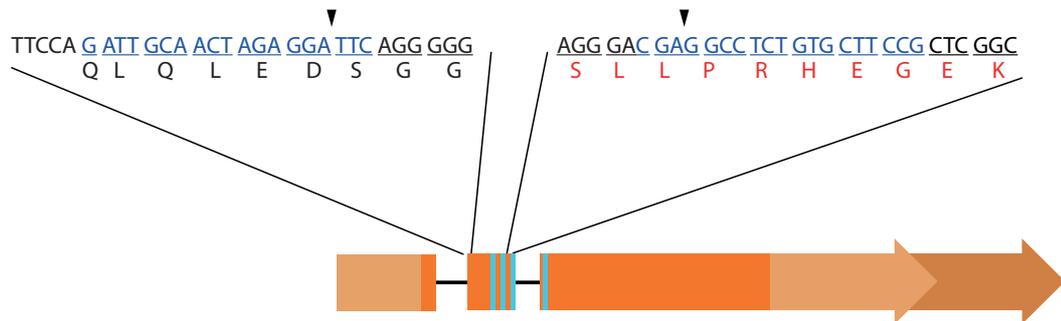
### 2.3.2. Generation of a new *crp* mutant allele

Then, we wanted to investigate the role of Crp in ISCs. However, the two available *crp* mutant lines to generate MARCM clones, *crp*<sup>KG00953</sup> and *crp*<sup>k00809</sup> (**fig. 2.2**), still showed protein in homozygous clones (see section 2.3.3). Therefore, we decided to generate null mutants using the CRISPR/Cas9 gene-editing technique. We used the pCDF4 plasmid, which has two different RNA polymerase III promoters to induce expression of two gRNAs and obtain cleavage in two genomic sites at the same time to produce a deletion. We designed one sgRNA at the 5' end of the second exon and one at the 3' end of the second exon. We decided to target these two sequences because the bHLH domain is found in the 3' end of the second exon and 5' end of the third exon (**fig. 2.4**). Both leucine zipper domains are in the third exon. With this approach, then, there could be different results, but we were mostly interested in two:

- Either of the sgRNAs creating a frameshift mutation that led to an early STOP codon disrupting the bHLH domain.
- Deletion of the whole second exon, resulting in a frameshift and an early STOP codon, or an in-frame deletion of most of the bHLH domain. With this outcome and without a frameshift or an early STOP codon, both leucine zipper domains would be intact.

We recovered two mutants that did not complement the lethality of two large deficiencies known to uncover the *crp* locus, *Df(2L)BSC278* and *Df(2L)r10*, as well as the recessive lethal alleles *crp*<sup>KG00953</sup> and *crp*<sup>k00809</sup> (P. Lall, M. Iztueta-Inchausti and J. de Navascués, unpublished). One of them, *crp*<sup>pl21.4</sup>, was found to have a small indel at the cleavage site close to the 3' end of the second exon, with no alteration at the cleavage site defined by the gRNA binding to the 5' end of the exon (P Lall and J de Navascués, unpublished). Indeed, it has already been reported that using this double target method could lead to a higher expression of one of the sgRNA, as the U6:3 promoter (used for the 3' sgRNA) has higher rates of mutagenesis than the U6:1 promoter (used for the 5' sgRNA) (Port and Bullock, 2016; Port *et al.*, 2014). However, in our case. The mutation in *crp*<sup>pl21.4</sup> caused a frameshift that created an early STOP codon in the ORF of Crp in the middle of the bHLH

domains and before the leucine zippers domains, which suggests that this allele is likely a molecular null.

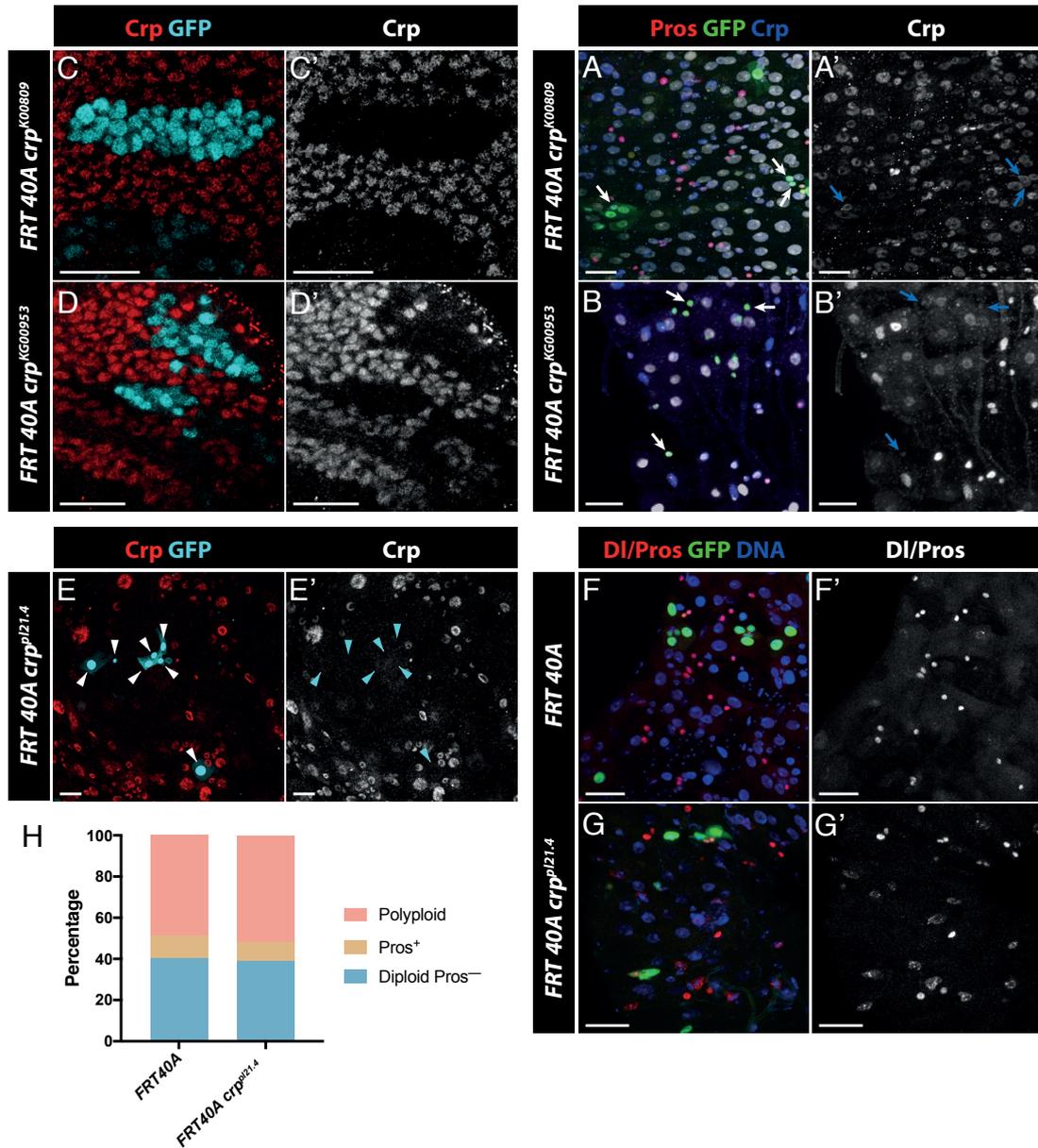


**Figure 2.4. Double *crp* mutagenesis experimental design**

Approximate position of the two targeted sequences for CRISPR mutagenesis. In cyan is indicated the bHLH domain of *crp*. One sgRNA was design at the 5' end of the second exon in sense with the transcription direction. The aminoacids encoded are found below each triplet. A second sgRNA was designed close to the 3' end of the second exon in the middle of the bHLH domain. The sgRNA was designed anti-sense with the transcription direction. Below it is found in red the aminoacids encoded by complementary (in sense) strand. The arrow heads indicate the Cas9 cleavage site.

### 2.3.3. *crp* expression arrests proliferation and differentiation

When we wanted to induce *crp* mutant MARCM clones with two null mutants, *crp*<sup>KG00953</sup> and *crp*<sup>k00809</sup>, we realized that while in the larval wing disc, Crp was not detected (**fig. 2.5A,B**), in the adult midgut some cells still expressed *crp* (**fig. 2.5C,D**), meaning that they are not null mutants as described previously (Wong *et al.*, 2015). As these *crp* alleles are insertional, it is possible that they interfere with communication between the *crp* promoter and its wing disc enhancers, but not with enhancers of the intestine, making them organ-specific mutants, rather than *bona fide* nulls. Therefore, as explained in section 2.3.2, we generated a new mutant, *crp*<sup>pl21.4</sup>. The clone induction using this mutant showed that cells were not expressing *crp* anymore (**fig. 2.5E**). However, the differentiation in the clones was not affected compared to WT clones (**fig 2.5 F,G**, quantification in **2.5H**).



**Figure 2.5. *crp* loss of function guts can differentiate normally**

**A,B.** Induction of (A) *crp*<sup>K00809</sup> and (B) *crp*<sup>KG00953</sup> MARCM clones in the wing disk. All cells in the clone do not express *crp*.

**C-D.** Cells in (C) *crp*<sup>K00809</sup> and (D) *crp*<sup>KG00953</sup> MARCM clones still show signal (arrows) using Crp antibody in the adult midgut.

**E.** *crp*<sup>p121.4</sup> MARCM clones did not express *crp* in the clonal in homozygous cells.

**F-G.** ISCs in *FRT40A crp*<sup>p121.4</sup> mutant clones can proliferate and differentiate into all cell types, similar as (F) *FRT40A* control ISCs.

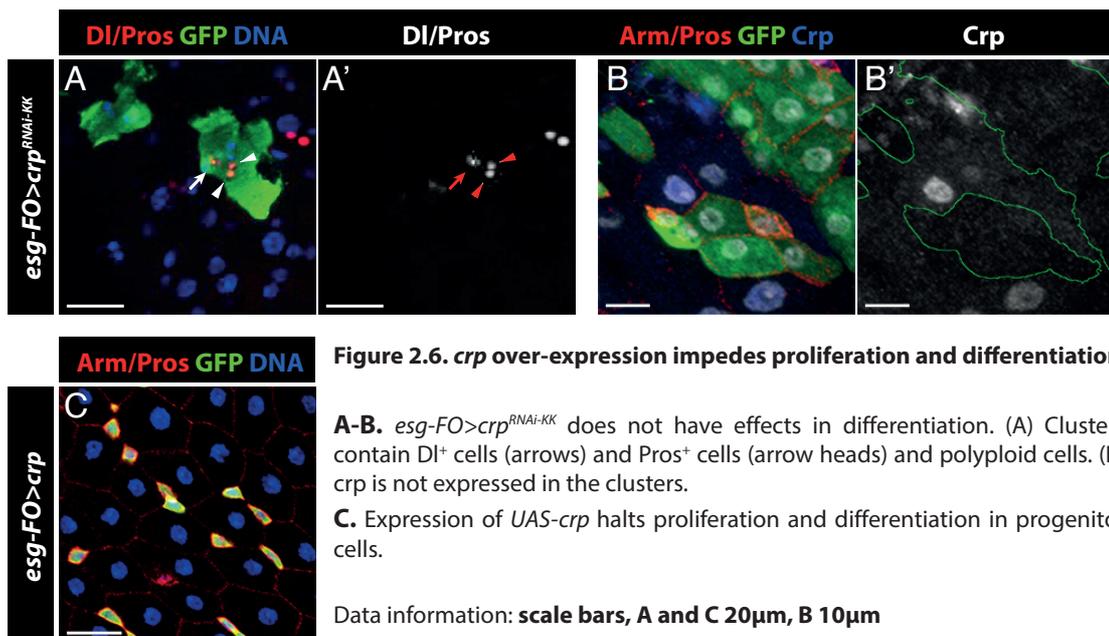
**H.** Quantification of the different cell types in the clonal area of *crp*<sup>p121.4</sup> clones. This quantification corresponds to preliminary data and the control *FRT40A* flies were from a different batch than the mutant *FRT40A crp*<sup>p121.4</sup> flies.

N = 1153 cells/86 clones for *FRT40A* clones, N = 306 cells/130 clones for *FRT40A crp*<sup>p121.4</sup>.

Data information: **scale bars, 20µm**

We also knocked-down *crp* expressing *UAS-crp<sup>RNAi<sub>KK108184</sub></sup>* with the *esg-FO* system for 10 days (**fig. 2.6A**). We could not observe any difference in the cells labelled with GFP compared with control flies, as the ISC/EB compartment seemed to generate new cells at normal rates and proportions. We then tested the efficiency of the *crp* RNAi line with a Crp antibody, and we could clearly see that there was not any signal in the GFP<sup>+</sup> area, while there was expression in the rest of the tissue (**fig. 2.6B**). This result suggests that Crp is dispensable for the maintenance of the tissue homeostasis.

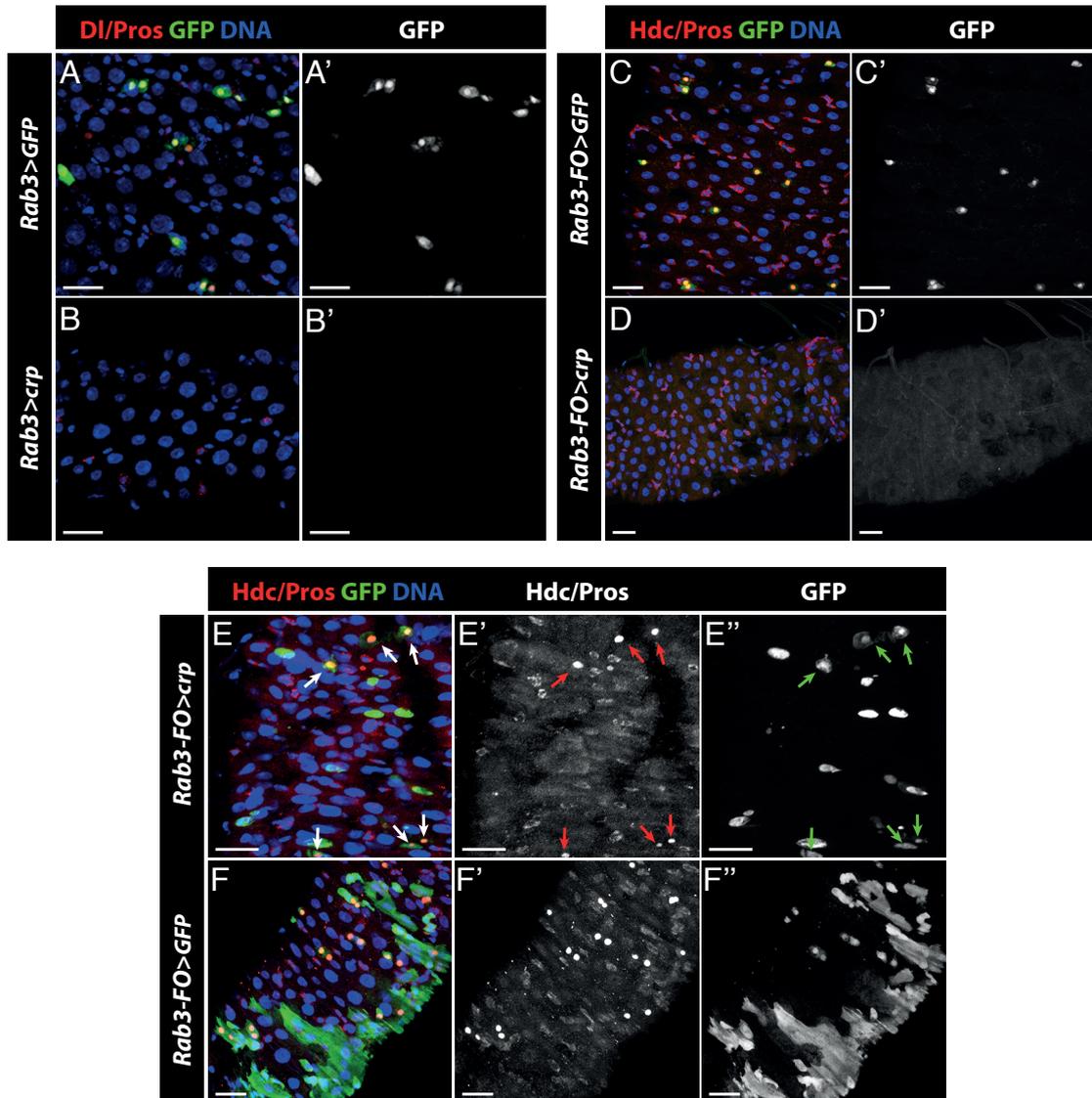
It has been previously described that Crp and Da compete for binding to the same E-boxes, and it was suggested that Da was preferentially bound to the DNA (King-Jones *et al.*, 1999). In this situation, only high levels of *crp* would have a function, and we might not find these levels of expression in homeostatic tissue. To check this, we over-expressed *UAS-crp* with the *esg-FO* system. Interestingly, when *UAS-crp* was expressed in the gut for 7 days, we could not observe any kind of differentiation (**fig. 2.6C**). More important, progenitor cells remain as single cells or in pairs. This suggests that *crp* over-expression arrested both differentiation and proliferation of ISCs. After one week, there was no differentiation in the whole gut. GFP<sup>+</sup> cells remained alone or in pairs, indicating that there had been no cell division. This suggests that Crp is not needed for differentiation nor maintenance of the stem cells.



#### 2.3.4. Overexpression of *crp* promotes EE cell death/elimination

Since the over-expression of *crp* in progenitor cells produced a complete arrest of proliferation and differentiation, we thought that this could be a mechanism to keep cells quiescent without interfering with their identity. Therefore, we decided to explore the effect of Crp in EE cells, as this is a quiescent population that expresses *crp*. To this end we used the EE-specific driver *Rab3-Gal4*. Rab3 is a small GTPase which is necessary for the vesicle fusion with its target (Bhuin and Roy, 2014; Fischer von Mollard *et al.*, 1990) and therefore, required in all secretory cells, including EE cells. Using this driver, we over-expressed *UAS-crp* and *UAS-GFP* and stained for Pros. Interestingly, the result was a complete loss of EEs (scored as Pros<sup>+</sup> and GFP<sup>+</sup> cells; **fig. 2.7A-B**). This could indicate that EE cells were de-differentiating. To check this possibility, we performed a lineage tracing experiment combining the *Rab3-Gal4* driver with the flipOut system (*Rab3-FO*). After 7 days of *UAS-crp* expression in the pool of secretory cells, almost no cells were labelled with GFP in the posterior midgut (**fig. 2.7C-D**), indicating that EE cells are dying rather than changing their fate. However, in the gastric region we could observe some Pros<sup>+</sup> and even some Hdc<sup>+</sup> cells labelled with GFP, indicating that some de-differentiation might be possible (**fig. 2.7E**).

Interestingly, in control *Rab3-F/O* flies we could observe GFP<sup>+</sup> cells that were Pros<sup>-</sup>. Moreover, we also observed large areas of tissue that were completely replaced with GFP<sup>+</sup> cells (**fig. 2.7F**). This indicates that the *Rab3-Gal4* lineage tracing targeted cells with stem properties. While it is possible that either the *UAS-flp* or the *Rab3-Gal4* had leaky expression in ISCs, we cannot rule out the possibility that EE cells are reverting to ISCs, and this might be a mechanism to have a reservoir of cells in case ISCs are compromised.



**Figure 2.7. Crp induces cell death in EE cells**

- A.** *Rab3*> *UAS-GFP* drives GFP expression in EE cells (Pros<sup>+</sup>).
- B.** *Rab3*> *UAS-crp* shows a total absence of GFP<sup>+</sup> or Pros<sup>+</sup> cells.
- C.** *Rab3-FO*> *UAS-GFP* shows that the majority of EE cells remain as differentiated secretory cells.
- D-E.** *Rab3-FO*> *UAS-crp* induce cell death in the majority of the secretory cells. (E) However, in the gastric area, cell death is less penetrant.
- F.** Lineage tracing from EE cells can produce large areas of GFP<sup>+</sup> cells.

Data information: **scale bars, 20μm**

## 2.4. Discussion

Herein, we have observed that the bHLH/Zip factor *crp* is expressed ubiquitously in the gut, although ECs express it at lower levels than all other cell types (Fig. 2.3C). We have found that loss of *crp* has no obvious effect on ISCs, while gain of function has different effects depending on cell type. In progenitor cells, there is an apparent arrest of proliferation and differentiation. However, the death of EE cells produced by *crp* over-expression opens up the possibility that cells are proliferating and differentiating normally, but differentiated cells immediately die. Nonetheless, we did not explore further this possibility as we could not see any sign of cell death or alteration in the tissue.

According to our observations, Crp and Da have no functional relationship. Previous studies reported that full-length human AP4 could not bind Da homologs E12, as the leucine zipper domain impedes HLH dimerization, and most likely cannot form dimers with any other bHLH through the HLH domain (Hu *et al.*, 1990). Moreover, Id1 and Id2 cannot bind with AP4 either (Sun *et al.*, 1991). Thus, the function of Crp is not regulated by Emc. However, Crp and Da has been both implicated in the regulation of the expression of the salivary gland secretion protein genes (*sgs*), but forming distinct complexes with unknown partners that bind the same E boxes (King-Jones *et al.*, 1999), although it has not been investigated if Da and Crp have redundant or antagonistic functions. In fact, it was found that Da complexes were preferentially bound to *sgs* promoter region. As hP4 can be both an activator or an inhibitor of gene expression (Kim *et al.*, 2006; Mermod *et al.*, 1988), it would be plausible that Da (and its bHLH partner) and Crp competes for DNA binding with opposing effects. This competition might also be found in our system, where in wild type, unchallenged conditions, Crp is not in sufficient amounts to compete with Da dimers, but when it is highly expressed, it can displace some Da:bHLH dimers. Moreover, Crp is unable to dimerize with other bHLH factors, but it homodimerizes with its leucine zipper domains (Hu *et al.*, 1990), and the homodimers bind the same sequences as Da (King-Jones *et al.*, 1999). However, opposing completely the function of Da would result in EC differentiation and we have not observed anything like that. We surmise that Crp homodimers could bind only to specific E-boxes depending on their sequence. Therefore, Crp:Crp dimers could be binding to Da:Sc E-boxes that regulate cell cycle or secretory differentiation genes,

arresting proliferation in the gut and EE differentiation, but not with Da:Da target sequences.

Our loss of function analysis and knock down experiments indicate that in homeostatic conditions, Crp function is dispensable, and only when it is highly expressed, it drives progenitor cells into a quiescent state. It is possible that Crp is only required in conditions of acute injury or infection, or in aged flies, which we have not explored.

It is interesting that the expression of *crp* in EE cells produce the death of this cells, since AP4 has been shown to silence the expression of caspases in mammals (Tsujiimoto *et al.*, 2005). Nevertheless, our observations suggest that Crp has, if anything, the opposite effect, stimulating cell death.

By contrast with the effects of its *Drosophila* homolog, hAP4 is a downstream target of cMyc (Jung *et al.*, 2008), which in turn is positively regulated by the Wnt pathway (He *et al.*, 1998). In the *Apc<sup>min</sup>* model of colorectal cancer, AP4 knockout mice have increased average survival (Jaeckel *et al.*, 2018). Moreover, Jaeckel *et al.* showed that these mice also manifest a reduced formation of adenomas, less cancer stem cells and a down-regulation of Notch and Wnt signalling pathways, indicating that AP4 promotes stemness and proliferation in the mammalian gut. However, in *Drosophila*, Crp seems to have the opposite effect to dMyc, as dMyc induces proliferation downstream of Wg signalling during midgut regeneration (Cordero *et al.*, 2012b; Ren *et al.*, 2013), while Crp inhibits proliferation. Therefore, it could be possible that in this context *crp* expression is not downstream of dMyc.

It is also interesting to note that in wild type flies, lineage tracing on differentiated EE cells suggested that secretory cells could de-differentiate into ISCs that could proliferate and differentiate as normal (**Fig. 2.7F**). This result opens the possibility that in very specific moments, where the integrity of the stem pool is at risk, EE cells could act as a reservoir for ISCs. This finding has not been reported before in *Drosophila*, although a study tried to force EE de-differentiation by ablating all progenitor cells through the knock-down of *prickle* in progenitor cells (Lu and Li, 2015). However, these authors could not see any progenitor cell after 60 days. Another possibility could be that *Rab3* is

expressed in rare occasions in ISCs, although then we would expect to find a larger number of ISCs expressing GFP.

In mammals, it has been described that some secretory progenitors could reverse into a stem state in stress conditions (Buczacki *et al.*, 2013). Moreover, recent reports show that small populations of completely differentiated EE cells can de-differentiate into ISCs and maintain the tissue homeostasis (Gross *et al.*, 2015; Yan *et al.*, 2017). Although we need additional controls to rule out that the reversion is an artefact, our results come to confirm that the EE fate is surprisingly close to that of the ISCs, and we propose that *Drosophila* could be a good model organism to study the de-differentiation of EE cells as a reservoir for ISCs.

## Conclusions

1. Da homodimers prevent differentiation in the adult fly intestine.
2. Emc promotes absorptive differentiation, probably through sequestration of Da.
3. Persistent expression of Sc prolongs the life of the transient pre-EE cells and increases their proliferation
4. Sc is active in ISCs, increasing *Dl* expression.
5. Sc can induce de-differentiation of EBs.
6. The committed state of EBs needs to be actively maintained, and at least part of this maintenance is performed by Emc, perhaps inhibiting basal levels of Sc.
7. Excess of Da in progenitor cells induces cell death. This function of Da is independent of Da homodimerisation.
8. *crp* is not required for the regulation of self-renewal nor differentiation in the adult fly midgut.
9. The expression of *crp* must be regulated, as its excess interferes with ISC proliferation and differentiation and survival of EE cells.

## Appendix 1: Cell counter

```
1. import sys
2. import xlrd
3. import os
4. import csv
5. import numpy as np
6. from scipy.stats import chisquare
7. from pandas.tools import plotting
8. import networkx as nx
9. from networkx.algorithms.components.connected import connected_components
10. import collections
11. from pyface.api import DirectoryDialog, OK
12. os.chdir('/Users/aleix/Documents/Python/Fastidious_libraries/')
13. import pointpicker_to_dist as ppd
14. import tiffiffile as tf
15. import pandas as pd
16. os.chdir('/Users/aleix/Desktop/Counting/')
17.
18. #-----
19. --
20. pixel_size = 1
21. pixels = 20
22. THR = 1
23.
24. #-----
25. --
26. def getFolder():
27.     '''GETFOLDER creates a GUI to obtain the path of the selected folder'''
28.     # GUI for getting a sample filepath
29.     dialog = DirectoryDialog(action="open", default_path='/Volumes/jd467/DATA/
30.     C O N F O C A L/ZEISS 700')
31.     dialog.open()
32.     if dialog.return_code == OK:
33.         folderpath=dialog.path
34.     return folderpath
35. #-----
36. --
37. # I first ask the number of conditions I am processing and then create a loop
38. # to create a variable
39. num_samples = input("How many samples to analyse?\n")
40.
41. os.chdir("/Users/aleix/Desktop/Counting")
42. samples = list()
43. folderpath = list()
44.
45. for f in range(0, int(num_samples)) :
46.     samples.append(input('Sample name: '))
47.     folderpath.append(getFolder())
48.
49. for f in range(0, int(num_samples)) :
50.     exec("file_list%s = os.listdir(folderpath[f])" % (f))
51.     exec("file_list%s = [x for x in file_list%s if x.endswith('.txt')]" % (f,f
52.     )) # keep only files ending with.txt
53.
54. ....
```

```

55. fpath = '/Users/aleix/Documents/Python/Analysis/'
56. # Everytime I run the program for a different batch it need to be copy here
57. batch = 'trials'
58.
59. click = os.path.join(fpath,batch,'Clicking_images')
60. click_tail = '_click.tif'
61.
62. pp_path = os.path.join(fpath,batch,'pointpicker')
63. pp_tail = '_pp.txt'
64. '''
65.
66.
67.
68. #-----
--
69.
70.
71. def pp_positions(source) :
72.     # Reading table and mapping the points
73.     x,y = np.loadtxt(os.path.join(folderpath,fn), skiprows=1, usecols=(1, 2),
unpack=True, dtype='uint16') # Separate columns
74.     XYdata = np.loadtxt(os.path.join(folderpath,fn), skiprows=1, usecols=(1, 2
)) #All together
75.     ppImage = np.zeros(im_proj[:, :,0].shape,dtype = 'float')
76.     ppImage[y, x] = 1
77.
78. def pp_numbers(source, sample) :
79.     countings = pd.read_table(source, usecols=['Type', 'X', 'Y'])
80.     ISC_GFP = (sum(countings.Type == 1))
81.     EB_GFP = (sum(countings.Type == 2))
82.     EE_GFP = (sum(countings.Type == 5))
83.     EC_GFP = (sum(countings.Type == 6))
84.     GFP_num = ISC_GFP + EB_GFP + EE_GFP + EC_GFP
85.     total_num = float(len(countings.index))
86.     t_ISC_GFP = (sum(countings.Type == 1)*100/total_num)
87.     t_EB_GFP = (sum(countings.Type == 2)*100/total_num)
88.     t_EE = (sum(countings.Type == 3)*100/total_num)
89.     t_EC = (sum(countings.Type == 4)*100/total_num)
90.     t_EE_GFP = (sum(countings.Type == 5)*100/total_num)
91.     t_EC_GFP = (sum(countings.Type == 6)*100/total_num)
92.     t_ISC = (sum(countings.Type == 7)*100/total_num)
93.     t_EB = (sum(countings.Type == 8)*100/total_num)
94.     exp = [sample]
95.     g_ISC_GFP = ISC_GFP*100/GFP_num
96.     g_EB_GFP = EB_GFP*100/GFP_num
97.     g_EE_GFP = EE_GFP*100/GFP_num
98.     g_EC_GFP = EC_GFP*100/GFP_num
99.
100.
101.     totaldata = np.hstack([exp,t_ISC_GFP,t_EB_GFP,t_EE_GFP,t_EC_GFP,t_I
SC,t_EB,t_EE,t_EC])
102.     GFPdata = np.hstack([exp,g_ISC_GFP,g_EB_GFP,g_EE_GFP,g_EC_GFP])
103.
104.     return totaldata, GFPdata
105.
106. def delaunay_distances(filename,pixel_size,THR):
107.
108.     from scipy.spatial import Delaunay
109.
110.     # read data from pointpicker output field
111.     XYdata = np.loadtxt(filename, skiprows=1, usecols=(2, 3))
112.
113.
114.     # obtain triangulation, get sides of triangles
115.     D = Delaunay(XYdata)
116.     V = D.vertices

```

```

117.         Vsorted = []
118.         A = V[:, 0]
119.         B = V[:, 1]
120.         C = V[:, 2]
121.         Vsorted.extend(zip(A, B))
122.         Vsorted.extend(zip(B, C))
123.         Vsorted.extend(zip(C, A))
124.         Vsorted = np.sort(Vsorted,axis=1)
125.
126.         # uniquify sides
127.         unique_idc = Vsorted.view(np.dtype( (np.void, Vsorted.dtype.itemsiz
e * Vsorted.shape[1]) ) )
128.         _, idx = np.unique(unique_idc, return_index=True)
129.         V_NR = Vsorted[idx]
130.
131.         # remove sides too close to opposing vertex (convex hull)
132.         excluded = []
133.         for side in D.convex_hull:
134.             for vertex in D.vertices:
135.                 if len(np.setdiff1d(vertex, side)) == 1:
136.                     A = np.vstack([D.points[np.setdiff1d(vertex, side)],ppd
.side_centr(D.points[side])])
137.                     dist_vtx2hull = ppd.dist_2p(A)
138.                     if dist_vtx2hull < ppd.dist_2p(D.points[side])/THR:
139.                         excluded.append(side)
140.
141.         relevant_sides = ppd.setdiff_rows(V_NR,np.sort(np.array(excluded),a
xis=1))
142.
143.         # get distances
144.         coord = D.points[relevant_sides]
145.         distances = [ppd.dist_2p(x) for x in coord]*pixel_size
146.         relevant_sides = relevant_sides[(np.asarray(distances) < 33), :]
147.
148.         .....
149.         # export to csv files
150.         import csv
151.         fl = open(filename.replace(".txt", ".csv"), 'w')
152.         writer = csv.writer(fl)
153.         for value in distances:
154.             writer.writerow([value])
155.         fl.close()
156.         '''
157.
158.         return XYdata, relevant_sides
159.
160.
161.     def connected_comp(l):
162.         # Algorithm to find connected components:
163.         # 1. take first set A from list
164.         # 2. for each other set B in the list do if B has common element(s)
with A join B into A; remove B from list
165.         # 3. repeat 2. until no more overlap with A
166.         # 4. put A into outpup
167.         # 5. repeat 1. with rest of list
168.         out = []
169.         while len(l)>0:
170.             first, rest = l[0], l[1:]
171.             first = set(first)
172.
173.             lf = -1
174.             while len(first)>lf:
175.                 lf = len(first)
176.
177.             rest2 = []
178.             for r in rest:

```

```

179.             if len(first.intersection(set(r)))>0:
180.                 first |= set(r)
181.             else:
182.                 rest2.append(r)
183.                 rest = rest2
184.
185.                 out.append(first)
186.                 l = rest
187.
188.         return out
189.
190.     def clone_population(TXYdata, single, multiple, sample):
191.         # identifies the size cluster and its composition
192.         type_clones = [TXYdata[multiple[x],0] for x in range(len(multiple))
193. ]
194.         if single.any():
195.             sin_type_clones = TXYdata[single,0]
196.             single_arrays = np.split(sin_type_clones, len(sin_type_clones))
197.
198.             complete_types = single_arrays + type_clones
199.         else:
200.             complete_types = type_clones
201.             complete_types = [list(complete_types[x]) for x in range(len(comple
202. te_types))]
203.
204.         cols = ['Size', 'ISC_GFP', 'EB_GFP', 'EE_GFP', 'EC_GFP']
205.         clone_composition = pd.DataFrame(columns=cols, dtype=float)
206.
207.         for f in range(len(complete_types)) :
208.             size = len(complete_types[f])
209.             isc = complete_types[f].count(1) | complete_types[f].count(1)
210.             eb = complete_types[f].count(2) | complete_types[f].count(2)
211.             ee = complete_types[f].count(5) | complete_types[f].count(5)
212.             ec = complete_types[f].count(6) | complete_types[f].count(6)
213.             data = np.hstack([size,isc,eb,ee,ec])
214.             data = pd.DataFrame([data, ], columns=list(cols))
215.             clone_composition = clone_composition.append(data, ignore_index
216. =True)
217.
218.         return clone_composition
219.
220.     def visualize_delaunay(XYdata, relevant_sides, clone_sides, complete_ve
221. rt, image_name):
222.
223.         import matplotlib
224.         import matplotlib.pyplot as plt
225.
226.         P = XYdata
227.
228.         X,Y = P[:,0],P[:,1]
229.
230.         fig = plt.figure(figsize=(20,20))
231.         axes = plt.subplot(1,1,1)
232.         plt.axis('off')
233.
234.         im = plt.imread(image_name)
235.         plt.imshow(im)
236.
237.         plt.scatter(X, Y, marker='o', color='c')
238.         plt.scatter(X[complete_vert], Y[complete_vert], marker='o', color='
239. r')
240.         plt.axis([0,512,0,512])

```

```

239.
240.
241.     lines = matplotlib.collections.LineCollection(P[relevant_sides], co
        lor='c')
242.     plt.gca().add_collection(lines)
243.     gfp_lines = matplotlib.collections.LineCollection(P[clone_sides], c
        olor='r')
244.     plt.gca().add_collection(gfp_lines)
245.
246.     plt.axis([0,512,0,512])
247.     plt.show()
248.     #plt.savefig(image_name+'.png', dpi=144, format='png')
249.
250.
251.     def GFP_cells(filename, sample, THR, image_name) :
252.         # It finds all GFP vertices and will find if they are connected wit
        h other GFP vertices
253.         import collections
254.
255.         # ISC-GFP are 1, EB-GFP are 2, EE-GFP are 5 and EC-GFP are 6
256.         XYdata, relevant_sides = delaunay_distances(filename, pixel_size, THR
        )
257.         TXYdata = np.loadtxt(filename, skiprows=1, usecols=(0, 2, 3))
258.         prg_GFP_A = TXYdata[relevant_sides[:,0],0] < 3
259.         prg_GFP_B = TXYdata[relevant_sides[:,1],0] < 3
260.         EE_GFP_A = TXYdata[relevant_sides[:,0],0] == 5
261.         EE_GFP_B = TXYdata[relevant_sides[:,1],0] == 5
262.         EC_GFP_A = TXYdata[relevant_sides[:,0],0] == 6
263.         EC_GFP_B = TXYdata[relevant_sides[:,1],0] == 6
264.
265.         # Generate boolean arrays with all GFP vertices
266.         GFP_A = prg_GFP_A | EE_GFP_A | EC_GFP_A
267.         GFP_B = prg_GFP_B | EE_GFP_B | EC_GFP_B
268.
269.         # finding vertices that connect two GFP cells
270.         GFP_sides = np.bitwise_and(GFP_A,GFP_B)
271.         GFP_sides = np.resize(GFP_sides,[GFP_sides.shape[0],1])
272.
273.         np.hstack([relevant_sides,GFP_sides])
274.
275.         clone_sides = relevant_sides[GFP_sides[:,0],:]
276.
277.         # Obtaining single cell clones.
278.         mult_GFP = np.unique(clone_sides)
279.
280.         GFP_singA = np.resize(GFP_A,[GFP_A.shape[0],1])
281.         GFP_singB = np.resize(GFP_B,[GFP_B.shape[0],1])
282.
283.         GFP_vertA = np.unique(relevant_sides[GFP_singA[:,0],0])
284.         GFP_vertB = np.unique(relevant_sides[GFP_singB[:,0],1])
285.
286.         all_vert = [mult_GFP , GFP_vertA , GFP_vertB]
287.         complete_vert = np.array(list(set(x for l in all_vert for x in l)))
288.         result = np.array(list(set(x for l in all_vert for x in l) - set(mu
        lt_GFP)))
289.         sin_list = [1 if x>=0 else x for x in result]
290.
291.         clones = connected_comp(clone_sides.tolist())
292.         clones = [list(clones[x]) for x in range(len(clones))]
293.         clone_numbers = [len(x) for x in clones]
294.         clone_numbers = sin_list + clone_numbers
295.
296.         clone_composition = clone_population(TXYdata, result, clones, sampl
        e)

```

```

297.         #visualize_delaunay(XYdata, relevant_sides, clone_sides, complete_v
ert, image_name)
298.
299.         return result, clone_numbers, clones, clone_composition
300.
301.         # Prepare the dataframes to fill
302.
303.         cols_t = ['Sample', 'ISC_GFP', 'EB_GFP', 'EE_GFP', 'EC_GFP', 'ISC', 'EB
', 'EE', 'EC']
304.         cols_cluster = ['Sample', 'ISC_GFP', 'EB_GFP', 'EE_GFP', 'EC_GFP']
305.         cols_composition = ['Size', 'ISC_GFP', 'EB_GFP', 'EE_GFP', 'EC_GFP']
306.         cols_clones = ['clone_size', 'number_clones', 'Fraction', 'inv_cumulati
ve', 'Samples']
307.         data_accum = pd.DataFrame(columns=cols_t, dtype=float)
308.         GFP_data_accum = pd.DataFrame(columns=cols_cluster, dtype=float)
309.         clone_composition_sum = pd.DataFrame(columns=cols_composition, dtype=fl
oat)
310.         cl_df_global = pd.DataFrame(columns=cols_clones, dtype=float)
311.         cl_ISC = pd.DataFrame(columns=cols_clones, dtype=float)
312.         cl_EB = pd.DataFrame(columns=cols_clones, dtype=float)
313.
314.         # Create lists with the multiple data frames created for each condition
315.
316.
317.
318.         for k in range(0, int(num_samples)) :
319.             exec("c%s_clone_composition_append = pd.DataFrame(columns=cols_comp
osition, dtype=float)" % (k))
320.
321.             file_list = os.listdir(folderpath[k])
322.             file_list = [x for x in file_list if x.endswith('.txt')]
323.             c_clone_composition_append = pd.DataFrame(columns=cols_composition,
dtype=float)
324.
325.
326.             acum_clone_numbers = []
327.             for f in range(0, len(file_list)) :
328.                 print ("processing image %d"%(f))
329.                 data, GFPdata = pp_numbers(os.path.join(folderpath[k], file_list
[f]), samples[k])
330.                 data = pd.DataFrame([data, ], columns=list(cols_t))
331.                 GFPdata = pd.DataFrame([GFPdata, ], columns=list(cols_cluster))
332.
333.                 data_accum = data_accum.append(data, ignore_index=True)
334.                 GFP_data_accum = GFP_data_accum.append(GFPdata, ignore_index=Tr
ue)
335.
336.                 sin_cl, clone_numbers, clones, clone_composition = GFP_cells(os
.path.join(folderpath[k], file_list[f]), samples[k], THR, os.path.join(folderpa
th[k], file_list[f]))
337.                 acum_clone_numbers = acum_clone_numbers + clone_numbers
338.
339.                 c_clone_composition_append = c_clone_composition_append.append(
clone_composition, ignore_index=True)
340.
341.                 cols = ['clone_size', 'number_clones']
342.                 counter = collections.Counter(acum_clone_numbers)
343.                 cl_df = pd.DataFrame.from_dict(counter, orient='index').reset_index
()
344.                 cl_df.columns = cols
345.
346.                 initial_cumulative = pd.DataFrame([[0,0]], columns=cols, dtype=floa
t)
347.                 cl_df = cl_df.append(initial_cumulative)

```

```

348.         cl_df['Fraction'] = cl_df.number_clones/sum(cl_df.number_clones)
349.         cl_df = cl_df.sort_values('clone_size')
350.         cl_df['inv_cumulative'] = 1-cl_df.Fraction.cumsum()
351.         cl_df['Samples'] = samples[k]
352.
353.
354.         c_clone_composition_append[['Size', 'ISC_GFP', 'EB_GFP', 'EE_GFP',
'EC_GFP']] = c_clone_composition_append[['Size', 'ISC_GFP', 'EB_GFP', 'EE_GFP',
, 'EC_GFP']].astype(float)
355.         c_clone_composition_sum = pd.DataFrame([c_clone_composition_append.
sum()], dtype=float)
356.         c_clone_composition_sum['Condition'] = samples[k]
357.
358.         cl_df_global = cl_df_global.append(cl_df, ignore_index=True)
359.         clone_composition_sum = clone_composition_sum.append(c_clone_compos
ition_sum, ignore_index=True)
360.
361.         # Counting quantity of ISC and EBs in clones
362.
363.         c1_ISC_counter = collections.Counter(c_clone_composition_append.ISC
_GFP)
364.         c1_ISC1 = pd.DataFrame.from_dict(c1_ISC_counter, orient='index').re
set_index()
365.         c1_ISC1.columns = cols
366.         c1_ISC1['Fraction'] = c1_ISC1.number_clones/sum(c1_ISC1.number_clon
es)
367.         c1_ISC1['inv_cumulative'] = 1-c1_ISC1.Fraction.cumsum()
368.         c1_ISC1['Samples'] = samples[k]
369.         c1_ISC = c1_ISC.append(c1_ISC1, ignore_index=True)
370.
371.         c1_EB_counter = collections.Counter(c_clone_composition_append.EB_G
FP)
372.         c1_EB1 = pd.DataFrame.from_dict(c1_EB_counter, orient='index').rese
t_index()
373.         c1_EB1.columns = cols
374.         c1_EB1['Fraction'] = c1_EB1.number_clones/sum(c1_EB1.number_clones)
375.
376.         c1_EB1['inv_cumulative'] = 1-c1_EB1.Fraction.cumsum()
377.         c1_EB1['Samples'] = samples[k]
378.         c1_EB = c1_EB.append(c1_EB1, ignore_index=True)
379.
380.         #####
381.
382.
383.         # When a data frame is created and new data is appended, then integers
become objects, therefore we need to convert them in order tp group them
384.         data_accum[['ISC_GFP', 'EB_GFP', 'EE_GFP', 'EC_GFP', 'ISC', 'EB', 'EE',
'EC']] = data_accum[['ISC_GFP', 'EB_GFP', 'EE_GFP', 'EC_GFP', 'ISC', 'EB', 'E
E', 'EC']].astype(float)
385.         GFP_data_accum[['ISC_GFP', 'EB_GFP', 'EE_GFP', 'EC_GFP']] = GFP_data_ac
cum[['ISC_GFP', 'EB_GFP', 'EE_GFP', 'EC_GFP']].astype(float)
386.
387.         groupby_type = data_accum.groupby('Sample')
388.
389.         # Total count of number clones/clone seize
390.         acum_clone_numbers = sorted(acum_clone_numbers, key=int)
391.         counter = collections.Counter(acum_clone_numbers)
392.
393.         # OUTPUT
394.         #####
395.
396.         # data_accum: Composition in % per stack of all cells
397.         # groupby_type: When both loops are in use, groupby_type groups the dat
a_accum
398.         # GFP_data_accum: Composition in % per stack of only GFP cells

```

```
399.     # cl_df_global: is a counter of how many clusters are for each size
400.     # clone_composition_sum: raw data of the composition of each cluster
```

## Appendix 2: D1 signal quantification

```
1. # -*- coding: utf-8 -*-
2. import os
3. import numpy as np
4. import cv2
5. from matplotlib import pyplot as plt
6. from scipy import ndimage
7. import skimage
8. import skimage.filters as skifi
9. import skimage.morphology as skimo
10. os.chdir('/Users/aleix/Documents/Python/Fastidious_libraries/')
11. import jqtricks as jq
12. import tiffiffle as tf
13. import pandas as pd
14. from skimage.feature import peak_local_max
15. from scipy import ndimage
16. os.chdir('/Users/aleix/Documents/Python/jq scripts')
17.
18. pp_path1 = '/Users/aleix/Desktop/Counting/2017/12. December/20171213_MARCM_emc
NIG_aDl_aPros_7-7+7_MF'
19. pp_path2 = '/Users/aleix/Desktop/Counting/2017/11. November/20171110_MARCM_scB
57_emcRNAi_aDl_aPros_4-7+7_MF'
20. pp_path = [pp_path1, pp_path2]
21. pp_tail = '.txt'
22.
23.
24.
25. folderpath1 = '/Users/aleix/Desktop/Confocal/2017/12.December/20171213_MARCM_e
mcNIG_aDl_aPros_7-7+7_MF/stacks'
26. folderpath2 = '/Users/aleix/Desktop/Confocal/2017/11.November/20171110_MARCM_s
cB57_emcRNAi_aDl_aPros_4-7+7_MF/stack'
27. folderpath = [folderpath1, folderpath2]
28.
29. genotype = [r'\mathit{emc^NIG}', r'\mathit{emc^NIG sc^B57}']
30.
31. def pp_analysis(folderpath, fn) :
32.     # Reading table and mapping the points
33.     countings = pd.read_table(os.path.join(folderpath,fn), usecols=['Type', 'X
', 'Y'])
34.     #x,y = np.loadtxt(os.path.join(folderpath,fn), skiprows=1, usecols=(1, 2),
unpack=True, dtype='uint16') # Separate columns
35.     #XYdata = np.loadtxt(os.path.join(folderpath,fn), skiprows=1, usecols=(1,
2)) #All together
36.     ppImage = np.zeros(im_proj[:, :, 0].shape, dtype = 'float')
37.     ppImage[countings.Y, countings.X] = 255
38.
39.     return countings, ppImage
40.
41. def DAPI(ch0, ch2, ISC_pos, ppImage):
42.     #This function will generate a mask for chanel 0 lacking the eroded nucleu
s (channel 2), as D1 intensity is weaker in the nucleus.
43.
44.     #Filters
45.
46.     median2 = ndimage.filters.median_filter(ch2, size=(2,2))
47.     val2 = skifi.threshold_otsu(median2)
48.     mask2 = (median2 > val2 ).astype(np.uint8)
49.
50.
51.     # Wathershed DAPI
52.
```

```

53.     #Now we want to add the cells that have been lost during the binarisation
54.     #using the pointpicker. What we will do is first merge both boolean arrays,
55.     #then take out the mask with cells, and produce and imdilate for the small
56.     #dots. Then merge again with the rest of the cells
57.     merge = (skimo.binary_dilation((mask2 | (ppImage > 0))-
58.     mask2, skimo.disk(3)).astype(np.uint8)) | mask2
59.
60.     # apply the Watershed algorithm
61.     labels = skimo.watershed((merge-
62.     ppImage), skimage.measure.label(ppImage), mask=merge)
63.
64.     ch0_noNucleus = ch0.astype(float)
65.
66.     for label in labels[ISC_pos.Y,ISC_pos.X]:
67.         # construct the label mask
68.         mask_SinCells = np.asarray([labels == label])[0,:,:)
69.         mask_smallerSinCells = skimo.binary_erosion(mask_SinCells,skimo.disk(2
70.         ))
71.         ch0_noNucleus[mask_smallerSinCells] = np.nan
72.
73.     return ch0_noNucleus
74.
75. def Pros(ppImage, ch0):
76.     #Image processing Prospero
77.
78.     #Apply a median filter and then an otsu theresholding
79.
80.     median0 = ndimage.filters.median_filter(ch0, size=(3,3))
81.     val0 = skifi.threshold_otsu(median0)
82.     mask0 = (median0 > val0 ).astype(np.uint8)
83.
84.     #Now we want to add the cells that have been lost during the binarisation
85.     #using the pointpicker. What we will do is first merge both boolean arrays,
86.     #then take out the mask with cells, and produce and imdilate for the small
87.     #dots. Then merge again with the rest of the cells
88.     merge = (skimo.binary_dilation((mask0 | (ppImage > 0))-
89.     mask0, skimo.disk(1)).astype(np.uint8)) | mask0
90.     # The dilation has to be in function of the pixel size / nuclear radius
91.
92.     # compute the exact Euclidean distance from every binary
93.     # pixel to the nearest zero pixel, then find peaks in this
94.     # distance map
95.     dist = ndimage.distance_transform_edt(merge)
96.
97.     # apply the Watershed algorithm
98.     labels = skimo.watershed(-
99.     dist, skimage.measure.label(ppImage), mask=merge)
100.     print("[INFO] {} unique segments found".format(len(np.unique(labels)) - 1)
101.     )
102.
103.     return labels
104.
105. def GFP(ppGFP, ch1):
106.     #Image processing GFP
107.
108.     #Apply Li's Minimum Cross Entropy theresholding
109.
110.     val1 = skifi.threshold_li(ch1)
111.     mask1 = ch1 > val1
112.     mask1_fill = ndimage.morphology.binary_fill_holes(mask1).astype(np.
113.     uint8)
114.     labeled_mask1 = skimo.label(mask1_fill)

```

```

107.         mask1_removed = skimo.remove_small_objects(labeled_mask1, min_size=
108.             5)
109.         second_mask = (mask1_removed > 0).astype(np.uint8)
110.         add_ppGFP = (skimo.binary_dilation((second_mask | (ppGFP > 0))-
111.             second_mask, skimo.square(2)).astype(np.uint8)) | second_mask
112.         close_mask1 = skimo.closing(add_ppGFP, skimo.square(3))
113.         # The dilation has to be in function of the pixel size / nuclear ra
114.         dius
115.         # compute the exact Euclidean distance from every binary
116.         # pixel to the nearest zero pixel, then find peaks in this
117.         # distance map
118.         dist = ndimage.distance_transform_edt(close_mask1)
119.         # apply the Watershed algorithm
120.         labels = skimo.watershed((close_mask1-
121.             ppGFP), skimage.measure.label(ppGFP), mask=close_mask1)
122.         return labels
123.
124.
125.
126.         cols_tidy = ['Cell_ID', 'Mean_Int_Cell', 'File_name', 'Genotype', 'Mark
127.             er']
128.         tidy_data = pd.DataFrame(columns=cols_tidy, dtype=float)
129.         for t in range(0, len(folderpath)) :
130.             file_list = os.listdir(folderpath[t]) #-> ndimage with I.shape -
131.                 > [time, plane, channel, y, x]
132.             file_list = [x for x in file_list if x.endswith('.tif')] # keep onl
133.                 y files ending with.tif
134.
135.             # Start the loop (prospero or general and then make sub-loops)
136.             for f in range(0,len(file_list)) :
137.                 im = tf.imread(os.path.join(folderpath[t],file_list[f]))
138.                 # Splitting the channel and Z dimensions
139.                 im_col = np.zeros([int(im.shape[0]/3), im.shape[1], im.shape[2]
140.                     , 3],dtype='uint8')
141.                 for x in (range(int(im.shape[0]/3))):
142.                     im_col[x,:,:0] = im[x*3,:,:]
143.                     im_col[x,:,:1] = im[x*3+1,:,:]
144.                     im_col[x,:,:2] = im[x*3+2,:,:]
145.                 im_col = np.transpose(im_col,[1,2,0,3])
146.                 im_proj = np.zeros([im_col.shape[0], im_col.shape[1], im_col.sh
147.                     ape[3]],dtype='uint8')
148.                 # project images
149.                 for c in range(im_col.shape[3]):
150.                     im_proj[:, :,c] = np.max(im_col[:, :, :,c],axis=2)
151.                 ch0 = im_proj[:, :,0]
152.                 ch1 = im_proj[:, :,1]
153.                 ch2 = im_proj[:, :,2]
154.                 countings, ppImage = pp_analysis(pp_path[t], file_list[f].repla
155.                     ce(".tif", "")+pp_tail)
156.                 EE_pos = countings.loc[countings['Type'] == 3]

```

```

163.         labels_pros = Pros(ppImage, ch0)
164.
165.
166.         mask = np.zeros(ppImage.shape, dtype="uint8")
167.
168.
169.         data = pd.DataFrame(columns=cols_tidy, dtype=float)
170.         data_pros = data
171.         pixel_int_list = []
172.
173.         # loop over the unique components for prospero cells
174.         for label in labels_pros[EE_pos.Y,EE_pos.X]:
175.             # construct the label mask
176.             labelMask = np.zeros(ppImage.shape, dtype="uint8")
177.             labelMask[labels_pros == label] = 255
178.             numPixels = cv2.countNonZero(labelMask)
179.
180.             mask_pros = cv2.add(mask, labelMask)
181.
182.             rows = [label, np.mean(ch0[labels_pros == label]), file_list[f], genotype[t], 'Pros']
183.
184.
185.             rows = pd.DataFrame([rows, ], columns=list(cols_tidy))
186.             data_pros = data_pros.append([rows,], ignore_index=True)
187.             pix_int = list(ch0[labels_pros == label])
188.             pixel_int_list = pixel_int_list + pix_int
189.
190.
191.         data_pros.insert(loc=2, column='Mean_Int_Image', value=np.mean(pixel_int_list))
192.
193.
194.         GFP_pos = countings.loc[countings['Type'].isin([1,2,5,6])]
195.         ppGFP = np.zeros(im_proj[:, :, 0].shape, dtype = 'float')
196.         ppGFP[GFP_pos.Y, GFP_pos.X] = 255
197.         labels_GFP = GFP(ppGFP, ch1)
198.
199.
200.
201.         ISC_pos = countings.loc[countings['Type'] == 1]
202.
203.         ch0_noNucleus = DAPI(ch0, ch2, ISC_pos, ppImage)
204.
205.         mask = np.zeros(ppImage.shape, dtype="uint8")
206.         data_Dl = data
207.         pixel_int_list = []
208.
209.         # loop over the unique components for Delta cells
210.         for label in labels_GFP[ISC_pos.Y,ISC_pos.X]:
211.             # construct the label mask
212.             labelMask = np.zeros(ppImage.shape, dtype="uint8")
213.             labelMask[labels_GFP == label] = 255
214.             numPixels = cv2.countNonZero(labelMask)
215.
216.             mask_Dl = cv2.add(mask, labelMask)
217.
218.             rows = [label, np.nanmean(ch0_noNucleus[labels_GFP == label]), file_list[f], genotype[t], 'Dl']
219.
220.
221.             rows = pd.DataFrame([rows, ], columns=list(cols_tidy))
222.             data_Dl = data_Dl.append([rows,], ignore_index=True)
223.             pix_int = list(ch0[labels_GFP == label])
224.             pixel_int_list = pixel_int_list + pix_int
225.

```

```

226.
227.
228.         #Calculate mean intensities per image and merge all the data (n
           p.nanmean as there are nan values)
229.
230.         data_D1.insert(loc=2, column='Mean_Int_Image', value=np.nanmean
           (pixel_int_list))
231.
232.
233.
234.         data = data_pros.append(data_D1, ignore_index=True)
235.
236.         data.insert(loc=6, column='SinCellvsProsMean', value=((data.Mea
           n_Int_Cell)/(data.Mean_Int_Image[0])))
237.         data.insert(loc=7, column='SignalMeanvsProsMean', value=((data.
           Mean_Int_Image)/(data.Mean_Int_Image[0])))
238.
239.         tidy_data = tidy_data.append([data,], ignore_index=True)

```

## References

- Abdelilah-Seyfried, S., Chan, Y. M., Zeng, C., Justice, N. J., Younger-Shepherd, S., Sharp, L. E., Barbel, S., Meadows, S. A., Jan, L. Y. and Jan, Y. N.** (2000). A gain-of-function screen for genes that affect the development of the drosophila adult external sensory organ. *Genetics* **155**, 733–752.
- Albert, B. Y. A. and Anderson, J. A.** (1984). On the existence of maximum likelihood estimates in logistic regression models. *Biometrika* **71**, 1–10.
- Aldaz, S., Morata, G. and Azpiazu, N.** (2003). The Pax-homeobox gene *eyegone* is involved in the subdivision of the thorax of *Drosophila*. *Development* **130**, 4473–4482.
- Alonso, M. C. and Cabrera, C. V.** (1988). The *achaete-scute* gene complex of *Drosophila melanogaster* comprises four homologous genes. *EMBO J.* **7**, 2585–91.
- Amcheslavsky, A., Jiang, J. and Ip, Y. T.** (2009). Tissue Damage-Induced Intestinal Stem Cell Division in *Drosophila*. *Cell Stem Cell* **4**, 49–61.
- Amcheslavsky, A., Nie, Y., Li, Q., He, F., Tsuda, L., Markstein, M. and Ip, Y. T.** (2014a). Gene expression profiling identifies the zinc-finger protein *Charlatan* as a regulator of intestinal stem cells in *Drosophila*. *Development* **141**, 2621–32.
- Amcheslavsky, A., Song, W., Li, Q., Nie, Y., Bragatto, I., Ferrandon, D., Perrimon, N. and Ip, Y. T.** (2014b). Enteroendocrine Cells Support Intestinal Stem-Cell-Mediated Homeostasis in *Drosophila*. *Cell Rep.* 1–8.
- Andersen, D. S., Colombani, J., Palmerini, V., Chakrabandhu, K., Boone, E., Röthlisberger, M., Toggweiler, J., Basler, K., Mapelli, M., Hueber, A. O., et al.** (2015). The *Drosophila* TNF receptor *Grindelwald* couples loss of cell polarity and neoplastic growth. *Nature* **522**, 482–486.
- Antonello, Z. A., Reiff, T., Ballesta-Illan, E. and Dominguez, M.** (2015). Robust intestinal homeostasis relies on cellular plasticity in enteroblasts mediated by miR-8-Escargot switch. *EMBO J.* **34**, 2025–41.
- Arbouzova, N. I. and Zeidler, M. P.** (2006). JAK/STAT signalling in *Drosophila*: insights into conserved regulatory and cellular functions. *Development* **133**, 2605–2616.
- Ashburner, M., Misra, S., Roote, J., Lewis, S. E., Blazej, R., Davis, T., Doyle, C., Galle, R., George, R., Harris, N., et al.** (1999). An exploration of the sequence of a 2.9-Mb region of the genome of *Drosophila melanogaster*: The *Adh* region. *Genetics* **153**, 179–219.
- Atchley, W. R. and Fitch, W. M.** (1997). A natural classification of the basic helix-loop-helix class of transcription factors. *Proc. Natl. Acad. Sci.* **94**, 5172–5176.
- Banezra, R., Davis, R. L., Lassar, A., Tapscott, S., Thayer, M., Lockshon, D. and**

- Weintraub, H.** (1990). Id: A Negative Regulator of Helix-Loop-Helix DNA Binding Proteins: Control of Terminal Myogenic Differentiation. *Ann. N. Y. Acad. Sci.* **599**, 1–11.
- Bang, A. G. and Posakony, J. W.** (1992). The *Drosophila* gene *hairless* encodes a novel basic-protein that controls alternative cell fates in adult sensory organ development. *Genes Dev.* **6**, 1752–1769.
- Bangi, E., Murgia, C., Teague, A. G. S., Sansom, O. J. and Cagan, R. L.** (2016). Functional exploration of colorectal cancer genomes using *Drosophila*. *Nat. Commun.* **7**, 1–16.
- Baonza, A. and García-Bellido, A.** (1999). Dual role of extramacrochaetae in cell proliferation and cell differentiation during wing morphogenesis in *Drosophila*. *Mech. Dev.* **80**, 133–146.
- Baonza, A., de Celis, J. F. and García-Bellido, A.** (2000). Relationships between extramacrochaetae and Notch signalling in *Drosophila* wing development. *Development* **127**, 2383–93.
- Bardin, A. J., Perdigoto, C. N., Southall, T. D., Brand, A. H. and Schweisguth, F.** (2010). Transcriptional control of stem cell maintenance in the *Drosophila* intestine. *Development* **137**, 705–14.
- Barker, N., Van Es, J. H., Kuipers, J., Kujala, P., Van Den Born, M., Cozijnsen, M., Haegerbarth, A., Korving, J., Begthel, H., Peters, P. J., et al.** (2007). Identification of stem cells in small intestine and colon by marker gene *Lgr5*. *Nature* **449**, 1003–1007.
- Barolo, S., Stone, T., Bang, A. G. and Posakony, J. W.** (2002). Default repression and Notch signaling: *Hairless* acts as an adaptor to recruit the corepressors Groucho and dCtBP to suppressor of *hairless*. *Genes Dev.* **16**, 1964–1976.
- Basu, S., Gavert, N., Brabletz, T. and Ben-Ze'ev, A.** (2018). The intestinal stem cell regulating gene *ASCL2* is required for L1-mediated colon cancer progression. *Cancer Lett.* **424**, 9–18.
- Beebe, K., Lee, W. C. and Micchelli, C. a.** (2010). JAK/STAT signaling coordinates stem cell proliferation and multilineage differentiation in the *Drosophila* intestinal stem cell lineage. *Dev. Biol.* **338**, 28–37.
- Beehler-Evans, R. and Micchelli, C. a.** (2015). Generation of enteroendocrine cell diversity in midgut stem cell lineages. *Development* **142**, 654–664.
- Bernard, C.** (1878). *Leçons sur les phénomènes de la vie communs aux animaux et aux végétaux.*
- Bhattacharya, A. and Baker, N. E.** (2009). The HLH protein *Extramacrochaetae* is required for R7 cell and cone cell fates in the *Drosophila* eye. *Dev. Biol.* **327**, 288–300.

- Bhattacharya, A. and Baker, N. E.** (2011). A network of broadly expressed HLH genes regulates tissue-specific cell fates. *Cell* **147**, 881–92.
- Bhattacharya, A., Li, K., Quiquand, M., Rimesso, G. and Baker, N. E.** (2017). The Notch pathway regulates the Second Mitotic Wave cell cycle independently of bHLH proteins. *Dev. Biol.* **431**, 309–320.
- Bhuin, T. and Roy, J. K.** (2014). Rab proteins: The key regulators of intracellular vesicle transport. *Exp. Cell Res.* **328**, 1–19.
- Biteau, B. and Jasper, H.** (2011). EGF signaling regulates the proliferation of intestinal stem cells in Drosophila. *Development* **138**, 1045–1055.
- Biteau, B. and Jasper, H.** (2014). Slit/Robo signaling regulates cell fate decisions in the intestinal stem cell lineage of Drosophila. *Cell Rep.* **7**, 1867–75.
- Biteau, B., Hochmuth, C. E. and Jasper, H.** (2008). JNK activity in somatic stem cells causes loss of tissue homeostasis in the aging Drosophila gut. *Cell Stem Cell* **3**, 442–55.
- Bjerknes, M. and Cheng, H.** (2006). Neurogenin 3 and the enteroendocrine cell lineage in the adult mouse small intestinal epithelium. *Dev. Biol.* **300**, 722–735.
- Blackwood, E. M. and Eisenman, R. N.** (1991). Max: A helix-loop-helix zipper protein that forms a sequence-specific DNA-binding complex with Myc. *Science* (80-. ). **251**, 1211–1217.
- Blackwood, E. M., Luscher, B. and Eisenman, R. N.** (1992). Myc and Max associate in vivo. *Genes Dev.* **6**, 71–80.
- Botas, J., Moscoso del Prado, J. and Garcia-Bellido, A.** (1982). Gene-dose titration analysis in the search of trans-regulatory genes in Drosophila. *EMBO J.* **1**, 307–310.
- Brand, M., Jarman, a P., Jan, L. Y. and Jan, Y. N.** (1993). asense is a Drosophila neural precursor gene and is capable of initiating sense organ formation. *Development* **119**, 1–17.
- Brou, C., Logeat, A., Lecourtois, M., Vandekerckhove, J., Kourilsky, P., Schweisguth, F. and Israel, A.** (1994). Inhibition of the DNA-binding activity of Drosophila supressor of hairless and its human homolog, KBF2/RBP-J'kappa', by direct protein interaction with drosophila Hairless. *Genes Dev.* **8**, 2491–2503.
- Brou, C., Logeat, F., Gupta, N., Bessia, C., LeBail, O., Doedens, J. R., Cumano, A., Roux, P., Black, R. A. and Israël, A.** (2000). A novel proteolytic cleavage involved in Notch signaling: The role of the disintegrin-metalloprotease TACE. *Mol. Cell* **5**, 207–216.

- Brown, N. L., Paddock, S. W., Sattler, C. a, Cronmiller, C., Thomas, B. J. and Carroll, S. B.** (1996). daughterless is required for Drosophila photoreceptor cell determination, eye morphogenesis, and cell cycle progression. *Dev. Biol.* **179**, 65–78.
- Buchon, N., Broderick, N. a., Chakrabarti, S. and Lemaitre, B.** (2009). Invasive and indigenous microbiota impact intestinal stem cell activity through multiple pathways in Drosophila. *Genes Dev.* **23**, 2333–2344.
- Buchon, N., Broderick, N. a, Kuraishi, T. and Lemaitre, B.** (2010). Drosophila EGFR pathway coordinates stem cell proliferation and gut remodeling following infection. *BMC Biol.* **8**, 152.
- Buczacki, S. J. A., Zecchini, H. I., Nicholson, A. M., Russell, R., Vermeulen, L., Kemp, R. and Winton, D. J.** (2013). Intestinal label-retaining cells are secretory precursors expressing lgr5. *Nature* **495**, 65–69.
- Cabrera, C. V and Alonso, M. C.** (1991). Transcriptional activation by heterodimers of the achaete-scute and daughterless gene products of Drosophila. *EMBO J.* **10**, 2965–73.
- Cabrera, C. V, Alonso, M. C. and Huikeshoven, H.** (1994). Regulation of scute function by extramacrochaete in vitro and in vivo. *Development* **120**, 3595–603.
- Campuzano, S.** (2001). Emc, a negative HLH regulator with multiple functions in Drosophila development. *Oncogene* **20**, 8299–307.
- Campuzano, S., Carramolino, L., Cabrera, C. V., Ruiz-Gómez, M., Villares, R., Boronat, A. and Modolell, J.** (1985). Molecular genetics of the achaete-scute gene complex of *D. melanogaster*. *Cell* **40**, 327–338.
- Cannon, W. B.** (1929). Organization for physiological homeostasis. *Physiol. Rev.* **9**, 399–431.
- Cao, J., Tang, M., Li, W. L., Xie, J., Du, H., Tang, W. B., Wang, H., Chen, X. W., Xiao, H. and Li, Y.** (2009). Upregulation of activator protein-4 in human colorectal cancer with metastasis. *Int. J. Surg. Pathol.* **17**, 16–21.
- Castanon, I., Von Stetina, S., Kass, J. and Baylies, M. K.** (2001). Dimerization partners determine the activity of the Twist bHLH protein during Drosophila mesoderm development. *Development* **128**, 3145–3159.
- Caudy, M., Vässin, H., Brand, M., Tuma, R., Jah, L. Y. and Jan, Y. N.** (1988). daughterless, a Drosophila gene essential for both neurogenesis and sex determination, has sequence similarities to myc and the achaete-scute complex. *Cell* **55**, 1061–1067.
- Celis, J. F. De, Celis, J. De, Ligoxygakis, P., Preiss, A., Delidakis, C. and Bray, S. J.** (1996). Functional relationships between Notch, Su(H) and the bHLH genes of the during imaginal development. *Development* **2728**, 2719–2728.

- Chen, S., Zhao, Y., Han, W., Chiu, S. K., Zhu, L., Wu, L. and Yu, K. N.** (2011). Rescue effects in radiobiology: Unirradiated bystander cells assist irradiated cells through intercellular signal feedback. *Mutat. Res. - Fundam. Mol. Mech. Mutagen.* **706**, 59–64.
- Chen, J., Xu, N., Huang, H., Cai, T. and Xi, R.** (2016). A feedback amplification loop between stem cells and their progeny promotes tissue regeneration and tumorigenesis. *Elife* **5**, 1–19.
- Chen, C., Cai, Q., He, W., Lam, T. B., Lin, J., Zhao, Y., Chen, X., Gu, P., Huang, H., Xue, M., et al.** (2017). AP4 modulated by the PI3K/AKT pathway promotes prostate cancer proliferation and metastasis of prostate cancer via upregulating L-plastin. *Cell Death Dis.* **8**, 1–14.
- Chen, J., Xu, N., Wang, C., Huang, P., Huang, H., Jin, Z., Yu, Z., Cai, T., Jiao, R. and Xi, R.** (2018). Transient Scute activation via a self-stimulatory loop directs enteroendocrine cell pair specification from self-renewing intestinal stem cells. *Nat. Cell Biol.* **1**.
- Choi, V. W. Y., Ng, C. Y. P., Cheng, S. H. and Yu, K. N.** (2012).  $\alpha$ -particle irradiated zebrafish embryos rescued by bystander unirradiated zebrafish embryos. *Environ. Sci. Technol.* **46**, 226–231.
- Choksi, S. P., Southall, T. D., Bossing, T., Edoff, K., de Wit, E., Fischer, B. E., van Steensel, B., Micklem, G. and Brand, A. H.** (2006). Prospero Acts as a Binary Switch between Self-Renewal and Differentiation in Drosophila Neural Stem Cells. *Dev. Cell* **11**, 775–789.
- Christy, B. A., Sanders, L. K., Lau, L. F., Copeland, N. G., Jenkins, N. A. and Nathans, D.** (1991). An Id-related helix-loop-helix protein encoded by a growth factor-inducible gene. *Proc. Natl. Acad. Sci.* **88**, 1815–1819.
- Clevers, H.** (2013). The intestinal crypt, a prototype stem cell compartment. *Cell* **154**, 274–284.
- Cohen, M. M.** (2003). The hedgehog signaling network. *Am. J. Med. Genet. A* **123A**, 5–28.
- Cohen, M., Georgiou, M., Stevenson, N. L., Miodownik, M. and Baum, B.** (2010). Dynamic Filopodia Transmit Intermittent Delta-Notch Signaling to Drive Pattern Refinement during Lateral Inhibition. *Dev. Cell* **19**, 78–89.
- Condeelis, J., Singer, R. H. and Segall, J. E.** (2005). THE GREAT ESCAPE: When Cancer Cells Hijack the Genes for Chemotaxis and Motility. *Annu. Rev. Cell Dev. Biol.* **21**, 695–718.
- Cordero, J. B., Stefanatos, R. K., Myant, K., Vidal, M. and Sansom, O. J.** (2012a). Non-autonomous crosstalk between the Jak/Stat and Egfr pathways mediates Apc1-driven intestinal stem cell hyperplasia in the Drosophila adult midgut. *Development* **139**, 4524–4535.

- Cordero, J. B., Stefanatos, R. K., Scopelliti, A., Vidal, M. and Sansom, O. J.** (2012b). Inducible progenitor-derived Wingless regulates adult midgut regeneration in *Drosophila*. *EMBO J.* **31**, 3901–3917.
- Costamagna, D., Berardi, E., Ceccarelli, G. and Sampaolesi, M.** (2015). Adult stem cells and skeletal muscle regeneration. *Curr. Gene Ther.* **15**, 348–63.
- Crews, S. T.** (1998). Control of cell lineage-specific development and transcription by bHLH-PAS proteins. *Genes Dev.* 607–620.
- Cronmiller, C. and Cummings, C. A.** (1993). The daughterless gene product in *Drosophila* is a nuclear protein that is broadly expressed throughout the organism during development. *Mech. Dev.* **42**, 159–169.
- Crook, N. E., Clem, R. J. and Miller, L. K.** (1993). An apoptosis-inhibiting baculovirus gene with a zinc finger-like motif. *J. Virol.* **67**, 2168–2174.
- Cubas, P. and Modolell, J.** (1992). The extramacrochaetae gene provides information for sensory organ patterning. *EMBO J.* **11**, 3385–3393.
- Cubas, P., De Celis, J. F., Campuzano, S. and Modolell, J.** (1991). Proneural clusters of achaete-scute expression and the generation of sensory organs in the *Drosophila* imaginal wing disc. *Genes Dev.* **5**, 996–1008.
- Culí, J. and Modolell, J.** (1998). Proneural gene self-stimulation in neural precursors: an essential mechanism for sense organ development that is regulated by Notch signaling. *Genes Dev.* **12**, 2036–2047.
- De Jossineau, C., Soulé, J., Martin, M., Anguille, C., Montcourrier, P. and Alexandre, D.** (2003). Delta-promoted filopodia mediate long-range lateral inhibition in *Drosophila*. *Nature*.
- de Navascués, J., Perdigoto, C. N., Bian, Y., Schneider, M. H., Bardin, A. J., Martínez-Arias, A. and Simons, B. D.** (2012). *Drosophila* midgut homeostasis involves neutral competition between symmetrically dividing intestinal stem cells. *EMBO J.* **31**, 2473–85.
- Delaunay, B.** (1934). Sur la sphere vide. *Bull. l'Académie des Sci. l'URSS*.
- Diggle, P. J. and Serra, J.** (1983). Image Analysis and Mathematical Morphology. *Biometrics* **39**, 536.
- Domínguez, M. and Campuzano, S.** (1993). asense, a member of the *Drosophila* achaete-scute complex, is a proneural and neural differentiation gene. *EMBO J.* **12**, 2049–60.
- Dutta, D., Dobson, A. J., Houtz, P. L., Gläßer, C., Revah, J., Korzelius, J., Patel, P. H., Edgar, B. A. and Buchon, N.** (2015). Regional Cell-Specific Transcriptome Mapping Reveals Regulatory Complexity in the Adult *Drosophila* Midgut. *Cell Rep.* **12**, 346–358.

- Edgar, B. A. and Orr-Weaver, T. L.** (2001). Endoreplication cell cycles: More for less. *Cell* **105**, 297–306.
- Eisenhoffer, G. T., Loftus, P. D., Yoshigi, M., Otsuna, H., Chien, C. Bin, Morcos, P. A. and Rosenblatt, J.** (2012). Crowding induces live cell extrusion to maintain homeostatic cell numbers in epithelia. *Nature* **484**, 546–549.
- Ellis, H. M., Spann, D. R. and Posakony, J. W.** (1990). extramacrochaetae, a negative regulator of sensory organ development in *Drosophila*, defines a new class of helix-loop-helix proteins. *Cell* **61**, 27–38.
- Farin, H. F., Van Es, J. H. and Clevers, H.** (2012). Redundant sources of Wnt regulate intestinal stem cells and promote formation of paneth cells. *Gastroenterology* **143**, 1518–1529.e7.
- Fidler, I. J.** (2003). The pathogenesis of cancer metastasis: The “seed and soil” hypothesis revisited. *Nat. Rev. Cancer* **3**, 453–458.
- Fischer von Mollard, G., Mignery, G. A., Baumert, M., Perin, M. S., Hanson, T. J., Burger, P. M., Jahn, R. and Südhof, T. C.** (1990). Rab3 is a small GTP-binding protein exclusively localized to synaptic vesicles. *Proc. Natl. Acad. Sci. U. S. A.* **87**, 1988–92.
- Flasse, L. C., Stern, D. G., Pirson, J. L., Manfroid, I., Peers, B. and Voz, M. L.** (2013). The bHLH transcription factor *Ascl1a* is essential for the specification of the intestinal secretory cells and mediates Notch signaling in the zebrafish intestine. *Dev. Biol.* **376**, 187–197.
- Fleming, R. J., Scottgale, T. N., Diederich, R. J. and Artavanis-Tsakonas, S.** (1990). The gene *Serrate* encodes a putative EGF-like transmembrane protein essential for proper ectodermal development in *Drosophila melanogaster*. *Genes Dev.* **4**, 2188–2201.
- Fortini, M. E. and Artavanis-Tsakonas, S.** (1994). The suppressor of hairless protein participates in notch receptor signaling. *Cell* **79**, 273–282.
- Fre, S., Huyghe, M., Mourikis, P., Robine, S., Louvard, D. and Artavanis-Tsakonas, S.** (2005). Notch signals control the fate of immature progenitor cells in the intestine. *Nature* **435**, 964–968.
- Fre, S., Bardin, A. J., Robine, S. and Louvard, D.** (2011). Notch signaling in intestinal homeostasis across species: the cases of *Drosophila*, Zebrafish and the mouse. *Exp. Cell Res.* **317**, 2740–7.
- Furriols, M. and Bray, S.** (2000). Dissecting the mechanisms of suppressor of hairless function. *Dev. Biol.* **227**, 520–532.
- Furriols, M. and Bray, S.** (2001). A model Notch response element detects suppressor of hairless-dependent molecular switch. *Curr. Biol.* **11**, 60–64.

- Fuse, N., Hirose, S. and Hayashi, S.** (1994). Diploidy of *Drosophila* imaginal cells is maintained by a transcriptional repressor encoded by *escargot*. *Genes Dev.* **8**, 2270–2281.
- Garcia-Bellido, A.** (1979). Genetic analysis of the achaete-scute system of *Drosophila melanogaster*. *Genetics* **91**, 491–520.
- Garrell, J. and Modolell, J.** (1990). The *Drosophila* extramacrochaetae locus, an antagonist of proneural genes that, like these genes, encodes a helix-loop-helix protein. *Cell* **61**, 39–48.
- Gassler, N.** (2017). Paneth cells in intestinal physiology and pathophysiology. *World J. Gastrointest. Pathophysiol.* **8**, 150–160.
- Gehart, H. and Clevers, H.** (2015). Repairing organs: Lessons from intestine and liver. *Trends Genet.* **31**, 344–351.
- Giagtzoglou, N., Alifragis, P., Koumbanakis, K. a and Delidakis, C.** (2003). Two modes of recruitment of E(spl) repressors onto target genes. *Development* **130**, 259–270.
- Goulas, S., Conder, R., Knoblich, J. A. and Figures, S.** (2012). The Par Complex and Integrins Direct Asymmetric Cell Division in Adult Intestinal Stem Cells. *Cell Stem Cell* **11**, 529–540.
- Gross, S., Balderes, D., Liu, J., Asfaha, S., Gu, G., Wang, T. C. and Sussel, L.** (2015). Nkx2.2 is expressed in a subset of enteroendocrine cells with expanded lineage potential. *Am. J. Physiol. - Gastrointest. Liver Physiol.* **309**, G975–G987.
- Guisoni, N., Martinez-Corral, R., Garcia-Ojalvo, J. and de Navascués, J.** (2017). Diversity of fate outcomes in cell pairs under lateral inhibition. *Development* **144**, 1177–1186.
- Guo, Z. and Ohlstein, B.** (2015). Bidirectional Notch signaling regulates *Drosophila* intestinal stem cell multipotency. *Science (80-. ).* **350**, aab0988-aab0988.
- Guo, M., Bier, E., Jan, L. Y. and Jan, Y. N.** (1995). tramtrack acts downstream of numb to specify distinct daughter cell fates during asymmetric cell divisions in the *drosophila* PNS. *Neuron* **14**, 913–925.
- Guo, M., Jan, L. Y. and Jan, Y. N.** (1996). Control of daughter cell fates during asymmetric division: Interaction of Numb and Notch. *Neuron* **17**, 27–41.
- Hariharan, I.** (2006). Growth regulation: a beginning for the hippo pathway. *Curr. Biol.* **16**, R1037-9.
- Hartenstein, V. and Posakony, J. W.** (1989). Development of adult sensilla on the wing and notum of *Drosophila melanogaster*. *Development* **107**, 389–405.

- Hartenstein, V., Takashima, S., Hartenstein, P., Asanad, S. and Asanad, K.** (2017). bHLH proneural genes as cell fate determinants of entero-endocrine cells, an evolutionarily conserved lineage sharing a common root with sensory neurons. *Dev. Biol.* **431**, 36–47.
- He, T. C., Sparks, A. B., Rago, C., Hermeking, H., Zawel, L., Da Costa, L. T., Morin, P. J., Vogelstein, B. and Kinzler, K. W.** (1998). Identification of c-MYC as a target of the APC pathway. *Science* (80-. ).
- Heitzler, P. and Simpson, P.** (1991). The choice of cell fate in the epidermis of *Drosophila*. *Cell* **64**, 1083–92.
- Hu, Y. F., Luscher, B., Admon, A., Mermod, N. and Tjian, R.** (1990). Transcription factor AP-4 contains multiple dimerization domains that regulate dimer specificity. *Genes Dev.* **4**, 1741–1752.
- Hu, X., Guo, W., Chen, S., Xu, Y., Li, P., Wang, H., Chu, H., Li, J., Du, Y., Chen, X., et al.** (2016). Silencing of AP-4 inhibits proliferation, induces cell cycle arrest and promotes apoptosis in human lung cancer cells. *Oncol. Lett.* **11**, 3735–3742.
- Huang, J., Wu, S., Barrera, J., Matthews, K. and Pan, D.** (2005). The Hippo signaling pathway coordinately regulates cell proliferation and apoptosis by inactivating Yorkie, the *Drosophila* Homolog of YAP. *Cell* **122**, 421–34.
- Huang, H., Li, J., Hu, L., Ge, L., Ji, H., Zhao, Y. and Zhang, L.** (2014). Bantam is essential for *Drosophila* intestinal stem cell proliferation in response to Hippo signaling. *Dev. Biol.* **385**, 211–219.
- Jackstadt, R., Röh, S., Neumann, J., Jung, P., Hoffmann, R., Horst, D., Berens, C., Bornkamm, G. W., Kirchner, T., Menssen, A., et al.** (2013). AP4 is a mediator of epithelial–mesenchymal transition and metastasis in colorectal cancer. *J. Exp. Med.* **210**, 1331–1350.
- Jaeckel, S., Kaller, M., Jackstadt, R., Götz, U., Müller, S., Boos, S., Horst, D., Jung, P. and Hermeking, H.** (2018). Ap4 is rate limiting for intestinal tumor formation by controlling the homeostasis of intestinal stem cells. *Nat. Commun.* **9**, 3573.
- Jafar-Nejad, H., Tien, A.-C., Acar, M. and Bellen, H. J.** (2006). Senseless and Daughterless confer neuronal identity to epithelial cells in the *Drosophila* wing margin. *Development* **133**, 1683–92.
- Jarman, A. P., Brand, M., Jan, L. Y. and Jan, Y. N.** (1993). The regulation and function of the helix-loop-helix gene, *asense*, in *Drosophila* neural precursors. *Development* **119**, 19–29.
- Jennings, B., Preiss, A., Delidakis, C. and Bray, S.** (1994). The Notch signalling pathway is required for Enhancer of split bHLH protein expression during neurogenesis in the *Drosophila* embryo. *Development* **120**, 3537–3548.

- Jenny, M., Uhl, C., Roche, C., Duluc, I., Guillermin, V., Guillemot, F., Jensen, J., Kedinger, M. and Gradwohl, G.** (2002). Neurogenin3 is differentially required for endocrine cell fate specification in the intestinal and gastric epithelium. *EMBO J.* **21**, 6338–6347.
- Jiang, H. and Edgar, B. A.** (2009). EGFR signaling regulates the proliferation of Drosophila adult midgut progenitors. *Development* **136**, 483–493.
- Jiang, H., Patel, P. H., Kohlmaier, A., Grenley, M. O., McEwen, D. G. and Edgar, B. A.** (2009). Cytokine/Jak/Stat signaling mediates regeneration and homeostasis in the Drosophila midgut. *Cell* **137**, 1343–55.
- Jiang, H., Grenley, M. O., Bravo, M. J., Blumhagen, R. Z. and Edgar, B. a.** (2011). EGFR/Ras/MAPK signaling mediates adult midgut epithelial homeostasis and regeneration in drosophila. *Cell Stem Cell* **8**, 84–95.
- Johnson, J. E., Birren, S. J. and Anderson, D. J.** (1990). Two rat homologues of Drosophila achaete-scute specifically expressed in neuronal precursors. *Nature* **346**, 858–861.
- Johnson, S. A., Zitserman, D. and Roegiers, F.** (2016). Numb regulates the balance between Notch recycling and late-endosome targeting in Drosophila neural progenitor cells. *Mol. Biol. Cell* **27**, 2857–2866.
- Jones, S.** (2004). An overview of the basic helix-loop-helix proteins. *Genome Biol.* **5**, 226.
- Jordà, M., Vinyals, A., Marazuela, A., Cubillo, E., Olmeda, D., Valero, E., Cano, A. and Fabra, À.** (2007). Id-1 is induced in MDCK epithelial cells by activated Erk/MAPK pathway in response to expression of the Snail and E47 transcription factors. *Exp. Cell Res.* **313**, 2389–2403.
- Jubb, A. M., Chalasani, S., Frantz, G. D., Smits, R., Grabsch, H. I., Kavi, V., Maughan, N. J., Hillan, K. J., Quirke, P. and Koeppen, H.** (2006). Achaete-scute like 2 (ascl2) is a target of Wnt signalling and is upregulated in intestinal neoplasia. *Oncogene* **25**, 3445–3457.
- Jung, P. and Hermeking, H.** (2009). The c-MYC-AP4-p21 cascade. *Cell Cycle* **8**, 982–989.
- Jung, P., Menssen, A., Mayr, D. and Hermeking, H.** (2008). AP4 encodes a c-MYC-inducible repressor of p21. *Proc. Natl. Acad. Sci. U. S. A.* **105**, 15046–15051.
- Kamps, M. P., Murre, C., Sun, X. and Baltimore, D.** (1990). A New Homeobox Gene Contributes the DNA Binding Domain of the t(1;19) Translocation Protein in Pre-B ALL. *Cell* **60**, 547–555.
- Karpowicz, P., Perez, J. and Perrimon, N.** (2010). The Hippo tumor suppressor pathway regulates intestinal stem cell regeneration. *Development* **137**, 4135–4145.

- Kazanjian, A., Noah, T., Brown, D., Burkart, J. and Shroyer, N. F.** (2010). Atonal homolog 1 is required for growth and differentiation effects of Notch/ $\gamma$ -secretase inhibitors on normal and cancerous intestinal epithelial cells. *Gastroenterology* **139**, 918–928.e6.
- Kim, T. H. and Shivdasani, R. A.** (2011). Genetic evidence that intestinal Notch functions vary regionally and operate through a common mechanism of math1 repression. *J. Biol. Chem.* **286**, 11427–11433.
- Kim, M. Y., Jeong, B. C., Lee, J. H., Kee, H. J., Kook, H., Kim, N. S., Kim, Y. H., Kim, J. K., Ahn, K. Y. and Kim, K. K.** (2006). A repressor complex, AP4 transcription factor and geminin, negatively regulates expression of target genes in nonneuronal cells. *Proc Natl Acad Sci U S A* **103**, 13074–13079.
- King-Jones, K., Korge, G. and Lehmann, M.** (1999). The helix-loop-helix proteins dAP-4 and daughterless bind both in vitro and in vivo to SEBP3 sites required for transcriptional activation of the Drosophila gene Sgs-4. *J. Mol. Biol.* **291**, 71–82.
- Klambt, C., Knust, E., Tietze, K. and Campos-Ortega, J. A.** (1989). Closely related transcripts encoded by the neurogenic gene complex enhancer of split of Drosophila melanogaster. *Embo J* **8**, 203–210.
- Kopczynski, C. C., Alton, A. K., Fachtel, K., Kooh, P. J. and Muskavitch, M. A.** (1988). Delta, a Drosophila neurogenic gene, is transcriptionally complex and encodes a protein related to blood coagulation factors and epidermal growth factor of vertebrates. *Genes Dev.* **2**, 1723–1735.
- Korzelius, J., Naumann, S. K., Loza-Coll, M. A., Chan, J. S., Dutta, D., Oberheim, J., Gläßer, C., Southall, T. D., Brand, A. H., Jones, D. L., et al.** (2014). Escargot maintains stemness and suppresses differentiation in Drosophila intestinal stem cells. *EMBO J.* **33**, 2967–2982.
- Kunisch, M., Haenlin, M. and Campos-Ortega, J. a** (1994). Lateral inhibition mediated by the Drosophila neurogenic gene delta is enhanced by proneural proteins. *Proc. Natl. Acad. Sci. U. S. A.* **91**, 10139–10143.
- Lai, S.-L. and Doe, C. Q.** (2014). Transient nuclear Prospero induces neural progenitor quiescence. *Elife* **3**, 1–12.
- Lam, R. K. K., Fung, Y. K., Han, W. and Yu, K. N.** (2015). Rescue effects: Irradiated cells helped by unirradiated bystander cells. *Int. J. Mol. Sci.* **16**, 2591–2609.
- Lan, Q., Cao, M., Kollipara, R. K., Rosa, J. B., Kittler, R. and Jiang, H.** (2018). FoxA transcription factor Fork head maintains the intestinal stem/progenitor cell identities in Drosophila. *Dev. Biol.* **433**, 324–343.
- Ledent, V., Paquet, O. and Vervoort, M.** (2002). Phylogenetic analysis of the human basic helix-loop-helix proteins. *Genome Biol.* **3**, 1–18.

- Lee, T. and Luo, L.** (2001). Mosaic analysis with a repressible cell marker (MARCM) for *Drosophila* neural development. *Trends Neurosci.* **24**, 251–254.
- Lee, W. C., Beebe, K., Sudmeier, L. and Micchelli, C. a.** (2009). Adenomatous polyposis coli regulates *Drosophila* intestinal stem cell proliferation. *Development* **136**, 2255–2264.
- Lehmann, R., Jiménez, F., Dietrich, U. and Campos-Ortega, J. A.** (1983). On the phenotype and development of mutants of early neurogenesis in *Drosophila melanogaster*. *Wilhelm Roux's Arch. Dev. Biol.* **192**, 62–74.
- Leptin, M.** (1991). twist and snail as positive and negative regulators during *Drosophila* mesoderm development. *Genes Dev.* **5**, 1568–1576.
- Li, K. and Baker, N. E.** (2018). Regulation of the *Drosophila* ID protein Extra macrochaetae by proneural dimerization partners. *Elife* **7**, 1–22.
- Li, C. H. and Lee, C. K.** (1993). Minimum cross entropy thresholding. *Pattern Recognit.* **26**, 617–625.
- Li, L. and Vaessin, H.** (2000). Pan-neural prospero terminates cell proliferation during *Drosophila* neurogenesis. *Genes Dev.* **14**, 147–151.
- Li, Q., Li, S., Mana-Capelli, S., RothFlach, R. J., Danai, L. V, Amcheslavsky, A., Nie, Y., Kaneko, S., Yao, X., Chen, X., et al.** (2014a). The conserved misshapen-warts-yorkie pathway acts in enteroblasts to regulate intestinal stem cells in *drosophila*. *Dev. Cell* **31**, 291–304.
- Li, Z., Guo, Y., Han, L., Zhang, Y., Shi, L., Huang, X. and Lin, X.** (2014b). Debra-mediated ci degradation controls tissue homeostasis in *drosophila* adult midgut. *Stem Cell Reports* **2**, 135–144.
- Li, Y., Pang, Z., Huang, H., Wang, C., Cai, T. and Xi, R.** (2017a). Transcription Factor Antagonism Controls Enteroendocrine Cell Specification from Intestinal Stem Cells. *Sci. Rep.* **7**, 1–12.
- Li, D., Liu, Y., Pei, C., Zhang, P., Pan, L., Xiao, J., Meng, S., Yuan, Z. and Bi, X.** (2017b). *miR-285*–Yki/Mask double-negative feedback loop mediates blood–brain barrier integrity in *Drosophila*. *Proc. Natl. Acad. Sci.* **114**, E2365–E2374.
- Lim, J., Jafar-Nejad, H., Hsu, Y. C. and Choi, K. W.** (2008). Novel function of the class I bHLH protein daughterless in the negative regulation of proneural gene expression in the *Drosophila* eye. *EMBO Rep.* **9**, 1128–1133.
- Lin, G., Xu, N. and Xi, R.** (2008). Paracrine Wingless signalling controls self-renewal of *Drosophila* intestinal stem cells. *Nature* **455**, 1119–1123.
- Lin, G., Xu, N. and Xi, R.** (2010). Paracrine Unpaired signaling through the JAK/STAT pathway controls self-renewal and lineage differentiation of *Drosophila* intestinal stem cells. *J. Mol. Cell Biol.* **2**, 37–49.

- Liu, Q., Rand, T. A., Kalidas, S., Du, F., Kim, H.-E., Smith, D. P. and Wang, X.** (2003). R2D2, a bridge between the initiation and effector steps of the Drosophila RNAi pathway. *Science* (80-. ). **301**, 1921–1925.
- Liu, X., Jiang, F., Kalidas, S., Smith, D. and Liu, Q.** (2006). Dicer-2 and R2D2 coordinately bind siRNA to promote assembly of the siRISC complexes. *RNA* **12**, 1514–1520.
- Liu, W., Singh, S. R. and Hou, S. X.** (2010). JAK-STAT is restrained by Notch to control cell proliferation of the drosophila intestinal stem cells. *J. Cell. Biochem.* **109**, 992–999.
- Lowe, N., Rees, J. S., Roote, J., Ryder, E., Armean, I. M., Johnson, G., Drummond, E., Spriggs, H., Drummond, J., Magbanua, J. P., et al.** (2014). Analysis of the expression patterns, subcellular localisations and interaction partners of Drosophila proteins using a pigP protein trap library. *Development* **141**, 3994–4005.
- Loza-coll, M. A., Southall, T. D., Sandall, S. L., Brand, A. H. and Jones, D. L.** (2014). Regulation of Drosophila intestinal stem cell maintenance and differentiation by the transcription factor Escargot. *EMBO J.* **33**, 2983–2996.
- Lu, Y. and Li, Z.** (2015). No Intestinal Stem Cell Regeneration after Complete Progenitor Ablation in Drosophila Adult Midgut. *J. Genet. Genomics* **42**, 83–86.
- Ma, W., Liu, B., Li, J., Jiang, J., Zhou, R., Huang, L., Li, X., He, X. and Zhou, Q.** (2018). MicroRNA-302c represses epithelial–mesenchymal transition and metastasis by targeting transcription factor AP-4 in colorectal cancer. *Biomed. Pharmacother.* **105**, 670–676.
- Maier, D., Nagel, A. C., Johannes, B. and Preiss, A.** (1999). Subcellular localization of Hairless protein shows a major focus of activity within the nucleus. *Mech. Dev.* **89**, 195–199.
- Marshman, E., Booth, C. and Potten, C. S.** (2002). The intestinal epithelial stem cell. *BioEssays* **24**, 91–98.
- Martínez, C., Modolell, J. and Garrell, J.** (1993). Regulation of the proneural gene achaete by helix-loop-helix proteins. *Mol. Cell. Biol.* **13**, 3514–3521.
- Meng, F. W. and Biteau, B.** (2015). A Sox Transcription Factor Is a Critical Regulator of Adult Stem Cell Proliferation in the Drosophila Intestine. *Cell Rep.* **13**, 906–914.
- Mermod, N., Williams, T. J. and Tjian, R.** (1988). Enhancer binding factors AP-4 and AP-1 act in concert to activate SV40 late transcription in vitro. *Nature* **332**, 557–561.
- Meyer, F. and Beucher, S.** (1990). Morphological segmentation. *J. Vis. Commun. Image Represent.*
- Micchelli, C. a and Perrimon, N.** (2006). Evidence that stem cells reside in the adult Drosophila midgut epithelium. *Nature* **439**, 475–9.

- Milano, J., McKay, J., Dagenais, C., Foster-Brown, L., Pognan, F., Gadiant, R., Jacobs, R. T., Zacco, A., Greenberg, B. and Ciaccio, P. J.** (2004). Modulation of Notch processing by  $\gamma$ -secretase inhibitors causes intestinal goblet cell metaplasia and induction of genes known to specify gut secretory lineage differentiation. *Toxicol. Sci.* **82**, 341–358.
- Mohr, O.** (1919). Character changes caused by mutation of an entire region of a chromosome in *Drosophila*. *Genetics* **275**, 275–282.
- Montagne, C. and Gonzalez-Gaitan, M.** (2014). Sara endosomes and the asymmetric division of intestinal stem cells. *Development* **141**, 2014–23.
- Moscoso del Prado, J. and Garcia-Bellido, A.** (1984). Cell interactions in the generation of chaetae pattern in *Drosophila*. *Wilhelm Roux's Arch. Dev. Biol.* **193**, 246–251.
- Muñoz, J., Stange, D. E., Schepers, A. G., Van De Wetering, M., Koo, B. K., Itzkovitz, S., Volckmann, R., Kung, K. S., Koster, J., Radulescu, S., et al.** (2012). The Lgr5 intestinal stem cell signature: Robust expression of proposed quiescent ' +4 ' cell markers. *EMBO J.* **31**, 3079–3091.
- Murre, C., McCaw, P. S., Vaessin, H., Caudy, M., Jan, L. Y., Jan, Y. N., Cabrera, C. V, Buskin, J. N., Hauschka, S. D. and Lassar, a B.** (1989a). Interactions between heterologous helix-loop-helix proteins generate complexes that bind specifically to a common DNA sequence. *Cell* **58**, 537–544.
- Murre, C., McCaw, P. S. and Baltimore, D.** (1989b). A new DNA binding and dimerization motif in immunoglobulin enhancer binding, daughterless, MyoD, and myc proteins. *Cell* **56**, 777–783.
- Murre, C., Voronova, A. and Baltimore, D.** (1991). B-cell- and myocyte-specific E2-box-binding factors contain E12/E47-like subunits. *Mol. Cell. Biol.* **11**, 1156–1160.
- Murre, C., Bain, G., van Dijk, M. A., Engel, I., Furnari, B. A., Massari, M. E., Matthews, J. R., Quong, M. W., Rivera, R. R. and Stuver, M. H.** (1994). Structure and function of helix-loop-helix proteins. *BBA - Gene Struct. Expr.* **1218**, 129–135.
- Nelson, B. R. and Reh, T. A.** (2008). Relationship between Delta-like and proneural bHLH genes during chick retinal development. *Dev. Dyn.* **237**, 1565–1580.
- Nelson, B. R., Hartman, B. H., Ray, C. A., Hayashi, T., Bermingham-McDonogh, O. and Reh, T. A.** (2009). Acheate-scute like 1 (Ascl1) is required for normal delta-like (Dll) gene expression and notch signaling during retinal development. *Dev. Dyn.* **238**, 2163–2178.
- Neuhold, L. A. and Wold, B.** (1993). HLH forced dimers: Tethering MyoD to E47 generates a dominant positive myogenic factor insulated from negative regulation by Id. *Cell* **74**, 1033–1042.
- Nieto, M. A.** (2002). The snail superfamily of zinc-finger transcription factors. *Nat. Rev. Mol. Cell Biol.* **3**, 155–166.

- Nigmatullina, L., Norkin, M., Dzama, M. M. and Sayols, S.** (2017). Id2 controls Lgr5 + intestinal stem cells progenitors specification during gut development. *EMBO J.* **37**, 869–885.
- Ohlstein, B. and Spradling, A.** (2006). The adult *Drosophila* posterior midgut is maintained by pluripotent stem cells. *Nature* **439**, 470–4.
- Ohlstein, B. and Spradling, A.** (2007). Multipotent *Drosophila* intestinal stem cells specify daughter cell fates by differential notch signaling. *Science* **315**, 988–92.
- Otsu, N.** (1979). A Threshold Selection Method from Gray-Level Histograms. *IEEE Trans. Syst. Man. Cybern.* **9**, 62–66.
- Patel, P. H., Dutta, D. and Edgar, B. A.** (2015). Niche appropriation by *Drosophila* intestinal stem cell tumours. *Nat. Cell Biol.* **17**, 1182–1192.
- Pellegrinet, L., Rodilla, V., Liu, Z., Chen, S., Koch, U., Espinosa, L., Kaestner, K. H., Kopan, R., Lewis, J. and Radtke, F.** (2011). Dll1- and Dll4-mediated notch signaling are required for homeostasis of intestinal stem cells. *Gastroenterology* **140**, 1230–1240.
- Perdigoto, C. N. and Bardin, A. J.** (2013). Sending the right signal: Notch and stem cells. *Biochim. Biophys. Acta* **1830**, 2307–22.
- Perdigoto, C. N., Schweisguth, F. and Bardin, A. J.** (2011). Distinct levels of Notch activity for commitment and terminal differentiation of stem cells in the adult fly intestine. *Development* **138**, 4585–95.
- Pérez-Garijo, A., Fuchs, Y. and Steller, H.** (2013). Apoptotic cells can induce non-autonomous apoptosis through the TNF pathway. *Elife* **2013**, 1–18.
- Petcherski, A. G. and Kimble, J.** (2000). Mastermind is a putative activator for Notch [1]. *Curr. Biol.* **10**, 471–473.
- Port, F. and Bullock, S. L.** (2016). Creating heritable mutations in *Drosophila* with CRISPR-cas9. In *Methods in Molecular Biology*, pp. 145–160.
- Port, F., Chen, H.-M., Lee, T. and Bullock, S. L.** (2014). Optimized CRISPR/Cas tools for efficient germline and somatic genome engineering in *Drosophila*. *Proc. Natl. Acad. Sci. U. S. A.* E2967–E2076.
- Potten, C. S., Hume, W. J., Reid, P. and Cairns, J.** (1978). The segregation of DNA in epithelial stem cells. *Cell* **15**, 899–906.
- Rawlins, E. L. and Hogan, B. L. M.** (2006). Epithelial stem cells of the lung: privileged few or opportunities for many? *Development* **133**, 2455–2465.
- Rebay, I., Fleming, R. J., Fehon, R. G., Cherbas, L., Cherbas, P. and Artavanis-Tsakonas, S.** (1991). Specific EGF repeats of Notch mediate interactions with Delta and serrate: Implications for notch as a multifunctional receptor. *Cell* **67**, 687–699.

- Rehfeld, J. F.** (1998). The new biology of gastrointestinal hormones. *Physiol. Rev.* **78**, 1087–108.
- Ren, F., Wang, B., Yue, T., Yun, E.-Y., Ip, Y. T. and Jiang, J.** (2010). Hippo signaling regulates *Drosophila* intestine stem cell proliferation through multiple pathways. *Proc. Natl. Acad. Sci. U. S. A.* **107**, 21064–21069.
- Ren, F., Shi, Q., Chen, Y., Jiang, A., Ip, Y. T., Jiang, H. and Jiang, J.** (2013). *Drosophila* Myc integrates multiple signaling pathways to regulate intestinal stem cell proliferation during midgut regeneration. *Cell Res.* **23**, 1133–46.
- Ren, W., Zhang, Y., Li, M., Wu, L., Wang, G., Baeg, G. H., You, J., Li, Z. and Lin, X.** (2015). Windpipe Controls *Drosophila* Intestinal Homeostasis by Regulating JAK/STAT Pathway via Promoting Receptor Endocytosis and Lysosomal Degradation. *PLoS Genet.* **11**, 1–26.
- Resende, L. P. F., Truong, M. E., Gomez, A. and Jones, D. L.** (2017). Intestinal stem cell ablation reveals differential requirements for survival in response to chemical challenge. *Dev. Biol.* **424**, 10–17.
- Riechmann, V., van Cruchten, I. and Sablitzky, F.** (1994). The expression pattern of Id4, a novel dominant negative helix-loop-helix protein, is distinct from Id1, Id2 and Id3. *Nucleic Acids Res* **22**, 749–755.
- Roach, G., Heath Wallace, R., Cameron, A., Emrah Ozel, R., Hongay, C. F., Baral, R., Andreescu, S. and Wallace, K. N.** (2013). Loss of *ascl1a* prevents secretory cell differentiation within the zebrafish intestinal epithelium resulting in a loss of distal intestinal motility. *Dev. Biol.* **376**, 171–186.
- Roignant, J.-Y. and Treismanuthor, J.** (2010). Pattern Formation In *Drosophila* Eye. *Int. J. Dev. Biol.* **53**, 795–804.
- Romani, S., Campuzano, S., Macagno, E. R. and Modolell, J.** (1989). Expression of achaete and scute genes in *Drosophila* imaginal discs and their function in sensory organ development. *Genes Dev.* **3**, 997–1007.
- Rooke, J., Pan, D., Xu, T. and Rubin, G. M.** (1996). KUZ, a conserved metalloprotease-disintegrin protein with two roles in *Drosophila* neurogenesis. *Science (80-. ).* **273**, 1227–1231.
- Rushlow, C. a, Hogan, a, Pinchin, S. M., Howe, K. M., Lardelli, M. and Ish-Horowicz, D.** (1989). The *Drosophila* hairy protein acts in both segmentation and bristle patterning and shows homology to N-myc. *EMBO J.* **8**, 3095–103.
- Ruzinova, M. B. and Benezra, R.** (2003). Id proteins in development, cell cycle and cancer. *Trends Cell Biol.* **13**, 410–418.
- Sallé, J., Gervais, L., Boumard, B., Stefanutti, M., Siudeja, K. and Bardin, A. J.** (2017). Intrinsic regulation of enteroendocrine fate by Numb. *EMBO J.* **36**, 1–18.

- Sallee, M. D., Littleford, H. E. and Greenwald, I.** (2017). A bHLH Code for Sexually Dimorphic Form and Function of the *C. elegans* Somatic Gonad. *Curr. Biol.* **27**, 1853–1860.
- Sandler, L.** (1972). On the genetic control of genes located in the sex-chromosome heterochromatin of *Drosophila melanogaster*. *Genetics* **70**, 261–74.
- Sato, T., Van Es, J. H., Snippert, H. J., Stange, D. E., Vries, R. G., Van Den Born, M., Barker, N., Shroyer, N. F., Van De Wetering, M. and Clevers, H.** (2011). Paneth cells constitute the niche for Lgr5 stem cells in intestinal crypts. *Nature* **469**, 415–418.
- Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B., et al.** (2012). Fiji: An open-source platform for biological-image analysis. *Nat. Methods* **9**, 676–682.
- Schmitz, A. A. P., Govek, E. E., Böttner, B. and Van Aelst, L.** (2000). Rho GTPases: Signaling, migration, and invasion. *Exp. Cell Res.* **261**, 1–12.
- Schuijers, J., Junker, J. P., Mokry, M., Hatzis, P., Koo, B. K., Sasselli, V., Van Der Flier, L. G., Cuppen, E., Van Oudenaarden, A. and Clevers, H.** (2015). Ascl2 acts as an R-spondin/wnt-responsive switch to control stemness in intestinal crypts. *Cell Stem Cell* **16**, 158–170.
- Schwartz, R., Engel, I., Fallahi-Sichani, M., Petrie, H. T. and Murre, C.** (2006). Gene expression patterns define novel roles for E47 in cell cycle progression, cytokine-mediated signaling, and T lineage development. *Proc. Natl. Acad. Sci.* **103**, 9976–9981.
- Schwitalla, S., Fingerle, A. A., Cammareri, P., Nebelsiek, T., Go, S. I., Ziegler, P. K., Canli, O., Heijmans, J., Huels, D. J., Moreaux, G., et al.** (2013). Intestinal Tumorigenesis Initiated by Dedifferentiation and Acquisition of Stem-Cell-like Properties. 25–38.
- Scopelliti, A., Cordero, J. B., Diao, F., Strathdee, K., White, B. H., Sansom, O. J. and Vidal, M.** (2014). Local control of intestinal stem cell homeostasis by enteroendocrine cells in the adult *Drosophila* midgut. *Curr. Biol.* **24**, 1199–211.
- Shaw, R. L., Kohlmaier, A., Polesello, C., Veelken, C., Edgar, B. A. and Tapon, N.** (2010). The Hippo pathway regulates intestinal stem cell proliferation during *Drosophila* adult midgut regeneration. *Development* **137**, 4147–58.
- Shroyer, N. F., Wallis, D., Venken, K. J. T., Bellen, H. J. and Zoghbi, H. Y.** (2005). Gfi1 functions downstream of Math1 to control intestinal secretory cell subtype allocation and differentiation. *Genes Dev.* **19**, 2412–2417.
- Shroyer, N. F., Helmrath, M. A., Wang, V. Y. C., Antalffy, B., Henning, S. J. and Zoghbi, H. Y.** (2007). Intestine-Specific Ablation of Mouse atonal homolog 1 (Math1) Reveals a Role in Cellular Homeostasis. *Gastroenterology* **132**, 2478–2488.

- Shu, Z. and Deng, W.-M.** (2017). Differential Regulation of Cyclin E by Yorkie-Scalloped Signaling in Organ Development. *G3 Genes|Genomes|Genetics* **7**, 1049–1060.
- Sikder, H. a., Devlin, M. K., Dunlap, S., Ryu, B. and Alani, R. M.** (2003). Id proteins in cell growth and tumorigenesis. *Cancer Cell* **3**, 525–530.
- Simpson, P.** (1990). Lateral inhibition and the development of the sensory bristles of the adult peripheral nervous system of *Drosophila*. *Development* **109**, 509–19.
- Sjöqvist, M. and Andersson, E. R.** (2017). Do as I say, Not(ch) as I do: Lateral control of cell fate. *Dev. Biol.* 1–13.
- Skeath, J. B. and Carroll, S. B.** (1991). Regulation of achaete-scute gene expression and sensory organ pattern formation in the *Drosophila* wing. *Genes Dev.* **5**, 984–995.
- Smith, J. E. and Cronmiller, C.** (2001). The *Drosophila* daughterless gene autoregulates and is controlled by both positive and negative cis regulation. *Development* **128**, 4705–14.
- Snippert, H. J., van der Flier, L. G., Sato, T., van Es, J. H., van den Born, M., Kroon-Veenboer, C., Barker, N., Klein, A. M., van Rheenen, J., Simons, B. D., et al.** (2010). Intestinal crypt homeostasis results from neutral competition between symmetrically dividing Lgr5 stem cells. *Cell* **143**, 134–144.
- Spradling, A. C., Stern, D., Beaton, A., Rhem, E. J., Laverly, T., Mozden, N., Misra, S. and Rubin, G. M.** (1999). The Berkeley *Drosophila* Genome Project gene disruption project: Single P-element insertions mutating 25% of vital *Drosophila* genes. *Genetics* **153**, 135–177.
- Spratford, C. M. and Kumar, J. P.** (2013). Extramacrochaetae imposes order on the *Drosophila* eye by refining the activity of the Hedgehog signaling gradient. *Development* **140**, 1994–2004.
- Spratford, C. M. and Kumar, J. P.** (2015a). Extramacrochaetae functions in dorsal-ventral patterning of *Drosophila* imaginal discs. *Development* **142**, 1006–1015.
- Spratford, C. M. and Kumar, J. P.** (2015b). Inhibition of Daughterless by Extramacrochaetae mediates Notch-induced cell proliferation. *Development* **142**, 2058–2068.
- Staley, B. K. and Irvine, K. D.** (2010). Warts and yorkie mediate intestinal regeneration by influencing stem cell proliferation. *Curr. Biol.* **20**, 1580–1587.
- Struhl, G. and Adachi, A.** (2000). Requirements for Presenilin-dependent cleavage of notch and other transmembrane proteins. *Mol. Cell* **6**, 625–636.
- Struhl, G. and Greenwald, I.** (1999). Presenilin is required for activity and nuclear access of notch in *drosophila*. *Nature* **398**, 522–525.

- Suijkerbuijk, S. J. E., Kolahgar, G., Kucinski, I. and Piddini, E.** (2016). Cell competition drives the growth of intestinal adenomas in *Drosophila*. *Curr. Biol.* **26**, 428–438.
- Sun, X. and Baltimore, D.** (1991). An Inhibitory Domain of El2 Transcription Factor Prevents DNA Binding in El2 Homodimers but Not in El2 Heterodimers. *Cell* **64**, 459–470.
- Sun, X. H., Copeland, N. G., Jenkins, N. A. and Baltimore, D.** (1991). Id proteins Id1 and Id2 selectively inhibit DNA binding by one class of helix-loop-helix proteins. *Mol. Cell. Biol.* **11**, 5603–11.
- Tanaka-Matakatsu, M., Miller, J., Borger, D., Tang, W. J. and Du, W.** (2014). Daughterless homodimer synergizes with Eyeless to induce Atonal expression and retinal neuron differentiation. *Dev. Biol.* **392**, 256–265.
- Tian, H., Biehs, B., Warming, S., Leong, K. G., Rangell, L., Klein, O. D. and De Sauvage, F. J.** (2011). A reserve stem cell population in small intestine renders Lgr5-positive cells dispensable. *Nature* **478**, 255–259.
- Troost, T., Schneider, M. and Klein, T.** (2015). A Re-examination of the Selection of the Sensory Organ Precursor of the Bristle Sensilla of *Drosophila melanogaster*. *PLoS Genet.* **11**, 1–21.
- Tsujimoto, K., Ono, T., Sato, M., Nishida, T., Oguma, T. and Tadakuma, T.** (2005). Regulation of the expression of caspase-9 by the transcription factor activator protein-4 in glucocorticoid-induced apoptosis. *J. Biol. Chem.* **280**, 27638–27644.
- Ueo, T., Imayoshi, I., Kobayashi, T., Ohtsuka, T., Seno, H., Nakase, H., Chiba, T. and Kageyama, R.** (2012). The role of Hes genes in intestinal development, homeostasis and tumor formation. *Development* **139**, 1071–1082.
- Vaessin, H., Grell, E., Wolff, E., Bier, E., Jan, L. Y. and Jan, Y. N.** (1991). prospero is expressed in neuronal precursors and encodes a nuclear protein that is involved in the control of axonal outgrowth in *Drosophila*. *Cell* **67**, 941–953.
- Valenta, T., Degirmenci, B., Moor, A. E., Herr, P., Zimmerli, D., Moor, M. B., Hausmann, G., Cantù, C., Aguet, M. and Basler, K.** (2016). Wnt Ligands Secreted by Subepithelial Mesenchymal Cells Are Essential for the Survival of Intestinal Stem Cells and Gut Homeostasis. *Cell Rep.* **15**, 911–918.
- van der Flier, L. G. and Clevers, H.** (2009). Stem Cells, Self-Renewal, and Differentiation in the Intestinal Epithelium. *Annu. Rev. Physiol.* **71**, 241–260.
- van der Flier, L. G., van Gijn, M. E., Hatzis, P., Kujala, P., Haegebarth, A., Stange, D. E., Begthel, H., van den Born, M., Guryev, V., Oving, I., et al.** (2009). Transcription Factor Achaete Scute-Like 2 Controls Intestinal Stem Cell Fate. *Cell* **136**, 903–912.

- Van Doren, M., Ellis, H. M. and Posakony, J. W.** (1991). The *Drosophila* extramacrochaetae protein antagonizes sequence-specific DNA binding by daughterless/achaete-scute protein complexes. *Development* **113**, 245–255.
- Van Doren, M., Powell, P. A., Pasternak, D., Singson, A. and Posakony, W.** (1992). Spatial regulation of proneural gene activity: Auto- and cross-activation of achaete is antagonized by extramacrochaetae. *Genes Dev.* **6**, 2592–2605.
- van Es, J. H., van Gijn, M. E., Riccio, O., van den Born, M., Vooijs, M., Begthel, H., Cozijnsen, M., Robine, S., Winton, D. J., Radtke, F., et al.** (2005). Notch/ $\gamma$ -secretase inhibition turns proliferative cells in intestinal crypts and adenomas into goblet cells. *Nature* **435**, 959–963.
- VanDussen, K. L. and Samuelson, L. C.** (2010). Mouse atonal homolog 1 directs intestinal progenitors to secretory cell rather than absorptive cell fate. *Dev. Biol.* **346**, 215–223.
- Vässin, H., Bremer, K. A., Knust, E. and Campos-Ortega, J. A.** (1987). The neurogenic gene Delta of *Drosophila melanogaster* is expressed in neurogenic territories and encodes a putative transmembrane protein with EGF-like repeats. *EMBO J.* **6**, 3431–40.
- Villares, R. and Cabrera, C. V.** (1987). The achaete-scute gene complex of *D. melanogaster*: Conserved Domains in a subset of genes required for neurogenesis and their homology to myc. *Cell* **50**, 415–424.
- Voronova, A. and Baltimore, D.** (1990). Mutations that disrupt DNA binding and dimer formation in the E47 helix-loop-helix protein map to distinct domains. *Proc. Natl. Acad. Sci. U. S. A.* **87**, 4722–4726.
- Walker, M. D., Park, C. W., Rosen, A. and Aronheim, A.** (1990). A cDNA from a mouse pancreatic b cell encoding a putative transcription factor of the insulin gene. *Nucleic Acids Res.* **18**, 1159–1166.
- Wang, L. and Baker, N. E.** (2015). Salvador-Warts-Hippo Pathway in a Developmental Checkpoint Monitoring Helix-Loop-Helix Proteins. *Dev. Cell* **32**, 191–202.
- Wang, C. and Xi, R.** (2015). Keeping intestinal stem cell differentiation on the Tramtrack. *Fly (Austin)*. **9**, 110–114.
- Wang, C., Guo, X., Dou, K., Chen, H. and Xi, R.** (2015). Ttk69 acts as a master repressor of enteroendocrine cell specification in *Drosophila* intestinal stem cell lineages. *Development* **142**, 3321–3331.
- Wang, L., Meng, Y., Xu, J.-J. and Zhang, Q.-Y.** (2018). The Transcription Factor AP4 Promotes Oncogenic Phenotypes and Cisplatin Resistance by Regulating LAPTM4B Expression. *Mol. Cancer Res.* **16**, 857–868.

- Wharton, K. A., Johansen, K. M., Xu, T. and Artavanis-Tsakonas, S.** (1985). Nucleotide sequence from the neurogenic locus Notch implies a gene product that shares homology with proteins containing EGF-like repeats. *Cell* **43**, 567–581.
- Wice, B. M. and Gordon, J. I.** (1998). Forced expression of Id-1 in the adult mouse small intestinal epithelium is associated with development of adenomas. *J. Biol. Chem.* **273**, 25310–25319.
- Wong, M. C., Castanon, I. and Baylies, M. K.** (2008). Daughterless dictates Twist activity in a context-dependent manner during somatic myogenesis. *Dev. Biol.* **317**, 417–429.
- Wong, M. M.-K., Liu, M.-F. and Chiu, S. K.** (2015). Cropped, Drosophila transcription factor AP-4, controls tracheal terminal branching and cell growth. *BMC Dev. Biol.* **15**, 20.
- Wosczyzna, M. N. and Rando, T. A.** (2018). A Muscle Stem Cell Support Group: Coordinated Cellular Responses in Muscle Regeneration. *Dev. Cell* **46**, 135–143.
- Xi, W.-D., Liu, Y.-J., Sun, X.-B., Shan, J., Yi, L. and Zhang, T.-T.** (2017). Bioinformatics analysis of RNA-seq data revealed critical genes in colon adenocarcinoma. *Eur. Rev. Med. Pharmacol. Sci.* **21**, 3012–3020.
- Xinghua, L., Bo, Z., Yan, G., Lei, W., Changyao, W., Qi, L., Lin, Y., Kaixiong, T., Guobin, W. and Jianying, C.** (2012). The overexpression of AP-4 as a prognostic indicator for gastric carcinoma. *Med. Oncol.* **29**, 871–877.
- Xu, N., Wang, S. Q., Tan, D., Gao, Y., Lin, G. and Xi, R.** (2011). EGFR, Wingless and JAK/STAT signaling cooperatively maintain Drosophila intestinal stem cells. *Dev. Biol.* **354**, 31–43.
- Yamaguchi, H., Wyckoff, J. and Condeelis, J.** (2005). Cell migration in tumors. *Curr. Opin. Cell Biol.* **17**, 559–564.
- Yan, K. S., Gevaert, O., Zheng, G. X. Y., Anchang, B., Probert, C. S., Larkin, K. A., Davies, P. S., Cheng, Z. fen, Kaddis, J. S., Han, A., et al.** (2017). Intestinal Enteroendocrine Lineage Cells Possess Homeostatic and Injury-Inducible Stem Cell Activity. *Cell Stem Cell* **21**, 78–90.
- Yang, Q., Bermingham, N. A., Finegold, M. J. and Zoghbi, H. Y.** (2001). Requirement of Math1 for secretory cell lineage commitment in the mouse intestine. *Science* (80- .). **294**, 2155–2158.
- Yang, D. J., Chung, J. Y., Lee, S. J., Park, S. Y., Pyo, J. H., Ha, N. C., Yoo, M. A. and Park, B. J.** (2010). Slug, mammalian homologue gene of Drosophila escargot, promotes neuronal-differentiation through suppression of HEB/daughterless. *Cell Cycle* **9**, 2789–2802.

- Yang, J., Ma, J.-P., Xiao, S., Zhang, X.-H., Xu, J.-B., Chen, C.-Q., Cai, S.-R. and He, Y.-L.** (2018). Evaluating the prognostic value and functional roles of transcription factor AP4 in colorectal cancer. *Oncol. Lett.* **15**, 7545–7554.
- Yasugi, T., Fischer, A., Jiang, Y., Reichert, H. and Knoblich, J. a.** (2014). A regulatory transcriptional loop controls proliferation and differentiation in Drosophila neural stem cells. *PLoS One* **9**, 1–7.
- Ye, Y., Lukinova, N. and Fortini, M. E.** (1999). Neurogenic phenotypes and altered notch processing in drosophila presenilin mutants. *Nature* **398**, 525–529.
- Yin, C. and Xi, R.** (2018). A Phyllopod-Mediated Feedback Loop Promotes Intestinal Stem Cell Enteroendocrine Commitment in Drosophila. *Stem Cell Reports* **10**, 43–57.
- Zaret, K. S., Caravaca, J. M., Tulin, A. and Sekiya, T.** (2010). Nuclear Mobility and Mitotic Chromosome Binding : Similarities between Pioneer Transcription Factor FoxA and Linker Histone H1 Nuclear Mobility and Mitotic Chromosome Binding Similarities between Pioneer Transcription Factor FoxA and Linker Histone H1. *Cold Spring Harb. Symp. Quant. Biol.* **75**, 219–226.
- Zarifi, I., Kiparaki, M., Koumbanakis, K. a, Giagtzoglou, N., Zacharioudaki, E., Alexiadis, A., Livadaras, I. and Delidakis, C.** (2012a). Essential roles of Da transactivation domains in neurogenesis and in E(spl)-mediated repression. *Mol. Cell. Biol.* **32**, 4534–48.
- Zarifi, I., Kiparaki, M., Koumbanakis, K. a., Giagtzoglou, N., Zacharioudaki, E., Alexiadis, A., Livadaras, I. and Delidakis, C.** (2012b). Essential Roles of Da Transactivation Domains in Neurogenesis and in E(spl)-Mediated Repression. *Mol. Cell. Biol.* **32**, 4534–4548.
- Zecca, M., Basler, K. and Struhl, G.** (1995). Sequential organizing activities of engrailed, hedgehog and decapentaplegic in the Drosophila wing. *Development* **121**, 2265–78.
- Zeng, X. and Hou, S. X.** (2015). Enteroendocrine cells are generated from stem cells through a distinct progenitor in the adult Drosophila posterior midgut. *Development* **142**, 644–653.
- Zeng, X., Chauhan, C. and Hou, S. X.** (2010). Characterization of Midgut Stem Cell- and Enteroblast-Specific Gal4 Lines in Drosophila. *Genesis* **48**, 607–611.
- Zeng, X., Lin, X. and Hou, S. X.** (2013). The Osa-containing SWI/SNF chromatin-remodeling complex regulates stem cell commitment in the adult Drosophila intestine. *Development* **140**, 3532–40.
- Zhai, Z., Kondo, S., Ha, N., Boquete, J. P., Brunner, M., Ueda, R. and Lemaitre, B.** (2015). Accumulation of differentiating intestinal stem cell progenies drives tumorigenesis. *Nat. Commun.* **6**, 1–13.

- Zhai, Z., Boquete, J. P. and Lemaitre, B.** (2017). A genetic framework controlling the differentiation of intestinal stem cells during regeneration in *Drosophila*. *PLoS Genet.* **13**, 1–27.
- Zhang, N., Yantiss, R. K., Nam, H., Chin, Y., Zhou, X. K., Scherl, E. J., Bosworth, B. P., Subbaramaiah, K., Dannenberg, A. J. and Benezra, R.** (2014). ID1 Is a Functional Marker for Intestinal Stem and Progenitor Cells Required for Normal Response to Injury. *Stem Cell Reports* **3**, 716–724.
- Zhao, G. Q., Zhao, Q., Zhou, X., Mattei, M. G. and de Crombrughe, B.** (1993). TFEC, a basic helix-loop-helix protein, forms heterodimers with TFE3 and inhibits TFE3-dependent transcription activation. *Mol. Cell. Biol.* **13**, 4505–12.
- Zhou, F., Rasmussen, A., Lee, S. and Agaisse, H.** (2013). The UPD3 cytokine couples environmental challenge and intestinal stem cell division through modulation of JAK/STAT signaling in the stem cell microenvironment. *Dev. Biol.* **373**, 383–393.