



**HAL**  
open science

# **Drosophila intestinal stem cell response to DNA damage and replication stress**

Benjamin Boumard

► **To cite this version:**

Benjamin Boumard. Drosophila intestinal stem cell response to DNA damage and replication stress. Cellular Biology. Université Paris sciences et lettres, 2021. English. NNT: 2021UPSLS052. tel-04193090

**HAL Id: tel-04193090**

**<https://pastel.hal.science/tel-04193090>**

Submitted on 1 Sep 2023

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

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



**THÈSE DE DOCTORAT**  
DE L'UNIVERSITÉ PSL

Préparée à l'Institut Curie

Dans le cadre d'une cotutelle avec Sorbonne Université Faculté des  
Sciences École Doctorale ED515 Complexité du Vivant (CdV)

**Réponses des cellules souches intestinales de Drosophile  
aux dommages à l'ADN et au stress de réplication**

*Drosophila* intestinal stem cell response to DNA damage and  
replication stress

Soutenue par

**Benjamin BOUMARD**

Le 12 Juillet 2021

Ecole doctorale n° 515

**Complexité du Vivant**

Spécialité

**Génomique, Cellule,  
Développement**

Composition du jury :

Agnès, AUDIBERT Professeure des Universités, Sorbonne Université	<i>Présidente</i>
Anne, ROYOU Directrice de recherche, IECB/Université de Bordeaux	<i>Rapporteuse</i>
Julia, CORDERO Senior Research Fellow, University of Glasgow	<i>Rapporteuse</i>
Parthive, PATEL Assistant Professor, University of Bristol	<i>Examineur</i>
Chunlong, CHEN Chargé de recherche, Institut Curie/PSL	<i>Examineur</i>
Allison, BARDIN Directrice de recherche, Institut Curie/PSL	<i>Directrice de thèse</i>

*“Better leave the light on” B.H.*

## REMERCIEMENTS

These last 4+ years in Paris and the completion of this PhD work would not have been possible without the support, help and advice of so many persons.

I want to thank Allison, first, for letting me discover scientific research in her lab in 2014. It was my first lab experience as an undergrad and I *guess* it had a positive enough impact to convince me to come back for my PhD a few years later. I am very grateful that I was free to choose whatever project I was interested in to work on. I am very thankful for Allison's amazing supervision and guidance through the ups and downs of the projects we worked on, as well as her advice on post-PhD career opportunities. I am grateful to have felt supported during these last few years, and to have received validation on scientific and some non-scientific decisions in my life. I am also very thankful to Allison for her never-ending joy about science, and for finding excitement in every little progress made. It was definitely helpful in pursuing the project, even in the most difficult times. Finally, I enjoyed working in the lab very much, I loved the inspiration that we could get from collaborative work on different projects and for the great atmosphere and friendship among team members. Allison has a lot of credit to take for creating such a diverse, inspiring and joyful environment.

I then want to thank Kasia for initiating the projects I have been working on, her work was decisive in the orientation of my PhD project and I am very thankful for all the advice she could give over the years. Kasia was also the first person to supervise me in the lab in 2014, she has a lot personal qualities that make the work with her very interesting, and I think she is a fantastic mentor. I also have to thank her for bringing Polish caramels to the lab, and other polish ~~vodka~~ delicacies.

The PhD work presented below would not have been possible without Nick who developed all the bioinformatics pipeline and conducted the analysis of the sequencing data. Thank you so much for taking the time in answering all my questions regarding the analysis. I would also like to thank Marius here, I hope we

will finally get to see you soon again in Paris. I am looking forward to have a couple of beers at the Bombardier with both of you.

I would like to thank my two best partners in crime in the lab, Lara and Marine, you are both incredible and I could never thank you enough for what you have meant for me. I am so happy I got to share the PhD experience from the beginning with Lara, you taught me so much and you are truly one of the most incredible persons I ever met. You are so inspiring in your words and actions, while being so humble about it, and you are already an incredible role model. I can't wait to see what you will achieve. Thank you Marine for your help in the lab and your incredible spirit, you are such a sweet and caring individual. You are also bringing so much character and determination in making the lab a nice welcoming place.

Of course, I would also like to thank the other essential lab members: Manon and Louis, the epigenetic squad, thanks a lot for your help and comments throughout my PhD. I wish all the best to Manon for the next year, and I am looking forward to seeing you defend. Annabelle, Natalia and Hojun, for perpetuating the lab spirit, you are good a inspiration as to what a good scientist should be and I have so much more to learn from you, (first, let's be honest, regarding planning and organization). I am looking forward to seeing the results of your projects, which I find very exciting. I wish we could have done a lab retreat with all of you.

Finally, I would like to thank Marwa for helping me carry out a few experiments in the last few months. It is a pleasure to supervise you for your master's project, and I wish you a lot of success. The last thank you goes to Gwenn, I did not get to meet you a lot yet, but seeing your interest in growing flowers and plants on my desk, I am sure you can only be an incredible person.

I would like to thank the members of the fly community of the BDD for nice interactions and regular help for reagents. I would like to thank particularly Bénédicte, Nicolas, Lucie et Sabine, for caring for our precious flies. I am also

very grateful to the microscopy platform of the BDD, they helped me to get the best pictures and the best methods to analyze my experiments.

I am grateful to Sarah, Renata and Pauline for accepting to be part of my thesis committee and giving important strategic advice on how to pursue my PhD project. Finally, I would like to warmly thank my mentors in other labs, Ya-Chieh Hsu, Michalis Averof, Michèle Crozatier and Nathalie Vanzo who supervised me for different master internships and guided me since.

I would like to thank the special people I had the chance to meet in Paris and without whom this PhD would have been much more difficult. In particular, Lara, Marine, Manon, Markus, Jacopo, Ana, Francesca, Annabelle, Özge, Cristobal and Mingqiang. Thank you so much for your trust, your humor and your support throughout these years. You are all very special to me, and I hope you know how much you matter to me. You have seen a lot of lows and I think the best version of myself can really express itself with you. I am also very happy I got to spend these last two years with Zheng, I could not express my feelings accurately enough, but maybe the simple words of Montaigne can: “Parce que c’était lui; parce que c’était moi”.

I would like to thank all my friends who have supported me through this experience: Didier, Marianne et Rémy. Agathe, Marion, Lise, Juliana, Ferdinand, Adam, Yannick, et Avril. I hope we will be able to celebrate many more events together.

Pour finir, je tiens à remercier mes parents et mes frères et sœurs pour leur amour et support sans conditions. Une pensée spéciale pour Enzo, Louise, Pauline, Maël et Cyrielle qui nous apportent à tous tellement de joie.

Merci beaucoup à tous avec tout ce qu’il faut d’amour.

## OUTLINE

<i>Chapter 1 : INTRODUCTION.....</i>	<i>8</i>
<b>A. Mutations – A Historical Perspective .....</b>	<b>8</b>
I. Mutations: emerging heritable traits driving evolution .....	9
II. Mutations: mappable variants of heredity .....	11
III. Mutations: inducible by external factors .....	14
IV. Mutations: somatic alteration of the organism .....	15
V. Mutations: the origin of tumorigenesis .....	16
VI. Mutations: molecular alteration of the genome sequence.....	18
<b>B. DNA damage and Repair .....</b>	<b>20</b>
I. Base modifications, replication errors and point mutations.....	21
II. DNA double strand break and choice of repair.....	23
III. Pol $\theta$ and Microhomology–Mediated End Joining .....	26
IV. Replication stress as a cause of DNA damage .....	29
<b>C. Somatic mutations and aging stem cells .....</b>	<b>37</b>
I. Stem cell function and aging .....	37
II. DNA damage/repair and aging.....	40
III. Somatic mutations in stem cells .....	40
<b>D. The Drosophila Midgut.....</b>	<b>42</b>
I. ISC identity and cell lineage decisions .....	42
II. Fine–tuning ISC proliferation.....	49
III. Stem cell long–term maintenance and aging.....	54
IV. Conclusion .....	57
<b>E. Contribution of this PhD work .....</b>	<b>58</b>
<i>Chapter 2 : STEM CELL MUTATIONS AND POL THETA CONTRIBUTION TO MUTAGENESIS.....</i>	<i>60</i>
<b>Characterization of Drosophila intestinal stem cell mutations and Pol Theta     contribution to mutagenesis. ....</b>	<b>61</b>
I. INTRODUCTION .....	61
II. Results.....	62
III. DISCUSSION.....	82
IV. MATERIALS AND METHODS .....	85
V. SUPPLEMENTARY FIGURES .....	90
VI. REFERENCES:.....	96
<i>Chapter 3 : A TISSUE–SPECIFIC BUFFERING MECHANISM OF REPLICATION STRESS .....</i>	<i>99</i>
Closing the GAP on replication stress:.....	100
<b>Tissue–specific buffering of nucleotide levels protects from replication stress     .....</b>	<b>100</b>
I. INTRODUCTION .....	100

II. RESULTS .....	102
III. DISCUSSION.....	117
IV. MATERIALS AND METHODS .....	119
V. SUPPLEMENTARY FIGURES .....	123
VI. REFERENCES .....	129
<i>Chapter 4 : DISCUSSION AND PERSPECTIVES.....</i>	<i>132</i>
I. On replication stress and genome instability in ISCs .....	133
II. Genome fragile sites and chromosomes rearrangements.....	136
III. On the functions of gap junctions .....	142
IV. Final statement.....	145
V. REFERENCES .....	145
INTRODUCTION–REFERENCES .....	149

**FIGURE LIST:**

Figure 1. An example of a new character identified in the Morgan’s lab: Beaded flies had altered wings.....	10
Figure 2. Diagram representing the pairing, size relation and shapes of Drosophila chromosomes. ....	11
Figure 3. A representation of Crossing over between two homologous chromosomes. From .....	12
Figure 4. Examples of tumor in Drosophila.....	17
Figure 5. Chronology of some of the theories and discoveries on mutations. e. ....	19
Figure 6. DNA Double–Strand Breaks main repair pathways with the factors they rely on. ....	24
Figure 7. Causes of replication stress and Consequences of replication fork collapse .....	31
Figure 8. Models of Notch signalling and cell fate decisions in the ISC lineage.. ....	44
Figure 9. Lineage regulators of intestinal stem cell differentiation towards enterendocrine and enterocyte cells. ....	46
Figure 10. Cell autonomous and Non–cell autonomous control of proliferation in the midgut. ....	50
Figure 11. Steroid and peptide hormone regulation of gut physiology.. ....	53
Figure 12. Model of the Tus/Ter system with Ter inserted before Notch .....	139

Here are only listed the figures of the Chapter 1 and 4. The Chapter 2 and 3 are constructed as individual articles with their own figure numeration. The Supplementary figures of each articles are at the end of each chapter, before the references.

## List of abbreviations:

**DSB** – Double–Strand Break  
**ssDNA** – Single–Strand DNA  
**BER** – Base Excision Repair  
**NER** – Nucleotide Excision Repair  
**MMR** – Mismatch Repair  
**NHEJ** – Non–Homologous End Joining  
**HR** – Homologous Recombination  
**SDSA** – Synthesis Dependent Strand Annealing  
**BIR** – Break Induced Replication  
**LOH** – Loss of Heterozygosity  
**Alt–EJ/Alt–NHEJ** – Alternative Non–Homologous End Joining  
**MMEJ** – Microhomology–Mediated End Joining  
**TMEJ** – Pol Theta–Mediated End Joining  
**CFS** – Common Fragile Sites  
**MMBIR** – Microhomology–Mediated Break Induce Replication  
**FoSTeS** – Fork Stalling and Template Switching  
**MiDAS** – Mitotic DNA Synthesis  
**TE** – Transposable Element  
**SV** – Structural Variant  
**SNV** – Single Nucleotide Variant  
**CNV** – Copy Number Variant  
**ROS** – Reactive Oxygen species  
**UV** – Ultraviolet radiation  
**EMS** – Ethyl–Methyl Sulfate  
**Pol $\theta$ /PolQ/mus308** – DNA Polymerase Theta  
**ISC** – Intestinal Stem Cell  
**EE** – Enteroendocrine cell  
**EEP** – Enteroendocrine Precursor cell  
**EB** – Enteroblast  
**EC** – Enterocyte  
**MSC** – Melanocyte Stem Cell  
**HSC** – Hematopoietic Stem Cell  
**Rnr** – Ribonucleoside reductase  
**RnrL** – Ribonucleoside reductase Large subunit  
**HU** – Hydroxyurea  
**Inx** – Innexin  
**GFP** – Green Fluorescent Protein  
**RFP** – Red Fluorescent Protein

# CHAPTER 1 : INTRODUCTION

A mutation is a modification of the genome sequence. It is acquired upon alteration of the DNA molecular composition or structure, so-called DNA damage, and the absence of repair or usage of erroneous DNA repair pathways. The DNA molecule can be altered in various ways, from base modification to DNA double strand break. As a consequence, the resulting mutations can affect a single gene through an indel or single nucleotide variant, or potentially multiple genes and chromosomes via structural variant (i.e., deletion, inversion or translocation) or chromosome rearrangement.

The factors causing DNA damage, the DNA repair pathways involved, and the types of mutation generated, have been extensively studied since before the discovery of the DNA structure. The purpose of this chapter is to summarize our knowledge of the mechanisms driving mutation accumulation and the consequences of such mutations on tissue homeostasis. I will also highlight the interrogations that remain to be addressed, which constituted the focus of this thesis.

*The references for the introduction are listed at the end of the manuscript*

## **A. Mutations – A Historical Perspective**

Although this PhD work focuses essentially on somatic mutations, the ideas behind mutagenesis emerged in the context of speciation and heredity<sup>1</sup>. In this part, I first wanted to briefly introduce the ideas and experimental work that led to the study of mutations, the establishment of genetics as a field of study and the questions surrounding tumorigenesis in the context of mutation acquisition. Importantly, the contribution of *Drosophila* genetics has revolutionized the way we tackle and understand the inheritance of traits, mutagenesis and tumorigenesis<sup>2,3</sup>.

## I. Mutations: emerging heritable traits driving evolution

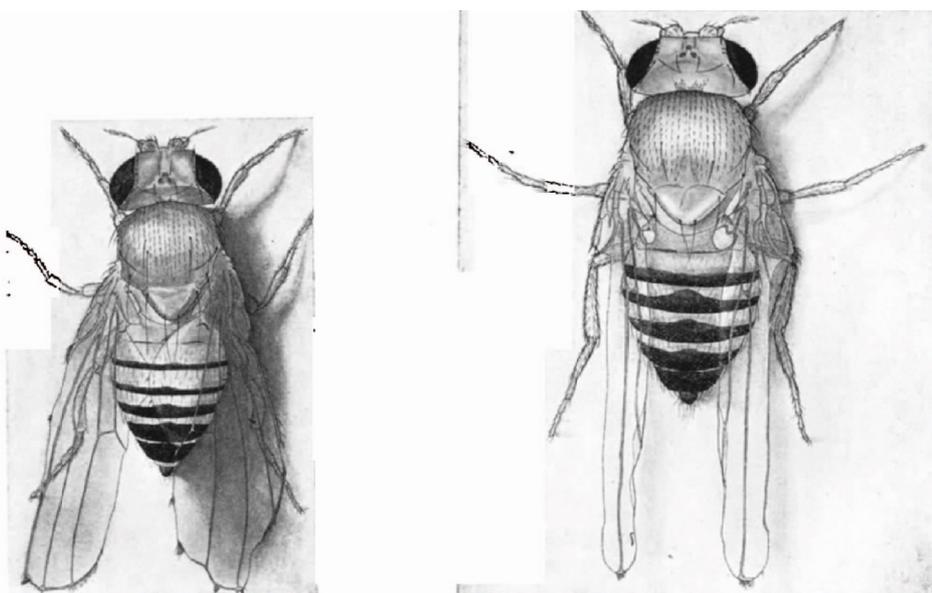
The theory of evolution by natural selection as stated in C. Darwin's publications, most notably in "On the Origin of Species", suggested the emergence of subtle variation of traits in a population. These fluctuating variations are continuous and likely present at any time in the population, but by the means of environmental pressure, reproductive isolation and the survival of the fittest, some variations of these traits can become preeminent, ultimately driving speciation. Opposed to Darwin, F. Galton's views sustained that brutal, dramatic changes can occur to drive speciation in a single generation.

Darwin also proposed a theory of heredity, the "pangenesis", by which traits could be transmitted from one generation to another through pangenes, but it lacked experimental evidence. In 1893, A. Weisman proposed the germplasm theory of heredity, suggesting the continuity of the germplasm located in the eggs and sperms to transmit traits between generations, and the discontinuity between germplasm and the somatoplasm. Importantly, any somatic modification of the organism (e.g. mutilation) would not be transmitted to the next generation.

The rediscovery of Mendel's theory of heredity in 1900 was followed by the study of inheritance in many species. It promoted the emergence of the field of genetics initiated in part by W. Bateson and W. Johannsen who proposed the main terminology: "genes" are the undefined units of transmission and "alleles" represent different variants of the same trait. In 1901, H. De Vries, who worked on the evening primrose *Oenothera Lamarckiana*, on which he followed the variation and emergence of "characters", or phenotypes, among the populations, witnessed sudden changes in different traits, some of which seemed to drive speciation within one generation. To explain the emergence of new characters, in "The Mutation Theory", De Vries named "mutations", the new forms that suddenly appeared between generations, and proposed that they drive speciation: "The

main principle of the mutation theory is that species and varieties have originated by mutation, but are, at present, not known to have originated in any other way.”<sup>4</sup>

De Vries experiments caused skepticism because they were in support of Galton’s view of the discontinuity of the evolution, but also initiated wide interest among biologists and geneticists. It appeared possible to follow the emergence of new characters through successive generations, and by doing so, to identify the mechanisms of speciation. T. H. Morgan, who had established his lab at Columbia in 1904, refocused to study the appearance of new characters in *Drosophila melanogaster* in 1907, encouraged by W. Castle<sup>2</sup>. The advantages of the fruit fly are numerous to study traits variation and heredity: they have a short life cycle (10 days) allowing for the succession of many generation in a short time period; they are small and can be bred easily in the lab and they produce a large amount of progeny. The first 3 years of work appeared quite disappointing in the search for new characters, as Morgan and his students could only characterize small variations of the pigmentations. However, in 1910, “in a pedigree culture of *Drosophila* which had been running for nearly a year through a considerable number of generations, a male appeared with white eyes. The normal flies have red brilliant red eyes”<sup>5</sup>. The new character identified was named “white”, and many more followed (Figure 1).

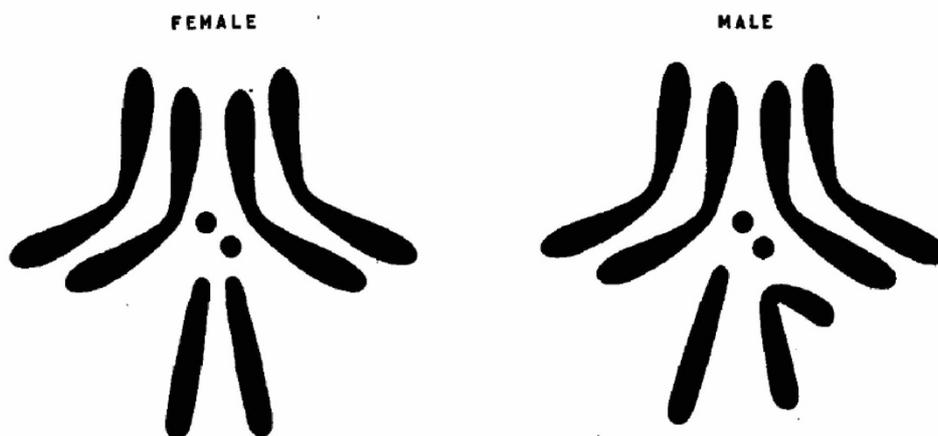


**Figure 1. An example of a new character identified in the Morgan’s lab: Beaded flies had altered wings. From Muller 1918.**

The new characters identified in Morgan's lab followed Mendelian rules of inheritance; however, single-step speciation events were never found. It appeared later that the "speciation" events identified by De Vries were mischaracterized. However, the term "mutation" persisted to design the modification of a gene, responsible for the emergence of new alleles. By the late 1910s, the Darwinian theory of evolution and the Mendelian theory of heredity were reconciled in a "Modern Synthesis". The combination of several characters (alleles) could support continuous variation within populations, and Darwin's natural selection would work on Mendelian units of heredity (genes) to drive speciation. These conclusions were supported by experiments in several model systems such as oats, wheat (Nilsson-Ehle<sup>6-8</sup>), maize (E. Easton 1910), mice<sup>9</sup> and *Drosophila*<sup>10</sup>.

## II. Mutations: mappable variants of heredity

While geneticists worked on theoretical carriers of heredity, which they named genes. In parallel, the work of cytologists helped to identify the chromosomes as the carriers of the material responsible for the heredity of characters. First, in 1902 by studying meiotic divisions, W. Stutton established that the constituents of the cell nucleus, the chromosomes, were responsible for the heredity. He observed at the same time as T. Boveri, that the chromosomes behaved as Mendelian factors during the meiosis, as they differentially segregated in daughter cells and progeny. Then, in 1905, E.B. Wilson and N.M. Stevens



**Figure 2.** Diagram representing the pairing, size relation and shapes of *Drosophila* chromosomes. From Bridges, 1916.

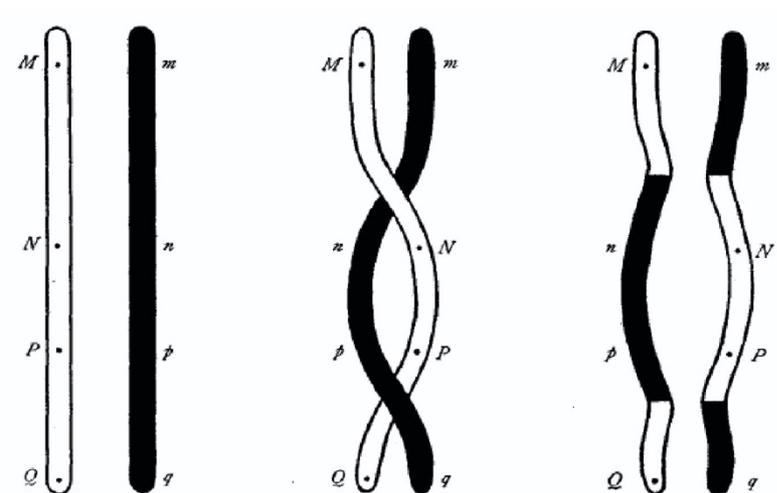
established independently, that a specific set of chromosomes were responsible for the sex determination in different organisms. N. Stevens, named the pair of chromosomes X and Y, and suggested in the beetle *Tenebrio molitor* and several other insect species, including *Drosophila* in 1908, that the Y chromosome was carrying the male determinants (Figure 2).

In the fly, Morgan could then determine that the *white* mutation was linked to a recessive character localized on the X chromosome. Because of the specific importance of the X chromosome in males, which only carry one; mutations on the X are easily detectable in males. It particularly helped H.J. Muller in the late 1910s to identify several lethal mutations on the X chromosome simply by following the female to male sex-ratio in the offspring<sup>11,12</sup>. Indeed, a male carrying a lethal mutation on the X chromosome would die, while a female would still have a non-lethal allele of the gene on the homologous chromosome.

After Morgan localized this first X-linked mutation, with the students in his lab, he could identify a few more like *rudimentary* and *yellow*, respectively responsible for a wing malformation and a change in the body color. From a series

of crosses with female flies carrying two different mutations, one on each X chromosome, they would observe that, different than the expected Mendelian ratio dictating that the progeny would inherit one of the two mutations, a variable fraction of flies

could also inherit both mutations or not receive any<sup>13,14</sup>. From these observations, Morgan postulated the probable exchange of mutations between chromosomes during meiosis, when the two homologous chromosomes were observed to be



**Figure 3. A representation of Crossing over between two homologous chromosomes.** From Weinstein 1918.

“strangely” intertwined before separation. He imagined the theoretical mechanism of crossing over by the end of 1911<sup>15</sup> (**Figure 3**).

Morgan and A. Sturtevant came to imagine that the probability of mutation exchange between the two chromosomes would depend on the distance between the two mutations on the presumed linear chromosome, this led them to create the first map of the X chromosome based on the crossing-over probability between 6 different mutations previously identified<sup>16</sup>. The crossing-over relationship between the mutations confirmed a somewhat linear organization of the genes along the chromosome, and the exchange frequency during meiosis could be interpreted as the distance between genes. Interestingly, if two mutations affecting the same trait were showed to be localized at the same site relative to the others, they could thus be interpreted as being two different alleles of the same gene. This new ability to identify mutations and localized them linearly on the chromosome revolutionized the field of genetics. It was used later in different species, for example in Maize by B. McClintock<sup>17,18</sup>, and rabbits by W.E. Castle<sup>19</sup>. For more than a century, geneticists have been using the same methods to identify and localize new mutations in different chromosomes based on crossing-over frequencies.

The development of cytology also allowed the direct observation of chromosomes and the establishment of physical chromosome maps. In addition, the correlation between the recombination of genetic traits and the physical crossing-over between homologous chromosomes during meiosis was established by McClintock and H. Creighton, who compared cytological and genetic maps of chromosomes<sup>18,20,21</sup>. The same conclusions were found in *Drosophila* soon after by C. Stern<sup>22</sup>.

Thus, the study of mutations allowed a better understanding of the mechanisms of inheritance in many species. It permitted the representation of genes compared to one another on chromosome maps. Additionally, it also helped to theoretically examine the question “What is a gene?”, and the

relationship between the gene and the heritable characters of the organism<sup>23,24</sup>. However, the identification of variants was tedious, as spontaneous mutants appeared infrequently. From the start in 1907, it took 3 years of experiments to isolate the *white* mutant, and about 50 X-linked mutations were characterized by 1916 in *Drosophila*<sup>25,26</sup>.

### III. Mutations: inducible by external factors

In order to increase the susceptibility of *Drosophila* strains to acquire *de novo* heritable mutations, H.J. Muller explored the possibility of using external factors to artificially induce more alterations. He established that X-ray irradiation was significantly increasing (x150) the rate of hereditary mutation in *Drosophila*<sup>27-29</sup>, thus external factors could indeed influence the mutation frequency. Similar experiments had been performed before in different species, since the discovery of X-ray in 1895. These had linked X-ray with fertility problem, fragmentation of the chromosomes (Bergonié and Tribondeau, 1904; Perthes 1904) and other clinical manifestation such as ulceration and cancer. It was also observed that actively growing tissues are the most sensitive to X-rays. Earlier attempts to increase the mutation rate with X-ray, for example in Morgan's lab, were unsuccessful. However, previous reports had already showed that X-rays could influence the crossing-over frequency in *Drosophila*<sup>30-34</sup>. At the same time, Muller also established the influence of temperature on mutation frequency<sup>28</sup>.

Shortly after Muller's findings, the mutagenic effect of X-rays were also shown in other species such as barley<sup>35</sup>, and mice<sup>36</sup>. However, the type of mutations induced by X-ray were not clearly defined. Muller first thought that the mutation increase was mostly due to point mutations affecting directly a single trait. However, it became clear later, with genetic crosses and cytological techniques, that X-rays also induced larger mutations affecting the chromosome structure such as inversions, deletions and translocation<sup>24,29,35,37-42</sup>. Such chromosomal aberrations had already been observed in the spontaneously arising

mutations<sup>43-45</sup>. Although X-rays induced the same type of mutations as the naturally occurring ones, Muller posited that natural radioactivity was unlikely to explain the emergence of spontaneous mutations<sup>46,47</sup>. Thus, he hypothesized the existence of other factors responsible for mutagenesis. Experiments on the consequence of several other external factors on mutation frequency and inheritance were carried out in *Drosophila* and other models. New physical and chemical mutagens were identified: for example radium<sup>35</sup>, formaline (Rapoport 1946), formaldehyde (Rapoport 1947,1948)<sup>48</sup>, ultraviolet light (UVs), nitrogen-mustard<sup>49,50</sup>, ethyl-methyl sulfate (EMS)<sup>51</sup> and other forms of radiation<sup>52,53</sup>, also producing point mutations and chromosome rearrangements<sup>54-56</sup>.

Thus, mutations can be artificially induced. Various mutagenic agents, likely present in the environment, can provoke hereditary trait changes. This possibility to induce mutation was then used to characterize and location new genes in classic genetic experiments and later led to the genetic screens methodology, for example using the alkylating agent EMS<sup>57</sup>.

#### **IV. Mutations: somatic alteration of the organism**

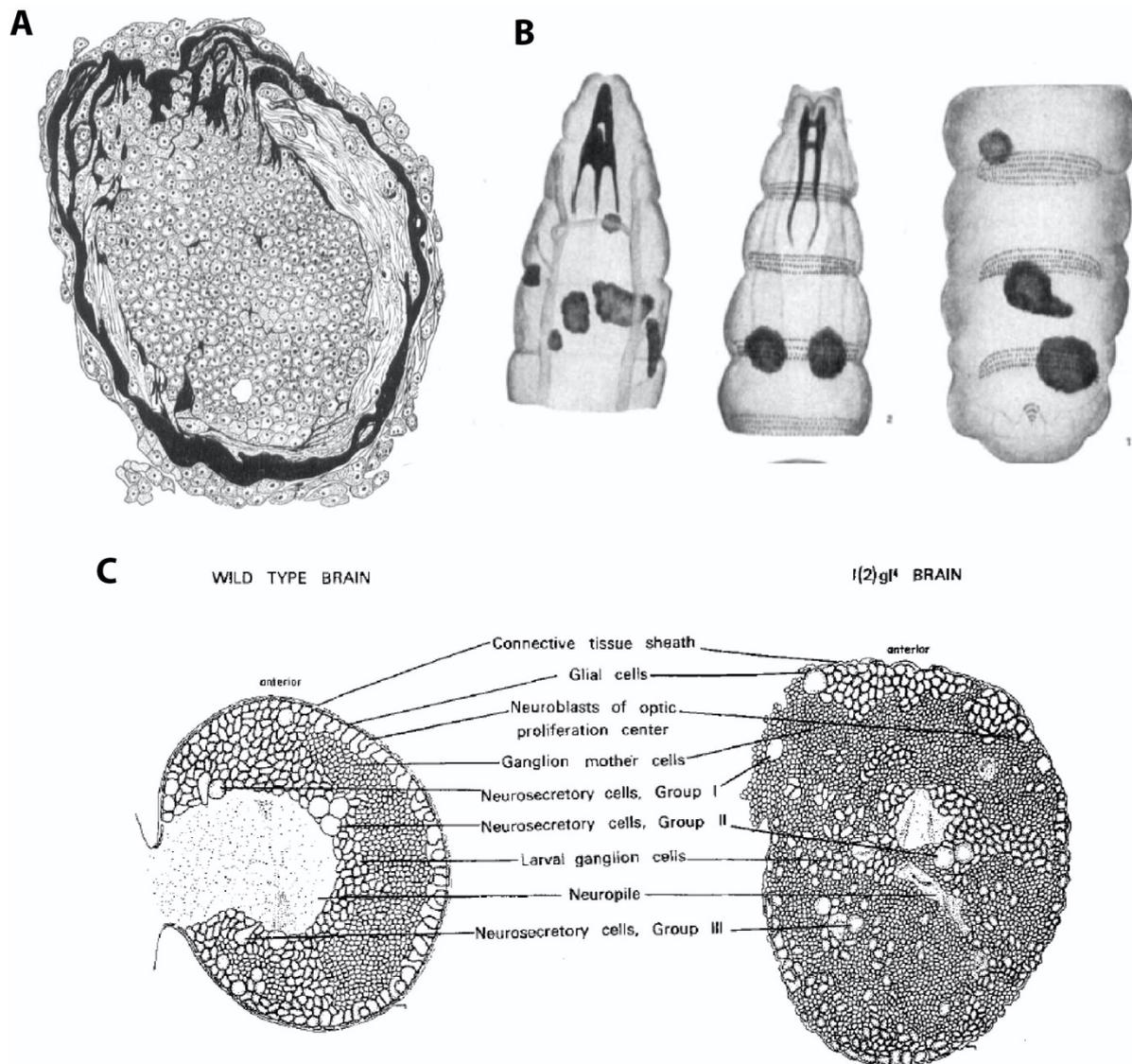
In addition to the identification of new heritable traits through mutation in the germ cells, some instances of mosaicism of the somatic tissues were reported. Mosaicism is seen as the coexistence in the same individual of two characters of the same trait in the tissue. This was notably observed in the variegated patterns of maize plants<sup>58,59</sup>. Some instances of mosaicism were also observed in *Drosophila*<sup>22</sup> and other species<sup>60</sup>. Importantly, X-rays and other mutagens were shown to induce tissue mosaicism<sup>61-64</sup>. The mosaicism was explained in part by somatic crossing-over of heterozygous traits in mitotic cells<sup>65</sup>, although *de novo* alteration of the gene in the somatic cells was also possible. It was later shown that variegation could also be due to moving DNA elements in the genome<sup>66</sup> or to non-genetic factors influencing gene expression, underlying a position effect on genes.

## V. Mutations: the origin of tumorigenesis

The implication of malignant tumors in human diseases was established in the 18<sup>th</sup> century. The characterization of cancer neoplasms was further carried out with new cytological methods in the end of the 19<sup>th</sup> century. It was postulated, for example by G. Pianese and D. Hansemann, that tumors were due to changes in cell behavior, and more specifically linked to a uncontrolled cell proliferation<sup>67,68</sup>. In 1914, T. Boveri was the first to hypothesize that cancer cell behavior was linked to chromosome alteration, as he observed that aneuploid cells divide atypically<sup>69</sup>. The development of *Drosophila* genetics yielded important findings related to Boveri's theory<sup>3</sup>:

First, a fraction of the larvae of *lethal(1)7* (known now as *deep orange, dor*), one of the first mutant characterized in Morgan's lab, spontaneously developed melanized tumors, likely responsible for the death of the animal before the adult stage<sup>25,70</sup> (**Figure 4A–B**). The progeny of the surviving flies also displayed the same phenotype in relatively similar proportions. It was the first evidence of hereditary tumors in *Drosophila*. Thus, the carriers of heredity, the chromosomes, contained mutations driving tumorigenesis. M.B. Stark, I. T. Wilson and later E.S. Russel, continued to study hereditary tumors in this model and identified several mutations responsible for neoplastic growth in the fly<sup>71–73</sup>. The spontaneous tumors arising in susceptible lines could affect different organs, including the intestinal track. In addition, spontaneous neoplastic growth was observed in different invertebrates species<sup>74–79</sup>. However, cancer development was more frequently observed in mammals, including humans, and the heredity of tumor incidence was also studied in different strains of mice<sup>80,81</sup>. Altogether, it suggested, that cancer was, in part, a hereditary disease, associated with particular variants.

The study of mutations that increased cancer susceptibility led to the description of several tumor-suppressor genes in *Drosophila*. The first one named *lethal (2) giant larvae (l(2)gl)* was characterized in 1967 by E. Gateff (Figure 4C). It eventually instated *Drosophila* as a good model system to study cancer-related questions, such as oncogene driven proliferation, drug screening, tumor metabolism and genome instability<sup>82,83</sup>.



**Figure 4. Examples of tumor in *Drosophila*.** **A.** Structure of gereditary tumors in *Drosophila*, accumulation of cells surrounded by dark spots. **B.** Representation of Hereditary tumors in the *Drosophila* larva. **C.** Comparison between wild-type brain *l(2)gl<sup>A</sup>* tumor brain in *Drosophila* larvae. Taken from Stark. 1918(A-B); Gateff and Schneiderman 1974 (C).

Additionally, it also appeared that external mutagens described before increased the frequency of cancer in flies and mammals. Thus, this led to the notion that somatic mutagenesis is a driver of tumorigenesis<sup>84,85</sup>. The influence of

mutagens on tumor incidence was also inferred from reports on the health conditions of patients after irradiation or in frequent contact with identified mutagens<sup>1</sup>. In the long term, they were more likely to develop cancers, and at a younger age. The link between mutagenesis and carcinogenesis was then established. Finally in 1950–1975s, as proposed decades before by T. Boveri, karyotype studies demonstrated the prevalence of chromosome rearrangements or ploidy alteration in human cancer cells<sup>86</sup>. Consequently, genome alteration drives tumor development and subsequent genome instability is a hallmark of cancer<sup>87</sup>.

## VI. Mutations: molecular alteration of the genome sequence

The identification of genes and mutations responsible for different variants linked with tumorigenesis, was achieved while the chemical and physical nature of the chromosome was unknown. The characterization of nucleic acids then allowed the description and classification of different types of mutations as well as an understanding of the specific effects of mutagens on DNA composition and structure.

Briefly, biochemical studies in the mid 40s to early 50s first identified that chromosomes were composed, in part, of deoxyribonucleotide acid (DNA), and DNA was found to carry the genetic material in bacteria, and viruses<sup>88,89</sup>. In the 1950s, the molecular composition of DNA was better characterized via chemical analysis<sup>90–92</sup>, and the double-helix structure<sup>93</sup> was postulated from X-ray crystallography experiments carried out by R. Franklin.

With the identification of the base composition of the DNA, the effects of mutagens, such as EMS and other alkylating agents, on DNA modifications and mutation acquisition was chemically established<sup>94–99</sup>, which advanced the theory of mutagenesis<sup>100</sup>. Finally, the deciphering of the genetic code<sup>101</sup> and of mechanisms regulating genes expression<sup>102</sup> in the early 1960s, helped to comprehend the consequences of base modifications on genic alteration, and to

connect gene sequence and expression with the phenotypes induced by mutations. However, the extensive description of mutations at the molecular level would later require the use of DNA sequencing technologies, which have been developed and improved since the 1970s<sup>103-106</sup>.

The development of molecular biology allowed the description of mutations at the molecular levels and also enabled the characterization of the consequences of mutations on gene function, consequently explaining the phenotypes observed since the 1910s (Figure 5).

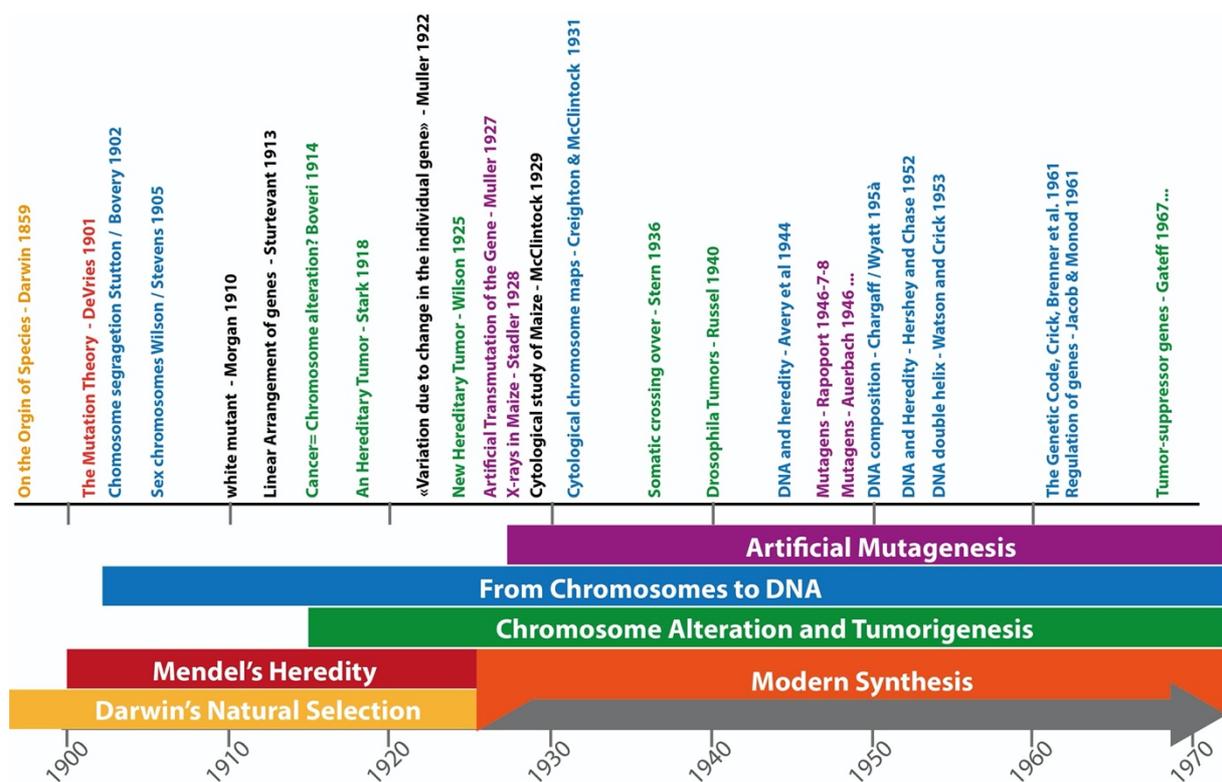


Figure 5. Chronology of some of the theories and discoveries on mutations. Details are in the text above.

## B. DNA damage and Repair

I presented previously how mutations of the genome were first identified and how different compounds could induce mutagenesis. With the discovery of the composition and structure of DNA, DNA damage, mutations and the consequences of mutagens could be assessed at the molecular level. Since DNA is the conserved carrier of genetic material in all living organisms, most species is subject to similar types of DNA damage, and the DNA repair mechanisms that have evolved to repair such alterations are well conserved, particularly in eukaryotes.

Many exogenous factors inducing DNA damage have been discovered, such as ionizing radiation, Ultraviolet (UV) radiation, nitrogen–mustard, formaldehyde and bacterial toxins. The specific exposure of different tissues to factors from the environment is responsible for many of the differences in the mutation landscape and tumor burden in distinct organs<sup>107,108</sup>. However, endogenous sources of DNA damage have also been identified: (1) Products of the cell metabolism like S-adenosylmethionine falsely methylate DNA, whereas reactive oxygen species (ROS) and aldehydes can oxidize or crosslink DNA bases, respectively. (2) Enzymes like topoisomerases or endogenous transposases can induce DNA double strand breaks (DSB). (3) Replicative or translesion DNA polymerases can incorporate mismatched bases through replication error. (4) Finally, replication stress can lead to replication fork collapse and DNA DSB (see below).

Here I first present briefly the DNA repair pathways involved in repairing bases or single strand nick, and the main factors involved in DSB repair. Then I detail two processes that have recently gained interest in the DNA repair and cancer fields: the Microhomology Mediated End–Joining (MMEJ) pathway and the contribution of replication stress in DNA damage and gross chromosomal rearrangements. We hypothesized that both may act in intestinal stem cells, thus they constitute the primary focus of my PhD thesis.

## I. Base modifications, replication errors and point mutations

Several pathways have evolved to repair DNA damage affecting DNA single strands or single bases, for example chemical modifications of the canonical bases A, T, C and G. The failure to repair such damage before the following DNA replication and division leads to point mutation acquisition in the daughter cell.

### Base excision repair

The base excision repair (BER) pathway repairs chemical modifications of bases that do not alter the DNA double helix structure: single strand break, alkylation, deamination, oxidation (for example guanine to 8-oxoguanine conversion, induced by ROS), methylation, loss of bases (mostly depurination caused by spontaneous hydrolysis)<sup>107-109</sup>.

BER removes the modified base via DNA glycosylase activity, and repairs the created nick in the single strand by adding the absent nucleotide or resynthesize a short stretch of DNA if several bases were missing/removed. Each step depends of a different enzyme.

### Nucleotide excision repair

The nucleotide excision repair (NER) pathway repairs chemical modifications of bases that distort the helix, for example: pyrimidine dimers induced by UVs, DNA crosslinked adducts or intrastrand crosslinks, oxidized purine with a cyclic conformation or non-canonical DNA structure like R-loops and G-quadruplexes<sup>107-110</sup>.

NER excises a short single-strand fragment of DNA surrounding the damage and resynthesizes the DNA sequence. The excision step is carried out by ERCC1, ERCC4 and ERCC5 best known as XPF and XPG, which are mutated in patients with Xeroderma Pigmentosum. These patients are highly sensitive to UVs, because they are unable to repair UV-induced DNA damage.

### *DNA replication error and translesion synthesis and Mismatch repair*

The main DNA polymerase for replication in eukaryotes are Pol $\alpha$ , Pol $\delta$  and Pol $\epsilon$ . Pol $\alpha$  primes the replication and Pol $\delta$  mostly participates in the elongation of the leading strand of DNA whereas Pol $\epsilon$  elongates the lagging strand, although Pol $\delta$  has also been shown to contribute to synthesis of the lagging strand as well. DNA replication is mostly error-free with an error rate of  $10^{-10}$ . Replicative polymerases have a high selectivity for good nucleotide pairing and also have 3'-5' proofreading activity that allows them to remove mis-incorporated nucleotides. The replication complex is also tightly linked to the mismatch repair (MMR) pathway, which is described below.

Replicative polymerases, however, are easily stalled on repeated DNA, or DNA with secondary structures or DNA lesions, such as single-strand breaks or interstrand crosslinks. In these contexts, some translesion polymerases have evolved to allow DNA replication through damaged DNA<sup>111</sup>. For example, Pol $\eta$  is able to accurately replicate DNA with thymidine dimers induced by UV, as such Pol $\eta$  is often found mutated in some Xeroderma Pigmentosum patients. Pol $\eta$  is also able to replicate through modified dG, for example intrastrand dG-dG dimers or 8-oxoG). Because of their functions in replication at DNA lesions, the translesion polymerases interact with the components of NER and BER. Pol $\kappa$ , Pol $\iota$  and Pol $\eta$  cooperate with NER in the repair of UV induced damage. The polymerase Rev1 is implicated in synthesis over abasic sites and bases with adducts. It has also been shown to replicate region of DNA forming G-quadruplexes or with triplet repeats.

Although replicative polymerases are blocked at DNA damage sites, translesion polymerases can proceed over lesions, but they are also significantly error-prone, generating more mismatches than the canonical replicative polymerase. For example, the error rate for Pol $\zeta$  is  $10^{-4}$ . Translesion synthesis generates a considerable number of mismatches. The mismatch repair pathway recognizes base-base inaccurate pairing, and small indels (insertions or deletions)

generated during canonical or translesion replication<sup>112</sup>. Like BER and NER, it involves several components, that are mostly conserved between prokaryotes and eukaryotes, recognizing specific lesions and handling the base excision and resynthesis. Mutations in some components of MMR are associated with colorectal cancer predisposition, such as MSH2, MSH6 or MLH1<sup>113</sup>.

Living organisms have established different strategies to repair DNA damage affecting the molecular base constitution, and defects in any of the pathways drives accumulation of point mutations and is linked with cancer predisposition. The predispositions are tissue-specific and related to the type of damage that particular organs can encounter, for example the skin is likely more exposed to UV and defects in NER and Pol $\eta$  are associated to skin cancer predisposition. Interestingly, inhibitors of these pathways can also be used in cancer treatments to make tumor cells more susceptible to DNA damage and provoke their apoptotic death<sup>114</sup>.

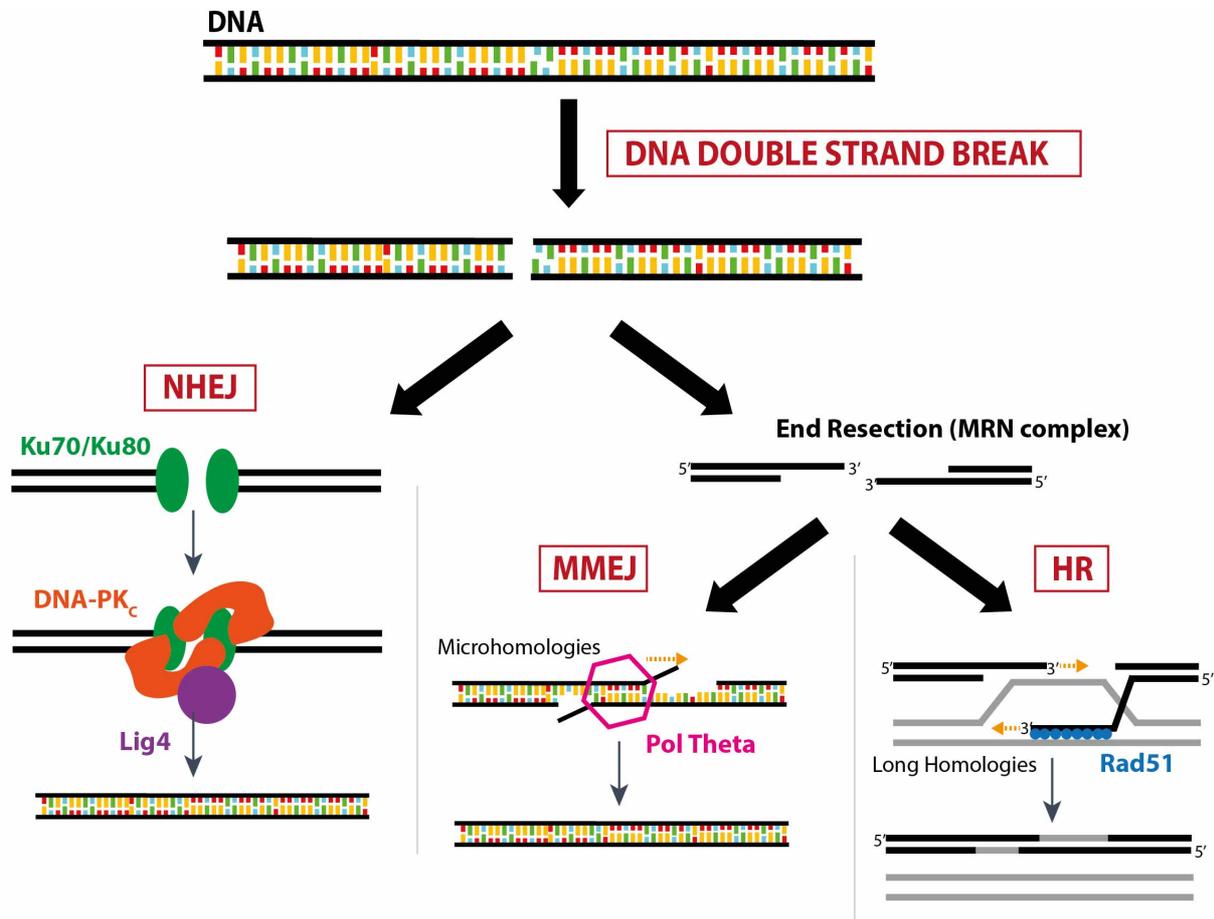
## II. DNA double strand break and choice of repair

The most deleterious insult to the DNA is DNA double-strand break (DSB), as it breaks the double-helix structure and fully interrupts the genome sequence. DSB can be induced by ionizing radiation or chemical reagents, or can be a consequence of replication fork collapse. Similar to the base modifying type of DNA damage, several mechanisms have evolved to repair a DNA DSB<sup>107-109,115</sup> (Figure 6):

### *Non-Homologous End Joining (NHEJ)*

NHEJ ligates together two blunt double strand DNA break ends. It relies on the binding of the Ku70-Ku80 heterodimer to the DSB ends, the recruitment of DNA-PK $\epsilon$  that promotes the close alignment of the broken ends and a final step of ligation dependent on Ligase 4.

NHEJ is mostly considered error prone as it can lead to the deletion of the sequence between the DNA breaks. It has also been involved in larger chromosome catastrophe events such as chromothripsis, responsible for large



**Figure 6. DNA Double-Strand Breaks main repair pathways with the factors they rely on.**

scale translocations, deletions and inversion, by ligating non-adjacent breaks from severely damaged chromosomes<sup>116</sup>.

### Homologous Recombination (HR)

The HR pathway is considered a more conservative DNA repair as it uses a homologous template to repair the DNA double-strand breaks. The DSB generates two double strand break ends. The repair first requires the resection of the 5' strand of the DNA break by the MRN complex (MRE11, RAD50, NBS1). The 3' ssDNA overhang is then bound by Rad5, which participates in the homology search. Finding of a >50bp homology sequence, likely from the sister chromatid or the homologous chromosome allows for an error-free repair<sup>117</sup>. The resolution

of the HR repair can proceed through different pathways, which are Crossing-Over, Synthesis-Dependent Strand Annealing (SDSA) or Break-Induce Replication (BIR, mostly depending on genomic context and homology length<sup>117</sup>.

- Crossing over requires the second strand of the DNA break to also be paired to the homologous sequence. It involves the action of a resolvase to disentangle the double Holliday-junction formed between the two DNA helices, and can lead to sequence exchange between broken DNA and its homologous template.
- SDSA relies on partial *de novo* synthesis of DNA to complete the broken sequence using the homologous template, before the strand is displaced to the second break end with homology search. The synthesis part of the repair can be quite extensive leading to gene conversion whereby a sequence of the homologous template is inherited by the now repaired DNA sequence.
- BIR takes places mostly in the context of single-ended double strand breaks, for example in the context of replication fork collapse. It depends on the complete *de novo* synthesis of the repaired sequence using the homologous template. This process is known to be particularly error-prone and relies on the Pif1 helicase and Pol32/PolD3, a subunit of the replicative polymerase Pol $\delta$ .

### Choice in the DSB repair

Several contexts and factors favor one pathway over the other for DSB repair<sup>115,118</sup>. NHEJ and HR relies on unrelated protein complexes and the mechanisms differ from the first step. NHEJ is a faster but error-prone repair pathway. HR requires extensive 5'→3' resection for homology search while NHEJ cannot process and repair DNA breaks with 3' overhangs larger than 4bp. Thus, the type of lesion and end resection dictate the pathway choice.

In addition, the cell-cycle phase influences the DNA repair pathway choice. NHEJ is predominant in G1 or noncycling cells and HR is preferred in S and G2

phase, when a sister chromatid is available. This choice is controlled in part by cyclin-dependent kinase that can phosphorylate EXO1 exonuclease and CtIP, which positively regulates MRE11 involved in DSB end resection and thus favoring HR during S and G2.

Several studies also suggest an influence of the chromatin context in the DNA repair pathway choice<sup>115,118</sup>, and age-related changes in the DNA repair choice and efficiency<sup>119</sup>.

### III. Polθ and Microhomology-Mediated End Joining

#### Another pathway for DSB repair

In addition to the canonical NHEJ and HR DSB repair pathways, evidence for an Alternative Non-Homologous End Joining (Alt-NHEJ) mechanism was uncovered<sup>120</sup>. In mammalian cells and yeast, a specific type of DNA repair junction was described, it contained short microhomologies at found the breakpoints and inserts from short, templated sequences at the site of the DSB repair event<sup>121,122</sup>. This mechanism was implicated in several types of chromosome alterations such as deletions<sup>123-126</sup>, translocation<sup>127-129</sup>, inversions or more complex events<sup>130</sup>, which are found in leukemia<sup>131,132</sup> and other cancers<sup>133-135</sup>. Alt-NHEJ also participates in the generation of antigen diversity in the V(D)J class-switch recombination of immune precursors<sup>136-139</sup>

Alt-NHEJ functions independently of the canonical NHEJ pathway components Lig4 and Ku70-Ku80<sup>122,140-144</sup>, and appears more mutagenic<sup>134,142,145</sup>. Because of the characteristics described above, this DSB repair pathway is termed Microhomology-Mediated End Joining (MMEJ), we will use this term.

#### MMEJ in yeast

In yeast, MMEJ functions with 5-25 bp of microhomologies. As mentioned before, it is independent of and repressed by the Ku complex<sup>140,146</sup>. MMEJ is a synthesis-dependent DNA repair mechanism. In yeast, it relies on the two DNA

polymerases Pol $\lambda$  and Pol $\delta$ <sup>147</sup> and implicates the non-necessary subunit of Pol $\delta$ , Pol32/POLD3<sup>148-150</sup>. Some components of the HR pathway were found to be required for the repair. First, the MRN-dependent DNA 5' end resection is essential to reveal homologies<sup>149,151</sup>. Furthermore, some reports suggest an implication of Rad52 in MMEJ<sup>148,152,153</sup>, whereas Rad52 seems to inhibit MMEJ along with Rad51 and RPA in more recent studies<sup>147,154,155</sup>. It is likely that the length of the homology dictates the requirement of the HR component Rad52, indeed Rad52 was found to be more associated with repair implicating microhomologies longer than 15 bp<sup>148</sup>.

### *Theta-Mediated End Joining*

In other eukaryotes, MMEJ was also found to play a role in DNA DSB repair and generated deletions or other structural variants. More specifically, in *Drosophila* DSB repair upon transposon excision was shown to contain microhomologies at the break-points and small inserted sequence in the mutants for NHEJ *lig4* and HR *spn-A (rad51)*<sup>156</sup>. It was found later, that most of this repair was due to MMEJ and was dependent on the specific DNA polymerase Theta (Pol $\theta$ )<sup>157,158</sup>, encoded in *Drosophila* by *mus308*, which was initially found to be involved in DNA crosslink repair<sup>159-162</sup>. The Pol $\theta$ -mediated repair mechanism was first seen as a back-up mechanism in the absence of NHEJ, but MMEJ signature was also observed in flies that were not deficient for NHEJ.

Homologues of Pol $\theta$  were found with similar functions in most eukaryotes, including humans<sup>163</sup>, mice<sup>164</sup>, *C. elegans*<sup>165</sup>, zebrafish<sup>166</sup> and *Arabidopsis thaliana*<sup>167</sup>. These studies confirmed Pol $\theta$ 's role in class-switch recombination but also the consequence of the error-prone MMEJ on genome instability<sup>168-172</sup>. However, Pol $\theta$  is absent in yeast. Thus, MMEJ pathways rely on different polymerases in yeast and higher eukaryotes. Because of the implication of Pol $\theta$ , MMEJ is often referred to as Theta-Mediated End Joining (TMEJ), however some other Pol $\theta$  independent MMEJ pathways could coexist in higher eukaryotes<sup>173-175</sup>.

Similarly, to MMEJ in yeast, TMEJ depends on a short DNA resection step<sup>176</sup>. Biochemical analysis also unveiled the Polθ-dependent mechanism of DSB repair. Polθ bound to single-strand DNA and used >2bp double strand microhomologies as a template for DNA replication<sup>177</sup>. Successive rounds of annealing, priming and release are thought to be responsible for short templated insertions at the break site; Polθ was also showed to insert random nucleotides through template-independent terminal transferase activity<sup>178-183</sup>. Polθ is the only DNA polymerase with a helicase domain whose function is not fully clear though integration of inserts requires both the helicase and polymerase domain<sup>184</sup>.

Thus, Polθ end joining repair is characterized by 2–15bp microhomologies (mostly >4bp) and 1–30bp insertions of random or templated sequence.

### *Polθ interaction with homologous recombination*

Polθ was found upregulated in Homologous Recombination deficient cancers such as *BRCA1* and *BRCA2*<sup>185</sup>. In addition, the expression of Polθ or PARP-1, a factor promoting Polθ recruitment to the DNA, was found to be associated with tumor resistance to cisplatin<sup>186,187</sup>. TMEJ and HR begin with resection of the DNA ends and have been shown to counteract each other. Rad51 blocks Polθ recruitment to DNA<sup>188</sup>, and the helicase domain of Polθ is implicated in RPA removal from the resected ssDNA to block HR<sup>189,190</sup>. Consistent with the fact that Polθ inhibits HR, it was shown to limit mitotic recombination<sup>189,191,192</sup>. In cancer, HR deficient cells depend on Polθ for their survival<sup>188,193</sup>, and the *Polθ* mutant signature has been associated with BRCA deficient cancers<sup>194</sup>. Targeting Polθ has become an important potential therapeutic strategy<sup>195,196</sup>.

### *Polθ and translesion synthesis*

In *Drosophila*, *mus308* mutants were identified as particularly sensitive to DNA crosslinking agents<sup>160</sup> like nitrogen mustard. The role of Polθ in crosslinking repair was also demonstrated in *C. elegans*<sup>165</sup> and *Arabidopsis*<sup>197</sup>. In addition, Polθ was involved in error-prone DNA repair in mice after UVs<sup>198</sup>, and in replication

through G-quadruplexes in *C.Elegans*<sup>199</sup>. Thus, Polθ likely acts as a translesion polymerase, although with high mutation rate<sup>163,198</sup>. Furthermore, Polθ was found to cooperate with the BER pathway<sup>200-202</sup>. In addition, MMR seems to function in Polθ inhibition<sup>203</sup> and Polθ mutant cells are dependent on MMR function for survival<sup>204</sup>.

Thus, Polθ interacts with several other pathways and is implicated in the DNA repair of different types of DNA damage in addition of its role in DNA DSB repair. We can then hypothesize that in addition to its well-characterized signature associated with structural variant formation in MMEJ, Polθ may also contribute to specific point mutations signature independently of DNA DSB repair. However, the contribution of Polθ in spontaneous mutation accumulation during aging and more particularly in stem cells has not been investigated. We later provide evidence that Polθ likely contributes to structural variant formation and point mutation in *Drosophila* intestinal stem cells.

*Interestingly, mus308 was found to be targeted by the shRNA interference pathway in the Drosophila ovary<sup>205,206</sup>, potentially indicating a specific inhibition of this pathway in the germline and embryo to limit mutagenesis.*

#### **IV. Replication stress as a cause of DNA damage**

Replication stress is an alteration of the DNA replication that can cause either replication slow-down, replication blockage or incomplete replication during S-phase<sup>207-209</sup>. Replication stress is the main cause of chromosome instability in cancer and has been associated with large genome copy-number variants (deletions, duplication), chromosome rearrangements and ploidy alterations<sup>209-211</sup>.

### *Replication stress and common fragile sites*

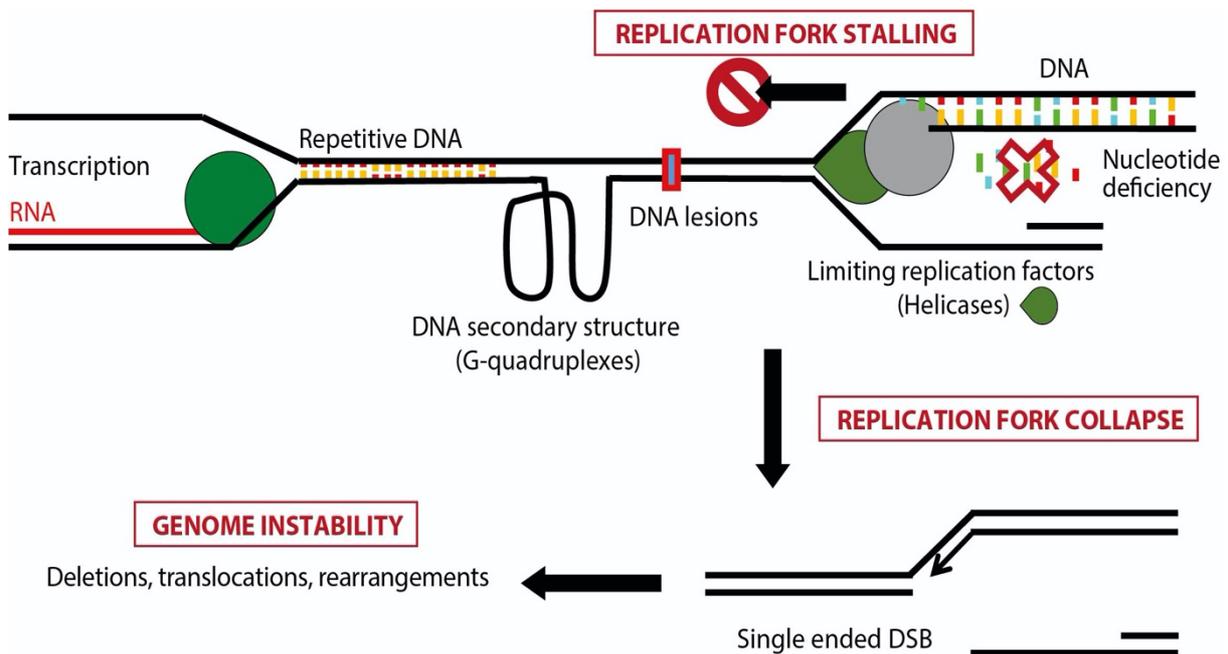
In cell culture experiments, several drugs were used to induce replication stress and study its consequences, the main ones were: Aphidicolin<sup>212</sup>, a DNA polymerase inhibitor, and Hydroxyurea<sup>213</sup> an inhibitor of the dNTP producing enzyme Ribonucleoside reductase. In addition, the deprivation of folate, which is a precursor of nucleotide metabolism, also induces replication stress<sup>214</sup>. Upon culture in conditions of replication stress, cells developed chromosomal breaks and gaps that resemble deletions and translocation. Interestingly, the genome rearrangements recurrently appeared at the same chromosomal location in different experiments. Thus, there are sites of the genome that are more likely to produce genome alteration in conditions of replication stress. In addition, these sites were also similar to known chromosomal alterations linked to cancer or genomic diseases<sup>215-218</sup>. These sites, named “Common fragile Sites” (CFS) were thus considered mutational “hot-spots” and particularly sensitive to replication stress.

### *Endogenous Causes of Replication stress*

Several studies, identified genomic characteristics of the Common Fragile Site<sup>207,209</sup>. First, CFS were found in genome regions with low-density of replication origins and late replication timing, as such they are likely to be unreplicated while the cell progresses to G2 and mitosis. In addition, fragile sites are more frequent in AT-rich region of the genome, interestingly upon induction of replication stress with hydroxyurea, DNA double strand breaks were enriched at poly(dA:dT) sequences of the genome<sup>219</sup>. Furthermore, CFS are enriched in large genes, because they are likely to be replicated late, but also because they are known to be more likely to generate DNA:RNA hybrids called R-loops<sup>207</sup>, which represent a DNA replication blockage as they create topological stress for the incoming replication fork<sup>220</sup>.

Moreover, genomic regions prone to form secondary structure such as inverted repeats and G-quadruplexes are likely to contribute to CFS. These regions

are known to be hard to replicate and promote replication fork stalling, they often require the usage of translesion polymerases<sup>111</sup>. In addition, CFS are enriched in condensed chromatin<sup>207</sup>. Replication can also be blocked by chromosome organization and topological stress at region bound by cohesin<sup>221,222</sup>. Topological stress is resolved by Topoisomerase II in eukaryotes, and inhibition of this enzyme leads to DSB formation at sites previously known to be prone to chromosome rearrangements in different cancers<sup>223</sup>, thus linking genome organization with



**Figure 7. Causes of replication stress and Consequences of replication fork collapse**

chromosome instability<sup>224,225</sup>. In addition, defects in condensin loading in mitosis was showed at common fragile sites of the genome, related to regions engaging in mitotic DNA synthesis<sup>226</sup>.

Finally, the availability in DNA precursors and the balance of replication proteins are essential to complete replication stress. The expression of MCM helicases, which are important for replication initiation and double helix unwinding during DNA synthesis, was compromised in aged hematopoietic stem cells, likely responsible for replication stress accumulation<sup>227,228</sup>. Similarly, defects in nucleotide metabolism is associated with replication stress associated DNA damage<sup>229,230</sup>.

Thus, cell metabolism, transcription program, replication timing and genome organization are main causes of replication stress (**Figure 7**). Because of the cell-type specific replication profile, chromosome organization and transcriptional programs, the sites prone to replication vary between cell types, explaining the differences in CFS identified from different cell populations<sup>207-209</sup>.

### *Oncogene induced replication stress*

Replication stress is more specifically associated with cancer and an oncogene-induced model for replication stress has been proposed<sup>210,231</sup>. Oncogenes act at different levels of the DNA replication regulation. In short, oncogenes such as CycE<sup>232,233</sup>, c-Myc<sup>234,235</sup> and Ras<sup>236</sup>, promote S-phase entry, resulting in shorter G1, and replication origin firing in early S-phase<sup>235,237</sup>. This has been associated with an increase in conflict with transcription in gene-rich regions<sup>235,238</sup>, and is responsible for genome rearrangement at early replicating fragile sites<sup>238</sup>. The acceleration of the cell cycle was also linked to defect in nucleotide metabolism, and oncogene expressing cells undergo replication in condition of nucleotide depletion, thus driving replication fork instability<sup>236,239,240</sup>.

### *Replication fork protection*

During replication, several mechanisms protect the replication fork and participate in replication completion upon fork stalling. Stalled replication forks are protected through recruitment of ssDNA binding protein RPA that participate in the activation of the ATR-mediated DNA damage response. Different mechanisms, which I will not extensively describe, can get replication back on track with limited effect on genome integrity<sup>237,241</sup>:

- Firing of new replication origins, to complete the replication of the affected genome segment.
- Recruiting translesions polymerases to bypass DNA damage blocking the replisome progression.

- Fork reversal or degradation requiring the disassembling of the replisome, limited exonuclease activity and recruitment of the homologous recombination components to find a new homology sequence to prime the reassembly of the replisome.

### *Unreplicated DNA and chromothripsis*

However, as mentioned previously, replication stress is responsible for large chromosome rearrangements, they can be produced through different mechanisms. First, in condition of replication stress, DNA synthesis likely continues after S-phase, into G2 and even mitosis. Unreplicated DNA in mitosis generates ultra-fine anaphase chromosome bridges between daughter cells that cannot properly segregate sister chromatids<sup>242,243</sup>. These unreplicated regions can be resolved through mitotic DNA synthesis (MIDAS) depending on RAD52 and HR resolvase MUS81<sup>244-247</sup>. However, the bridges can be broken and the two daughter cells either lose or the inherit extra-chromosomal part<sup>248</sup>. Chromosome bridges are responsible for large copy-number variant formation like deletion and translocation. It can also lead to the complete or near-complete gain or loss of a chromosome. The presence of unbalanced quantity of a homologous chromosome and free broken pieces of chromosomes leads to micronuclei formation in the cell and often subsequent chromosome loss. This mechanism called breakage-fusion-bridge (BFB) is responsible for a part of the chromosome alteration in cancer cells, and can be generated through a single cell division with a chromosome bridge<sup>248</sup>. Similar to the consequences observed after nuclear envelope breakage that lead to micronuclei formation<sup>249</sup>, the breakage of anaphase bridges could contribute to the massive genomic alterations described as chromothripsis in cancer cells<sup>250</sup>.

### CNV induced by replication fork collapse

Replication fork stalling can lead to replication fork collapse generating a single-ended DNA double strand break, different DNA repair mechanisms can be involved in its repair.

Different systems have been engineered to artificially induce a replication fork stalling at a specific locus and follow the outcomes generated by the fork collapse<sup>251-253</sup>. These systems demonstrated that replication blockage promotes gross chromosomal rearrangement and mitotic recombination. Upon, replication stress induced DNA double strand break, the recruitment of the homologous recombination pathway can lead to classical homologous recombination repair, for example using the homologous chromosome as a template. Alternatively, in Break-Induced replication, the homologous chromosome is used as a template for synthesis, and as already mentioned, this process is highly mutagenic.

N. Willis and R. Scully adapted the Tus/*Ter* system to induce a replication fork blockage at a specific locus in mammalian cells<sup>254</sup>. This system is naturally used in *E. coli* to terminate replication: the Tus protein binds to the *Ter* sequences and block the replisome coming from one direction, thus ending the replication of the bacterial circular chromosome at this site. By expressing Tus in cells containing the *Ter* sequence in a cassette with modified GFP and RFP sequences, they could decipher the outcomes of replication fork collapse.

Most of the events recovered were classical homologous recombination events depending on the Rad51/BRCA1 pathway and using homology stretches inserted in the GFP/RFP cassette<sup>254</sup>. However, some events had a more complex nature with inversions and/or duplications, with micro-homologies at the breakpoints and/or nucleotide addition at the junctions, which could suggest a role for the Pol $\theta$ -dependent End Joining pathway<sup>255,256</sup>. We mentioned earlier that the Pol $\theta$  pathway plays an important role in replication stress resolution in HR deficient cells<sup>188,193</sup>. But Pol $\theta$  also contributes to replication stress resolution and replication progression in HR proficient cells<sup>257,258</sup>. However, some other mechanisms were proposed to explain complex structural variant formation.

It appeared that many events relied on successive templated sequence with micro-homologies, suggesting a replicative mechanism of repair with template switching during DNA synthesis. This mechanism, named Fork Stalling and Template Switches (FoSTeS) was postulated to explain the formation of complex events after replication fork collapse and was characterized by successions of genomic duplication/triplication, inversions and translocation<sup>175,259,260</sup>. Similarly, another mechanism called MMBIR for Microhomology-mediated Break Induced Replication was proposed to generate such complex events<sup>261</sup>. MMBIR results from defective synthesis-dependent HR and lead to template switching using microhomologies<sup>173,262</sup>. Similarly to BIR, MMBIR relies on the POL32/POLD3 subunit Pol $\delta$ <sup>263</sup> or the other translesion polymerase Rev1<sup>262</sup>, in a mechanism reminiscent of the MMEJ observed in yeast<sup>147</sup>. Whether Pol $\theta$  plays a role in MMBIR has not been excluded.

The complex rearrangements following replication stress were largely described in the context of cancer<sup>264</sup> but could also be related to copy number alteration in non-cancer genomes, for events happening in the germline or during embryonic development<sup>261,265-267</sup>. Whether replication stress also involves chromosome rearrangements in proliferating stem cells remains unclear.

### **Replication stress and cancer therapeutics**

Because of cancer cells' susceptibility to DNA damage and in particular to replication stress, several therapeutic strategies are being developed to induce increased levels of replication stress in cancer cells or to blocks the mechanisms that enable them to survive high levels of replication stress<sup>268</sup>. For example, nucleotide synthesis inhibition to induced more replication stress is used to treat glioblastoma and leukemia, and is envisioned for other cancer<sup>269-273</sup>. Strategies targeting PARP-1 and Pol $\theta$  which have been involved in replication stress resistance in BRCA1/BRCA2 cells are also being tested<sup>195,196</sup>.

Thus, we highlighted the important contribution of replication stress to genome instability in cancer. Nucleotide depletion is a potent driver of oncogene-induced replication stress and is targeted for cancer therapeutics. However, how replication stress could affect adult stem cells and contribute to mutation accumulation during aging is unclear. Below, we investigated the contribution of nucleotide depletion on adult stem cell physiology and DNA damage. We also interrogate tissue-specificity in sensitivity to replication stress. We uncovered a non-cell autonomous buffering mechanism of nucleotide levels dictating DNA damage susceptibility.

## C. Somatic mutations and aging stem cells

Tissue-specific stem cells are found in most mammalian adult tissues. They are essential for tissue function and to maintain tissue homeostasis, i.e. enabling the tissue renewal during aging and promoting regeneration in case of tissue damage. They achieve the maintenance of tissue homeostasis via their ability to differentiate into the several cell types of the tissue and through self-renewal, essential to sustain a functional stem cell pool. Stem cells achieve the balance between renewal and differentiation through asymmetric division whereby they produce a differentiating cell and a cell that will maintain stem cell features. A large section of stem cell literature focuses on understanding the genetic, epigenetic and systemic mechanisms controlling stem cell proliferation and differentiation, and the influence of the stem cell microenvironment, called the “niche”, in regulating stem cell function. The hematopoietic system was the first to be extensively studied *in vivo* and *in vitro*, but more recently the characterization of stem cell populations in the intestine, the skin and other epithelia has provided unvaluable insight on what defines a stem cell and how it is regulated. In *Drosophila*, the adult midgut has proven to be an excellent model to address the same questions. (see Part D.).

### I. Stem cell function and aging

Although adult stem cells are mostly found in the tissue throughout life, age-related changes in stem cell function have been observed in different systems.

#### *Stem cell loss*

Upon aging, stem cell can be lost either because they lose the ability to self-renew and differentiate, or because of stem cell death. In the skin, hair graying is link to the loss of melanocyte stem cells (MSC)<sup>274</sup>. The depletion of the stem cell pool is likely linked to the progressive differentiation to non-renewing pigment producing melanocytes, and linked to different causes: (1) DNA damage,

as genotoxic stress induced by ionizing radiation has been shown to drive MSC differentiation<sup>275</sup>; (2) organismal stress, since the stress-induced noradrenaline production by neurons surrounding the dermal papilla of the hair follicle specifically prompted the depletion of melanocyte stem cells<sup>276</sup>.

Stem cell can also be lost as a result of exhaustion, for example in the case of subsequent rounds of forced proliferation, which could be induced upon repetitive injury to the tissue. In the *Drosophila* midgut, the stem cell pool tend to decrease following successive regenerative response after bacterial infection<sup>277</sup>. This phenomenon was also observed upon injury in the zebrafish brain, whereby the neuron progenitors were lost overtime<sup>278</sup>. Similarly, progenitor loss is a marker of brain aging in mice<sup>279</sup>. In the hematopoietic system, several rounds of isolation and transplantation or stress decrease the stem cell ability to produce clones<sup>280-282</sup>.

A variety of stem cells are maintained in a quiescent state, changes in the niche signaling can drive them out of quiescent to differentiate, thus leading to stem cell loss, this has been observed in muscle stem cells, for which increase in FGF signaling from the niche drives satellite cells out of quiescence and depletion of the stem pool<sup>283</sup>. Reactivation of old quiescent hematopoietic stem cells also provokes their depletion<sup>284</sup>. However, stem cell loss through differentiation or exhaustion is likely tissue-specific and in no way a generalized consequence of stress or aging in all systems.

### **Proliferation blockage**

Upon aging, some stem cells lose the ability to proliferate in homeostatic condition and or for a regenerative response upon injury or simply maintain tissue homeostasis. A decline in stem cell ability to proliferate was observed in the muscle, whereby the overactivation of TGF- $\beta$ <sup>285</sup> or loss of Notch signaling activation<sup>285,286</sup> limits satellite cell proliferation and tissue renewal. In other tissues, the modification of the niche is promoting stem cell quiescence and conveys the

loss of ability to proliferate for homeostasis and regeneration<sup>284</sup>. In some instances, stem cell aging leads to stem cell senescence.

Systemic stress can also force stem cells into quiescence and limit their regenerative potential, for example the stress hormone corticosterone blocks hair-follicle stem cell proliferation and reduce hair growth in mice<sup>287</sup>.

### **Loss of differentiation potential**

One characteristic of many tissue-specific stem cells is the ability to differentiate into the diverse cell types of the tissue. However, upon aging this ability can be lost. Not only because stem cell can lose their ability to proliferate (see above), but because their differentiation potential can be skewed toward one lineage. Thus, it creates an imbalance in the differentiated cells produce. This could be seen in the hematopoietic system, in which the aged hematopoietic stem cell produce more cells of the myeloid lineage at the expense of the lymphoid lineage<sup>288</sup>.

### **Overproliferation and tumorigenesis**

Lastly, stem cells can engage in overproliferation and tumorigenesis, for example in the skin<sup>289</sup>. Overproliferation can be driven by intrinsic factors, such as persistent inflammation or intrinsic modification forcing cell-cycle progression. Interestingly, the tumorigenic potential and the type of tumor generated upon oncogene expression depends on the stem cell type and origin in the tissue.<sup>290,291</sup> It is now well established that many neoplasia arise from stem cell misregulation and overproliferation, and that in the tumor some cell maintain stem cell properties to fuel tumor growth, the cancer stem cells<sup>290</sup>.

### **The Hallmarks of aging**

The study of aging stem cells uncovered several changes that could be driving stem cell loss of function and any outcome presented above<sup>292</sup>. It has

been observed that aged stem cells present epigenetic alteration<sup>293,294</sup>, metabolic dysfunction<sup>295</sup>, defective cellular stress response (proteostatic stress, oxidative stress), modified communication with their niche and a different ability to respond to regulating signals (mentioned above). In addition, they have distinct response to DNA damage, and also an accumulation of somatic mutation was observed in many tissues<sup>284,296</sup>.

## II. DNA damage/repair and aging

The link between aging and DNA damage/repair was first established with the study of different progeroid syndromes, marked by an accelerated aging and a higher susceptibility to cancer. Patients with congenital mutations in NER, MMR or DSB repair pathways, for example in Xeroderma Pigmentosum, Cockayne syndrome, Ataxia Telangiectasia, or Bloom's and Werner's syndrome, have higher susceptibility to cancer, neurodegenerative diseases, immunodeficiency and display rapid age-related decline as seen by hair graying, alopecia or fibrosis<sup>296</sup>.

More specifically, in different tissues, old stem cells have higher levels of DNA damage, and efficient DNA repair is a limiting factor in stem cell maintenance. This was observed for example in the skin<sup>297</sup> and the hematopoietic system<sup>227,298-301</sup>. Thus, DNA damage susceptibility and DNA repair ability contribute to tissue stem cell aging<sup>107,108,284,296</sup>.

## III. Somatic mutations in stem cells

It has been established that DNA damage alters stem cell function during aging. Whole-genome sequencing approaches demonstrated that tumorigenesis is associated with high levels of mutations and enabled the characterization of tumor specific mutational landscape. However, until recently, the spectrum of mutations in healthy somatic tissues was underappreciated, mostly because detecting

somatic mutations is technically challenging. Somatic mutations likely emerge independently in different cells creating tissue mosaicism, generating limited evidence supporting any variants in whole-tissue sequencing. Only somatic mutations present in expanded clones inside the tissue would be detected. However, recent advances in stem cell culture methods (e.g. organoid growth from single cells<sup>302</sup>) and sequencing techniques (e.g. single cell whole genome sequencing<sup>303</sup>) helped to detect somatic variants in stem cells and differentiated cells.

The mutational landscapes of different healthy epithelium like skin<sup>304</sup>, oesophagus<sup>305,306</sup>, intestine<sup>302,307,308</sup>, breast<sup>309</sup>, lung<sup>309</sup>, endometrium<sup>310</sup>, but also of the liver<sup>302,311</sup>, brain<sup>312</sup> and blood<sup>313</sup>, have been partially characterized. These studies demonstrated a linear accumulation of mutations with age. In addition, they showed tissue-specificity in the mutational profiles, with some mutational signatures directly linked to cell-specific metabolism or characteristic exogenous factors of the tissue environment. Interestingly, clonal accumulations containing mutations in tumor suppressor gene, such as NOTCH and P53 were detected with aging<sup>305,306,308</sup>, giving us insights into the mechanisms driving cancer initiation in apparent healthy tissues<sup>284,314,315</sup>.

Although somatic mutations were detected in various tissues, we are still beginning to understand the extent of mutagenesis in healthy tissues and the consequences on stem cell and tissue function, clonal dominance and tumor initiation. How different mutation accumulation is between tumor and healthy tissues is unclear. In addition, the contribution in adult stem cells mutagenesis of different DNA damaging factors, such as replication stress, and the role of different DNA repair pathway in specific mutation accumulation have not been investigated.

## **D. The *Drosophila* Midgut**

*This section was adapted from a review submitted recently to Current Opinion in Cell Biology (Boumard B. and Bardin A.J., 2021). It presents aspects of the regulation of intestinal stem cells proliferation and differentiation, and age-related changes. Although the focus was mostly on recent publications, and as such voluntarily omits some past work, I think that it represents a good overview of the questions addressed using the *Drosophila* midgut as a model system. The last part on stem cell aging and genome instability was extended as it is directly relevant to the experiments presented below.*

Adult tissues rely largely on the activity of stem cells to maintain tissue homeostasis and orchestrate a regenerative response to injury response. Adult stem cell proliferation, self-renew, and lineage differentiation are coordinated by numerous signaling pathways through epigenetic, transcriptional, post-transcriptional and post-translational mechanisms. Functional changes of stem cell properties during aging have been associated with a failure of these regulatory mechanisms.

Given its relative simplicity along with precise genetic tools, the midgut of the fruit fly, *Drosophila melanogaster*, and its adult intestinal stem cells (ISCs) have proven to be an excellent model system to decipher stem cell lineage decisions, proliferation control during homeostasis and stress, and age-related functional decline. Here we present a review of recent findings on how adult intestinal stem cells differentiate, interact with their environment, and change during aging.

### **I. ISC identity and cell lineage decisions**

Previous studies over the past 15 years have characterized the ISC lineage and division properties allowing for differentiation and stem cell self-renewal.

Comparable to the mammalian intestine and lung epitheliums, the *Drosophila* ISCs can produce two differentiated cell types: hormone-producing enteroendocrine cells (EEs) and absorptive enterocytes (ECs) through dividing enteroendocrine precursors (EEPs) and directly differentiating enteroblasts (EBs), respectively<sup>316-319</sup>. Here we highlight recent findings advancing our understanding of cell fate and lineage decisions in ISCs (**Figure 8 & Figure 9**).

### *Spindle orientations and asymmetric vs symmetric stem cell fate decisions*

Adult stem cells undergo routine renewal as well as regenerative repair, a process that often requires expansion of the stem cell pool. While previous studies proposed a role for spindle orientation in asymmetric/symmetric fate acquisition<sup>320</sup>, recent work has clarified this further: ISC divisions having a planar spindle orientation parallel to the basement membrane are promoted through Jun Kinase (JNK) signalling, which directly controls the spindle regulatory components Kif1a, Wdr62, and Mud<sup>321</sup>. Planar spindle orientations during division correlates with symmetric ISC fate acquisition and are more prevalent during regeneration, adaptive resizing, and aging<sup>321</sup>. As cell fate acquisition in the two ISC daughter cells depends largely on Notch/Delta signalling, future studies will be needed to determine how JNK-driven planar ISC divisions may alter ISC fate via Notch/Delta signalling and how they may relate to previous studies indicating a contribution of Bmp signalling on symmetric ISC divisions<sup>322</sup>.

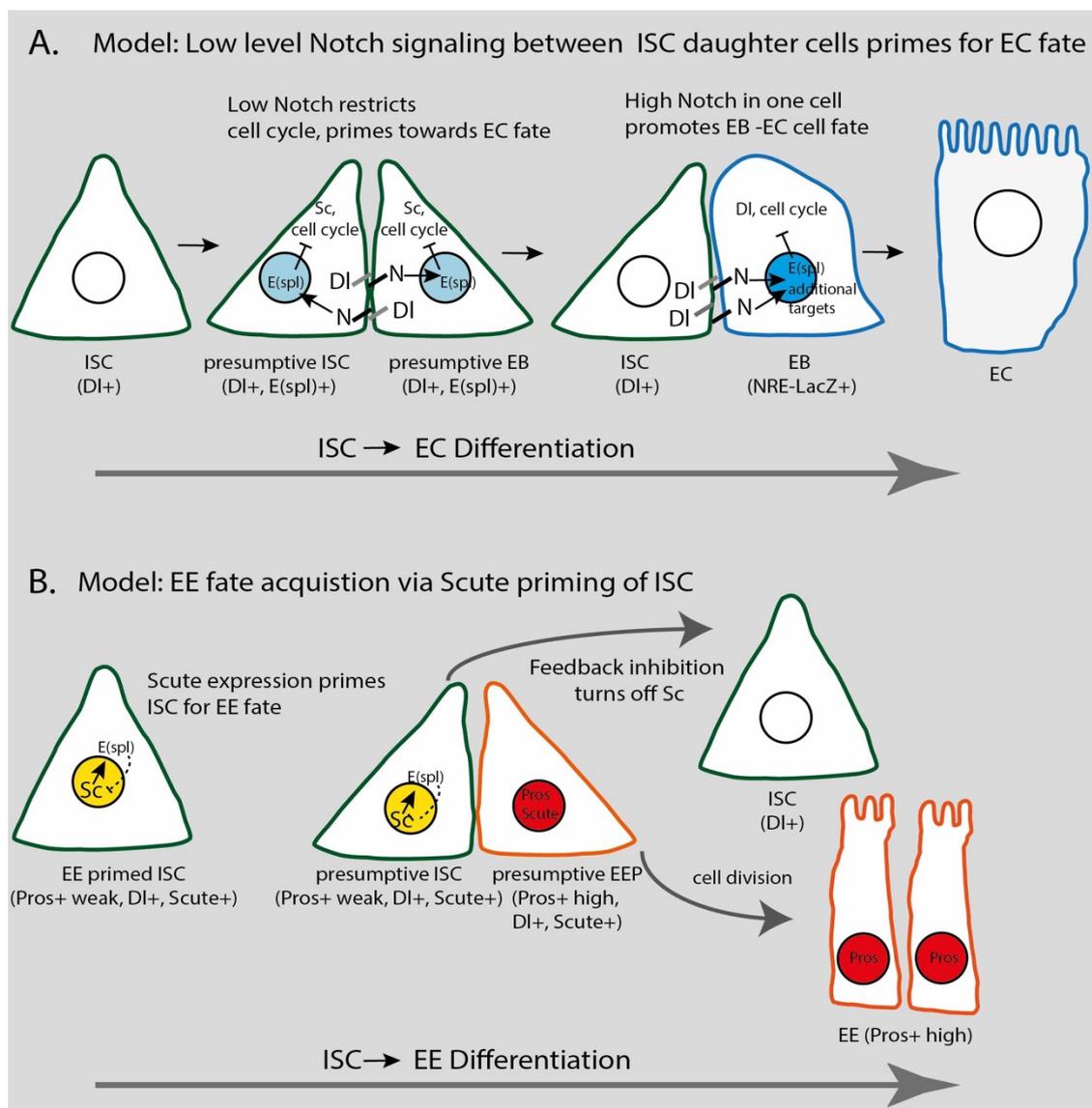
### *Maintenance of ISC identity and early lineage decisions*

ISC lineage decisions rely heavily on Notch signalling promoted by its ligand Delta (DI). *Notch* inactivation in stem cells lead to neoplastic accumulation of ISCs and EEs, suggesting that Notch activity is essential for EC but not EE differentiation, and that Notch pathway limits ISC proliferation<sup>323-325</sup>.

A complete understanding of how Notch signalling acts has been challenging, in part due to previous mis-conceptions about the cell lineage. Indeed, evidence now suggests that a dividing enteroendocrine precursor cell (EEP), is made by the

stem cell and further divides once to make 2 EE cells<sup>319</sup>. Notch signalling was previously thought to be off in the ISC, though recent data suggest that low level of Notch signalling and activation of a subset of target genes may have essential functions in the ISC to regulate its maintenance, proliferation, and lineage decisions.

Indeed, more recent studies interrogated the identity of DI<sup>+</sup> Pros<sup>+</sup> cells using lineage tracing and single cell analyses and support the notion that in wild-type contexts, these are EE primed ISCs or EEPs<sup>318,319,326,327</sup>. Notch signalling occurring between two daughter cells leads to low level activation of bHLH *E(spl)*-



**Figure 8. Models of Notch signalling and cell fate decisions in the ISC lineage.** Several lines of evidence point toward an essential function of Notch in EC differentiation and ISC maintenance (A), while Scute primes ISC for EE differentiation (B).

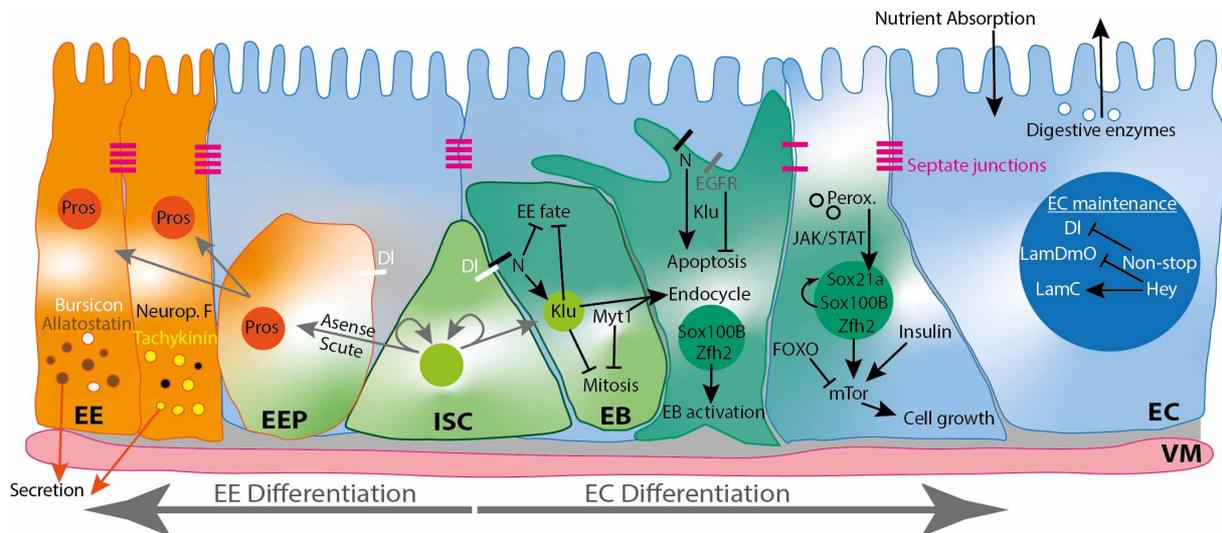
*C* target genes<sup>328</sup>. bHLH E(spl)-C proteins with their corepressor, Groucho, act to repress cell cycle genes in the stem cells<sup>328</sup>. We propose that this low-level Notch activation may also be essential to maintain these ISCs in an EC primed state by repression of the proneural transcription factors Scute, and likely Asense (**Figure 8A**), consistent with reports of high levels of Scute upregulation in *Notch* mutant contexts<sup>319,329</sup>.

How then is an EE cell produced? Elegant studies demonstrated that Scute becomes expressed in a subset of ISCs, likely implicating a Scute auto-feedforward loop<sup>319</sup>. When a sufficient level of Scute is present, the ISC switches to become EE primed and will divide and produce an ISC and an EEP, which upon further cell division will make EE cells (**Figure 8B**). Additional studies with live imaging and fate markers will determine the potential role of Pros segregation during ISC division as well as other factors implicated in EE fate choice such as Slit-Robo signalling, Numb, and mechanical input<sup>318,319,326,327,330</sup>. Future work will undoubtedly reveal additional surprises in cell fate control of the ISC lineage.

Independent of DI/Notch signalling, additional transcriptional and post-transcriptional control helps maintain ISC identity. A role for the cohesin component, Rad21, in ISC maintenance has been established, likely through its contributions to chromatin structure and regulation of differentiation genes<sup>331</sup>. Tramtrack69, a transcriptional repressor, is also required to maintain ISC identity and its inactivation converts ISCs to neural stem cell-like state<sup>332</sup>. In addition to control at the transcriptional level, recent studies also uncovered a role of post-transcriptional regulation, whereby P-bodies in ISCs sequester and block the translation of mRNA encoding differentiation genes such as Pdm1, expressed at low levels in the ISC<sup>333</sup>. These findings highlight the necessity of transcriptional and post-transcriptional coordination required to preserve stem cell identity.

## How to build an enterocyte

Recent work has better defined how lineage differentiation of ISC is directed towards EC or EE terminal cell fates (**Figure 9**). EBs can remain dormant in the tissue before being "activated" to differentiate. Activation leads to EB cell growth, endoreplication, and morphological changes that include lamellipodia formation and the increased expression of septate junction components, such as Tsp2A and Mesh required for EB integration in the epithelium<sup>326,334,335</sup>. Previous studies indicated that EB differentiation requires coordinated activity of numerous transcription factors, including Esg, Sox21a, GATAe, and Pdm1 (reviewed in<sup>330</sup>). New findings demonstrate that an early event in the EB differentiation is the expression of the transcription repressor, Klumpfuss (Klu), which inhibits EE determinant genes<sup>326,336</sup>. In addition, Sox100B has a direct role activating transcription of Sox21a in EBs, essential for EC fate acquisition<sup>337-340</sup>. Further morphological changes and growth of EBs are also promoted by the zinc finger transcription factor Zfh2<sup>340</sup>. Future studies will be needed to better understand the regulatory interplay between Klu, Zfh2, Sox100B, and previously identified



**Figure 9. Lineage regulators of intestinal stem cell differentiation towards enteroendocrine and enterocyte cells.** Intestinal stem cells (ISCs) are essential to maintain tissue homeostasis, they mostly divide asymmetrically to self-renew and produce a differentiating progenitor. Intestinal stem cells are multipotent as they can generate two differentiated cell types: absorptive enterocytes (ECs) representing ~80% of the progeny and hormone secreting enteroendocrine cells (EEs), ~20%. The differentiation is achieved through enteroblasts (EBs) and dividing enteroendocrine precursors (EEPs), respectively. Here we focused on recent advances, please see the text for further details.

transcription factors important for EB to EC differentiation such as Sox21a and GATAe.

Interestingly, there is an emerging role for organelle activity and metabolic state in specific steps of differentiation (**Figure 9**). Peroxisome function, which is induced upon injury, promotes repair and EC differentiation by impinging on endocytosis and late endosome maturation. Defective peroxisomes halt differentiation of EBs by altering late endosomes and blocking JAK/STAT activation required for downstream Sox21a expression<sup>341</sup>. Defects in the metabolism of mitochondria also inhibit EB growth and EC differentiation via activation of FOXO which prevents mTor signalling<sup>342</sup>. How other metabolic processes impinge on cell states will be an important future direction of study.

Another critical aspect of EB differentiation is a shift from the mitotic cell cycle in ISCs, to the endocycle in differentiating EBs, resulting in polyploid ECs (**Figure 9**). Recent studies suggest an important role of the Cdk1 inhibitory kinase, Myt1, which inhibits Cyclin A in EBs to promote G2 arrest and a switch to endocycles<sup>343</sup>. In addition to its roles mentioned above in repressing EE fate, Klu is also proposed to inhibit the mitotic cell cycle in EBs through binding to and repressing *CycE* and *CycB*<sup>336</sup>. Alteration of nucleotide metabolism has also been suggested to be critical for EC differentiation, likely through control of endocycle<sup>344</sup>.

Finally, it was recently established that EB fate is also under the control of cell death pathways. Under routine homeostatic conditions, EBs are subjected to both cell death-promoting caspase activity and inhibition of this process through the EB-specific expression of the caspase inhibitor, Diap1<sup>345</sup>. A balance of EB survival vs apoptosis is accomplished by Notch signalling priming EBs for cell death and EGFR signaling promoting cell survival. This mechanism ensures that in healthy, homeostatic tissues, excess numbers of EBs are culled, whereas in regenerative conditions, EB survival is enhanced<sup>345</sup>. Additional non-apoptotic functions of the initiator caspase Dronc also promote EB differentiation<sup>346,347</sup>.

These studies bring to light the complexity of regulation of the ISC to EC differentiation process, integrating intrinsic transcription and post-transcriptional regulation as well as tissue-level information.

### How to keep an enterocyte

Once differentiated, EC nuclear organization and program of transcription are actively preserved. The Non-STOP Identity Complex (NIC), a deubiquitylase that targets histone H2B-Ub, and the transcription factor Hey appear to be important for maintenance of the EC state, with their inactivation resulting in de-repression of *Delta*, a gene normally expressed only in ISCs<sup>348,349</sup>(**Figure 9**). Thus, the differentiated cell state is actively maintained through chromatin regulation. Additional studies are required to understand the influence of changes to chromatin, their impact on lineage choices, and fate restriction.

### EE fate and diversity

Several factors regulating ISC differentiation towards EE fate have been established and are also described more extensively elsewhere<sup>330,350</sup>. The transcriptional program leading to EE differentiation relies on the bHLH factor Scute that primes ISCs toward EE fate (**Figure 8** and **Figure 9**). Upon cell division, the EE primed ISC will give rise to an EE precursor cell (EEP) along with a renewed ISC (Figure 2). Single-cell RNA-seq data has provided additional evidence that EE cells originate through distinct EEPs<sup>326,327</sup>. It was recently suggested that EE priming in progenitor cells is epigenetically regulated by Polycomb complex proteins<sup>351</sup>. The inactivation of *Polycomb* leads to downregulation of EE genes, which is likely indirect given Polycomb's well-characterized role in transcriptional repression.

Single cell RNA sequencing of the whole gut or an EE-enriched cell population illustrated the functional diversity of EE cells<sup>326,327</sup>. 10 major subtypes of EEs were identified with strong regional specialization within the midgut. The differentiation of most of the characterized subtypes relies on the expression of 14 different

transcription factors and arise from distinct sub-lineages with differential requirements for Notch signaling<sup>327,352</sup>. Each subtype is associated with the expression of 2 to 5 peptide hormones as well as specific hormone receptors. While some functions of EE peptide hormones are known<sup>330</sup>, undoubtedly future studies will elucidate additional short- and long-range signalling activities of these molecules.

## II. Fine-tuning ISC proliferation

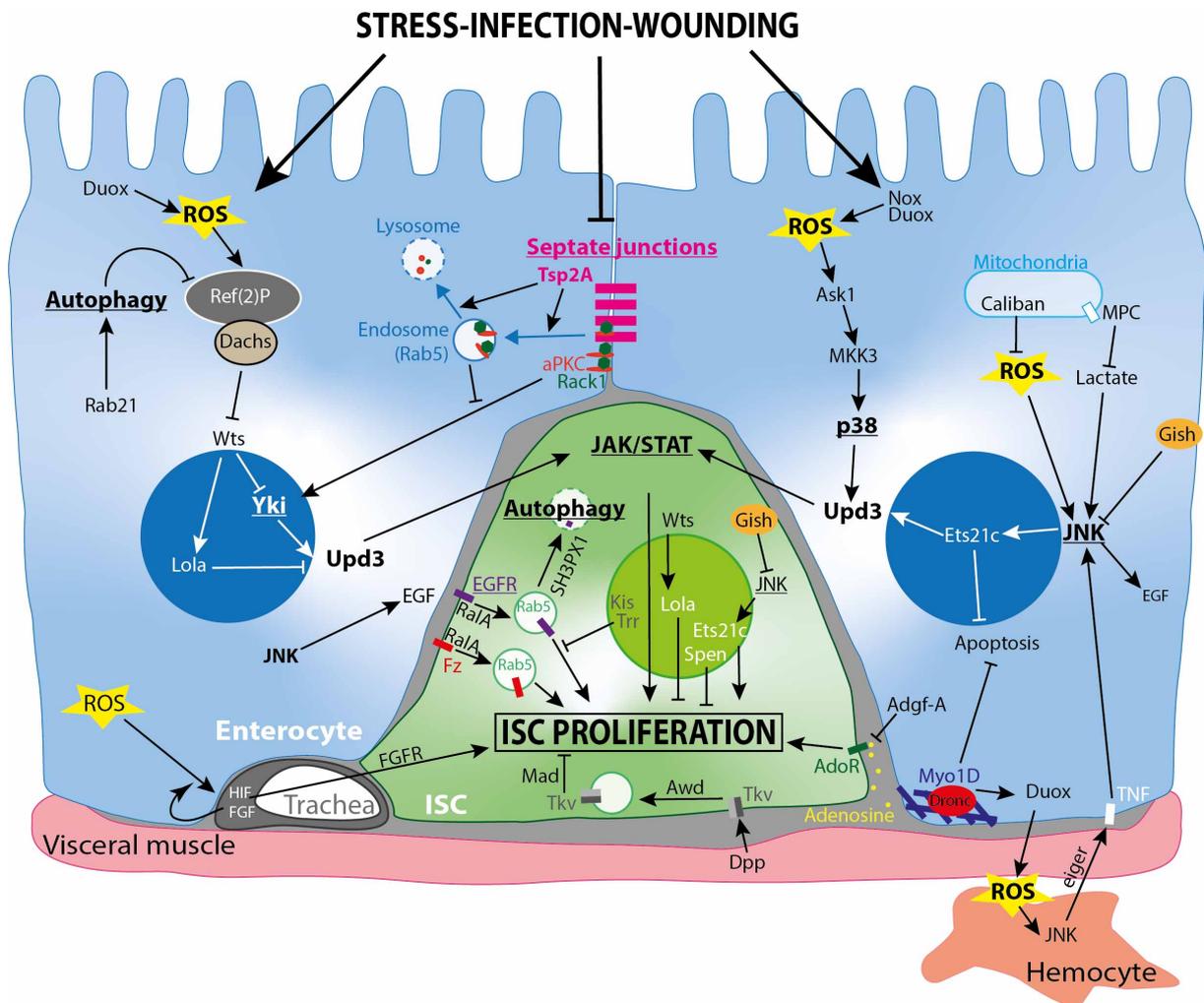
In the recent years, many signaling pathways have been shown to regulate ISC proliferation in homeostasis and in response to tissue injury. Although some cell-autonomous factors directly regulate ISC divisions, non-cell autonomous signals are important to control stem cell proliferation in response to tissue and organismal needs. Jak-Stat and EGFR pathways are the primary mitogenic regulators driving ISCs to increase proliferative capacity in response turnover and injury, though ISCs receive important input from other pathways including Wnt/Wg, BMP/TGF $\beta$  Hippo, JNK, and p38 among others (for a more detailed review, see<sup>330,350</sup>). How these different pathways are integrated in ISCs to balance cell proliferation rates is still somewhat elusive. We will discuss here recent advances in our understanding of cell-autonomous and cell-non-autonomous regulation as well as long-range input from other tissues (**Figure 10 & Figure 11**).

### Cell-autonomous regulation of proliferation

Recent studies indicate that the trafficking and degradation of signaling receptors are exquisitely regulated in ISCs. After tissue damage-induced proliferation, internalization of the BMP receptor, Tkv, requires Awd, a facilitator of dynamin endocytic function; this promotes Mad activation driving ISC return to quiescence<sup>353</sup>. Similarly, endocytosis of the Wnt receptor (Fz3) and EGFR is regulated by the RalA GTPase, which is essential for the regenerative response of ISCs after damage<sup>354,355</sup>. Likewise, under homeostatic conditions, EGFR protein

levels are tightly regulated in ISCs and limited through SH3PX1-dependent autophagic degradation<sup>356</sup>. EGFR protein levels are also mediated by the chromatin remodelers, Kismet and Trithorax-related that promote expression of Cbl, an E3-ligase that degrades EGFR, thereby restraining ISC proliferation during homeostasis<sup>357</sup>. Identifying other regulators of receptor and ligand intra- and extracellular dynamics will be necessary to fully understand ISC proliferative responses in homeostasis and upon injury.

Genetic screens have identified genes with striking phenotypes in controlling ISC proliferation (Figure 10). Interestingly, the transcription factor Lola, limits ISC proliferation



**Figure 10. Cell autonomous and Non-cell autonomous control of proliferation in the midgut.** In homeostatic conditions, stem cell proliferation is adjusted to the needs of the tissue. Upon tissue stress or damage, stem cell proliferation is triggered as a regenerative response. To balance proliferation and quiescence, ISCs integrate JAK/STAT, JNK, WNT/Wg, Hippo, EGFR, FGFR, AdoR and BMP/Dpp signaling pathways, cues of which are provided cell-autonomously and non-cell-autonomously. This scheme represents some sensing mechanisms driving mitogenic ligands and interactions recently identified. For simplicity, the arrows can represent direct protein-protein interactions, transcriptional control or genetic interactions. Please see the main text for details.

proliferation by downregulation of cell cycle regulators and seems to act downstream of Hippo/Wts, yet independently of the canonical component Yki<sup>358</sup>. In addition, the loss of activity in progenitor cells of Spen, involved in RNA processing and transcriptional repression, results in a large excess of ISCs, which requires Insulin signaling<sup>359</sup>. In addition, a novel role on ISC proliferation was found for Adenosine receptor (AdoR) signaling that stimulates Ras<sup>360</sup>. Acting during homeostasis, AdoR signaling is amplified after tissue damage due to enhanced extracellular adenosine levels<sup>360</sup>. Further knowledge of stem cell-intrinsic regulatory programs governing cell cycle and proliferative status will be informative.

### **NON-cell autonomous regulation**

Enterocytes play primary roles in sensing epithelial damage and produce mitogens, including Unpaireds (Upds) and EGFs to activate ISC proliferation non-cell-autonomously. How tissue damage is sensed in ECs to induce a regenerative response has been further investigated in the last few years.

One important stress signal in ECs is reactive oxygen species (ROS; **Figure 10**). Recent studies have better defined molecular mechanisms upstream and downstream of ROS in ECs<sup>361</sup>. Tissue damage induces ROS via Nox and Duox enzymes activating the Ask1-MJJ3-p38 pathway<sup>362</sup>. However, ROS can also activate JNK signaling, whose downstream effects in ECs rely at least in part on the transcription factor Ets21c<sup>363</sup>. Ets21c induces Upd3 expression in ECs that mediates Jak/STAT-dependent proliferation in adjacent ISCs. JNK activation in ECs can also be motivated by ROS production due to defects in mitochondrial metabolism or alteration in pyruvate metabolism<sup>364,365</sup>. In addition, a recent study demonstrates that in response to JNK signaling caused by tumor growth, EGFs are cleaved and activated by rhomboid to upregulate EGFR signaling in ISCs<sup>366</sup>. JNK, therefore, is a broad sensor of numerous types of tissue damage. Strikingly, gut epithelial ROS can signal to nearby other tissues: In response to damage-induced ROS, tracheal cells promote stem cell proliferation via the FGF/FGFR

pathway<sup>367</sup>. ROS produced in dying ECs cells can recruit hemocytes, that also produce ROS and further promotes JNK activation in nearby ECs<sup>368</sup>.

In addition to tissue damage *per se*, diverse cellular defects in ECs cause a stress response that non-autonomously stimulates ISC proliferation (**Figure 10**). Functional septate junctions (SJ) are essential sensors of EC health; their perturbation leads to activation of Yki and Upd3 expression<sup>334,369–371</sup>. Mechanistically, the SJ component, Tsp2A, promotes endocytosis and lysosomal degradation of aPKC<sup>334</sup>. Since aPKC antagonizes Hippo, its accumulation upon Tsp2A knockdown results in activation of Yki. A sensor function for SJ is likely acting during stress, aging, and defective trafficking or autophagy all of which indirectly lead to loss of SJ proteins contributing to a proliferative response in these contexts<sup>334,366,372,373</sup>. Conversely, autophagy has an additional role in preventing ISC overproliferation: autophagic degradation of Dach5, an activator of Yki, is required to prevent Yki–Upd3–driven overproliferation of ISCs in response to ROS generated by commensal bacteria contexts<sup>373</sup>.

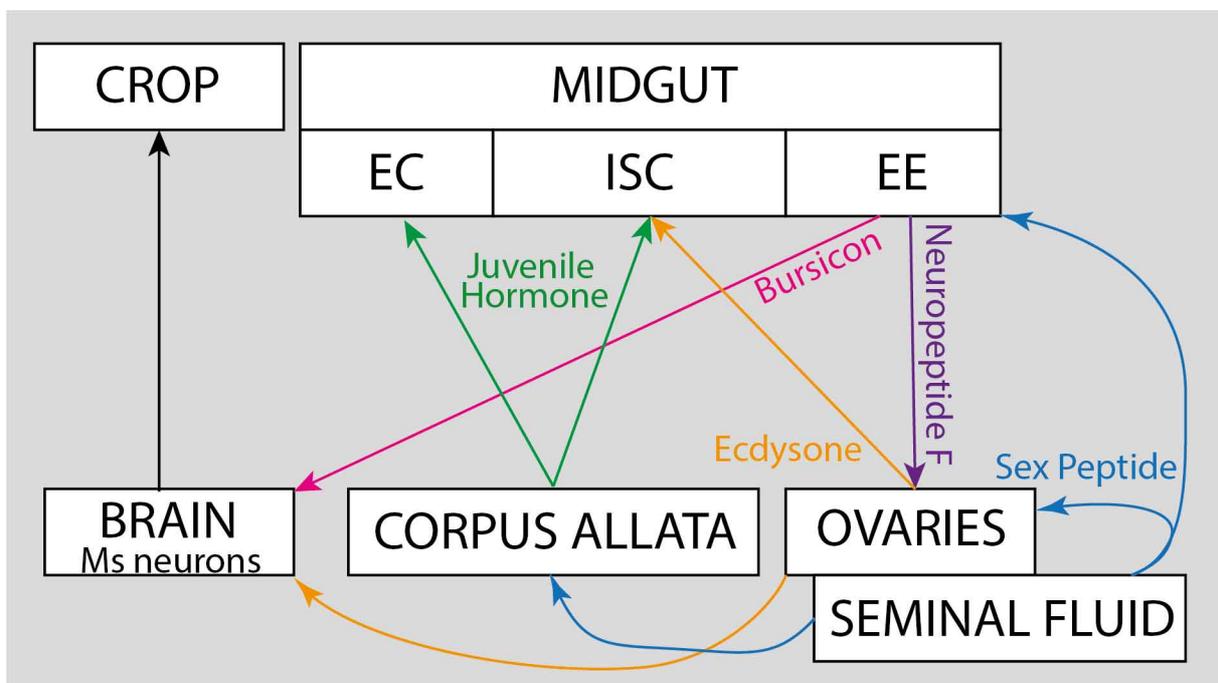
Overall, these new studies demonstrate how ISCs integrate cell-intrinsic and extrinsic cues from neighboring cells in the tissue to fine-tune proliferation allowing both homeostatic cell replacement and regenerative responses. In addition to coordination within the gut epithelium, it is now appreciated that longer-range signaling from the other tissues including the hemocytes, brain, corpus allata, ovaries, and testes can alter stem cell proliferative properties and gut physiology.

### **Steroid and peptide hormone regulation of gut physiology**

The adult gut is influenced by a variety of systemic factors, including steroid hormones (**Figure 11**). Mating provokes physiological and behavioral changes in females that are mediated by Sex peptide (SP), Juvenile Hormone (JH) and Ecdysone allowing an increase in nutrient intake required to fulfill the energy-demanding egg production.

After mating, SP in the male seminal fluid induces a fast response in females responsible for an increase in Ecdysone production by the ovary and an increase in JH production by the corpus allata through a neuronal relay<sup>374,375</sup>. Both SP and Ecdysone directly promote ISC proliferation and gut expansion<sup>376,377</sup>. Mechanistically, Ecdysone is imported in ISCs via the Ecdysone importer (Eci), and binds with Usp to the Ecdysone Receptor (EcR) driving the proliferative response of the ISC through inducing the expression of EGF signaling genes (*spi*, *krn*, *rho*) via *Eip75B*<sup>378</sup>, and by promoting the Pri mediated post-translational modification of the mitogenic factor Shavenbaby<sup>379</sup>.

Ecdysone also promotes food intake through myosuppressin (Ms) expressing neurons mediated crop enlargement<sup>380</sup>. In addition, the enteroendocrine cell (EE) pool increases upon mating. EE secretion of the neuroendocrine peptide Bursicon drives: (1) Ms neurons activation responsible for food intake<sup>380</sup>, revealing a brain-gut-brain axis regulated by Ecdysone; and (2) lipid metabolism and energy catabolism in the Enterocytes and in the fat body<sup>375,381</sup>. Through different pathways, enteroendocrine cells modulate the



**Figure 11. Steroid and peptide hormone regulation of gut physiology.** Mating provokes physiological and behavioral changes in females that are mediated by Sex peptide (SP), Juvenile Hormone (JH) and Ecdysone allowing an increase in nutrient intake required to fulfill the energy-demanding egg production. See the text for details.

organism metabolism<sup>382</sup>. EE cells feedback to the ovaries and promote germline stem cells division via production of neuropeptide F in response to SP signaling<sup>374</sup>. In males as well midgut metabolic changes are linked to sperm production and food intake regulation<sup>383</sup>.

### III. Stem cell long-term maintenance and aging

Despite the adult fly only living 6 to 8 weeks, the midgut shows significant age-related changes in tissue organization and cellular function during this time. At the tissue level, aging promotes JNK-driven stem cell overproliferation and results in dysplasia development<sup>384</sup>. A change in mitotic spindle orientation mediated by JNK signaling has been suggested to be responsible for increasing ISC symmetric divisions, and extending the pool of proliferative cells in the gut<sup>321</sup>. Aging is also associated with microbiota dysbiosis powering persistent ROS production and immune signalling<sup>385</sup>, both known mitogenic signals. The interaction with diverse microbiota influences intestinal stem cell function and tissue regeneration<sup>386-388</sup>. Both commensal and pathogenic bacteria affect metabolism and immunity of the gut and organism<sup>389,390</sup>. Furthermore, with age, gut compartmentalization is altered<sup>391</sup>, the epithelial structure is damaged and the intestinal barrier integrity is lost, likely resulting from dysfunctional septate and tricellular junction<sup>372,392-396</sup>. In some instances, repeated activation of stem cell proliferation leads to the exhaustion of the stem cell pool<sup>277</sup>. Studies of aging and its impact on *Drosophila* gut and ISCs have been recently comprehensively reviewed<sup>397</sup>. At the cellular level, aging is impacting nuclear organization, metabolism and genome stability.

#### *Nuclear organization and epigenetic regulation during aging*

The nuclear organization and chromatin structure were shown to be altered upon aging in different model systems<sup>292</sup>. ECs were shown to undergo alteration of chromatin organization with changes in histone H3K9me3 and HP1<sup>398</sup>. Similarly,

as mentioned above, Hey and NON-Stop promote genomic organization in ECs and their expression in ECs decreases during aging, correlating with changes to Lamin structures<sup>348,349</sup>.

How the chromatin landscape and nuclear organization changes in stem cells during aging has only begun to be investigated. Recent ATAC-seq data suggests that aging leads to mild alteration of chromatin accessibility at promoters enriched for binding motifs of Polycomb interactors. The authors suggested that changes in ISC H3K27me2 levels, catalyzed by the Polycomb Repressive Complex 2 (PRC2), led to a bias toward EE fate by upregulation of EE-specific gene expression<sup>351</sup>. In addition to previous characterization of chromatin in young<sup>357</sup> and aged<sup>351</sup> stem cells, further studies are needed to establish the full extent of epigenetic alterations in ISC and differentiated cells and to unravel direct targets of PRC2 in EE regulation throughout adult life.

### **Proteostatic and environmental stress**

Loss of proteostasis has been associated with aging in numerous tissues and organisms<sup>292</sup>. Upon induction of protein aggregates thought to disrupt proteostasis, ISCs undergo a cell cycle arrest, which depends on Keap1-Nrf2/CncC and the cyclin-dependent kinase inhibitor, Dacapo. This "proteostatic checkpoint" allows upregulation of proteasome genes and clearance of aggregates. Interestingly, in aged ISCs, cell cycle arrest no longer occurs upon aggregate induction, suggesting a loss of this checkpoint. Overexpression of Nrf2/CncC pathway improves age-related increase in cell proliferation and decline in gut barrier function<sup>399</sup>.

### **Changes of mitochondrial metabolism with age**

As mentioned above, mitochondria have important roles in ISCs and their quality control during aging is essential in *Drosophila* ISCs<sup>400</sup> as well as in other cell types in a variety of species during aging<sup>292</sup>. The regulation of mitochondrial metabolism is crucial for ISC proliferation and differentiation<sup>342</sup>, and

downregulation of the mitochondrial pyruvate carrier promotes ISC proliferation<sup>401,402</sup>. Upon tissue damage, ISC proliferation requires an intensification of mitochondrial activity; the increase of mitochondrial  $\text{Ca}^{2+}$  drives restoration of ATP levels, reduction of NADH, and reactive oxygen species production by the respiratory chain<sup>403</sup>. In aged ISCs, mitochondrial  $\text{Ca}^{2+}$  levels as well as ATP production are reduced marking a switch from mitochondrial respiration to aerobic glycolysis. This Warburg-like metabolic reprogramming drives ISC hyperproliferation and resembles the metabolic rewiring of oncogene-transformed cells<sup>403</sup>. Additional studies will be required to understand how changes in mitochondria and their metabolic function during aging feedback on cell-cell signaling pathways to impact aging.

### *DNA damage and genome instability*

Previous work has demonstrated increased marks for oxidative DNA damage and DNA breaks during aging in intestinal stem cells<sup>404,405</sup>. A requirement of the DNA damage response mediated by ATR and ATM was also demonstrated for ISC proliferation and maintenance during aging<sup>406</sup>. In ECs, deficiency in the DNA damage and DNA repairs responses is driving stem cell DNA damage and hyperproliferation<sup>407</sup>. A recent study also showed an aged-related increase in the level of O-GlcNac, a nutrient-driven post-translational modification of proteins associated with oxidative stress<sup>408</sup>. This, in turn, leads to DNA damage upregulation in ISCs upon oxidative stress and high sugar diet, mediated at least in part by O-GlyNacylation, therefore linking nutrient uptake and glucose metabolism to DNA damage<sup>408</sup>.

In addition to increased marks of DNA damage in ISCs during aging, our previous studies found a rise in somatic mutations in the gut<sup>409</sup>. Notably, a functional consequence of this is that 10–15% of aged male flies acquire tumor-like neoplasia due to the spontaneous mutation of the X-linked tumor suppressor *Notch*<sup>409</sup>. *Notch* was found to be inactivated mostly by deletions (~2–500kb) or

more complex genomic rearrangements of the same size range. The complex rearrangements, which are combinations of deletion, inversions and duplication/triplications, contributed to 26,6% of the structural variants at the *Notch* locus (8/30 samples)<sup>410</sup>. Recently, we have expanded our characterization of the genome alterations affecting aged intestinal stem cells<sup>410,411</sup>. Whole-genome sequencing of male neoplasia reveal that structural variants and point mutations occur genome-wide<sup>410</sup>.

In addition to structural variants and point mutations, de novo insertion of transposable elements (TEs) were also detected genome-wide as well as within the *Notch* locus, likely responsible for neoplasia formation in some instances<sup>411</sup>. Our finding suggested tissue-specific variation in TE subclass mobility, with some TEs being more mobile in the gut than the brain or germline<sup>411</sup>. Altogether, this demonstrates that the ongoing DNA damage in ISCs during aging can have significant genomic consequences and alter tissue homeostasis.

#### IV. Conclusion

Studies over the last 15 years using the *Drosophila* adult midgut and ISCs as a model have uncovered fundamental underlying principles of stem cells and their interactions with the external environment and other tissues. Through candidate approaches, genetic screening, and scRNA-seq analyses, the field now has a solid framework of the transcriptional regulators control of lineage decisions and proliferation. Nevertheless, future studies adapting novel genomic new techniques in this system will undoubtedly provide a better characterization of epigenetic landscapes and their functional relevance in the ISC lineage. The last few years have also demonstrated the versatility of the gut to restore homeostasis after tissue damage and adapt the tissue to organismal needs. This process requires a complex integration of signals in the different cell types and tissues to tune stem cell proliferation and adapt cell differentiation and turnover. Recent advances in live-imaging technologies will contribute to the understanding of the

dynamic nature of cell death, cell proliferation and tissue remodeling in real time<sup>412,413</sup>. Future studies will also tease apart the precise role that stem cell proliferation has on aging phenotypes in this tissue. As many age-related alterations in stem cells are rescued by the inhibition of stem cell proliferation, the extent to which aging phenotypes are primary defects or secondary ones, due to deregulated proliferation program needs to be clarified. Importantly, aging is also characterized by intestinal dysbiosis. The insight gained on processes controlling adult stem cells using the *Drosophila* midgut will provide important first principles and testable hypotheses for addition studies in other model organisms.

## **E. Contribution of this PhD work**

In the next chapters, I present the data obtained during my PhD in the form of two paper manuscript. Upon arriving in the lab, I became interested in characterizing the mutational landscape of intestinal stem cells and understanding the factors causing DNA damage and mutation accumulation in the gut.

### ***In the first part of my thesis work I investigated:***

- ◇ The timing of *Notch* inactivating events driving neoplasia formation.
- ◇ The consequences of aging on the stem cell response to DNA damage and on the DNA repair efficiency in progenitors.
- ◇ The landscape of somatic mutations affecting ISCs, essentially in collaboration with K. Siudeja, N. Riddiford and L. Alzouabi.
- ◇ The contribution of the DNA polymerase Theta in neoplasia formation and the spontaneous generation of structural variants and point mutations.

Some of the results presented in this part are already published in Siudeja et al., and Riddiford et al, (see Chapter 2, Figure 3 and Figure 3 Supplementary 1). However, I considered important to present these data, to have a better understanding of the whole work, and to specifically highlight my contribution. We plan to reorganize the manuscript presented in order to submit a new paper that would focus on the unpublished data.

Interestingly, we uncovered several mechanisms that constitute the foundation to new investigations conducted independently by other scientists in the lab.

*In the second part of my thesis work*, I investigated the consequences of replication stress on ISC maintenance and tissue-specific sensitivity to nucleotide depletion using the midgut and the developing wing as comparative models. More specifically:

- ◇ I developed a cell-specific approach to induce replication stress
- ◇ I demonstrated the high sensitivity of ISCs to nucleotide depletion, leading to extensive DNA damage and stem cell loss
- ◇ I showed that progenitors of the wing disc benefit from gap junctions to buffer nucleotide levels and limit replication stress
- ◇ I characterized the expression and localization of gap junction in the midgut, with the help of M. El-Hajj.
- ◇ I concluded that stem cells are devoid of the identified nucleotide buffering mechanism, thus explaining their sensitivity to replication stress.

This part (Chapter 3) constitutes a new manuscript that we aim to submit soon

CHAPTER 2 : STEM CELL  
MUTATIONS AND POL THETA  
CONTRIBUTION TO MUTAGENESIS

# Characterization of *Drosophila* intestinal stem cell mutations and Pol Theta contribution to mutagenesis.

Benjamin Boumard<sup>1</sup>, Nick Riddiford<sup>1</sup>, Allison Bardin<sup>1</sup>

<sup>1</sup> Institut Curie, PSL Research University, CNRS UMR 3215, INSERM U934, Stem Cells and Tissue Homeostasis Group, Paris, France.

## I. INTRODUCTION

Biological tissues are subject to various exogenous and endogenous factors causing DNA damage and driving irreversible changes of the genome resulting from lack of repair or erroneous DNA repair pathway usage. Genome instability is a hallmark of cancer<sup>1</sup>, but recent studies also demonstrated the rich landscape of somatic mutation accumulating in adult stem cells of healthy tissues<sup>2</sup>. These mutations likely contribute to the age-related functional decline of adult stem cells and to the early steps of tumorigenesis. The sequencing of cancer genomes and dissection DNA repair pathways have provided insight into the DNA damage and repair factors contributing to tumor-specific mutational signatures. However, whether stem cell susceptibility to genome instability changes with aging and how it is affecting stem cell function is unclear. Likewise, the sources of DNA damage and the nature of the erroneous DNA repair pathways involved in mutation accumulation in adult stem cells are not fully understood.

The fruit fly intestine, comparable to that of mammals, maintains tissue homeostasis during adult life through self-renewal and differentiation of intestinal stem cells (ISCs)<sup>3,4</sup>. Our previous study indicated that *Drosophila* intestinal stem cells frequently acquire somatic mutations during aging<sup>5</sup>. We found that an important consequence of somatic mutation is that 10–15% of aged male develop neoplasia characterized by the accumulation of ISC and enteroendocrine cells (EEs), a consequence of X-linked tumor-suppressor gene *Notch* inactivation, mirroring tumor initiation. Thus, we have established the midgut as a good model to study genome instability in adult stem cells.

Here we explore factors driving genome instability in adult stem cells. We find that mutations driving neoplasia formation through somatic *Notch* inactivation arise in adult stages. By applying whole-genome sequencing to adult guts, our data presented here and in parallel studies<sup>6,7</sup> suggest a role for the Alternative Non-Homologous End Joining (Alt-EJ) pathway in generating structural variants in the fly intestine<sup>8</sup>. We further analyzed the role of this pathway focusing DNA Pol $\theta$ , the DNA repair polymerase implicated in Alt-EJ, and define its contributions to somatic mutation in intestinal stem cells.

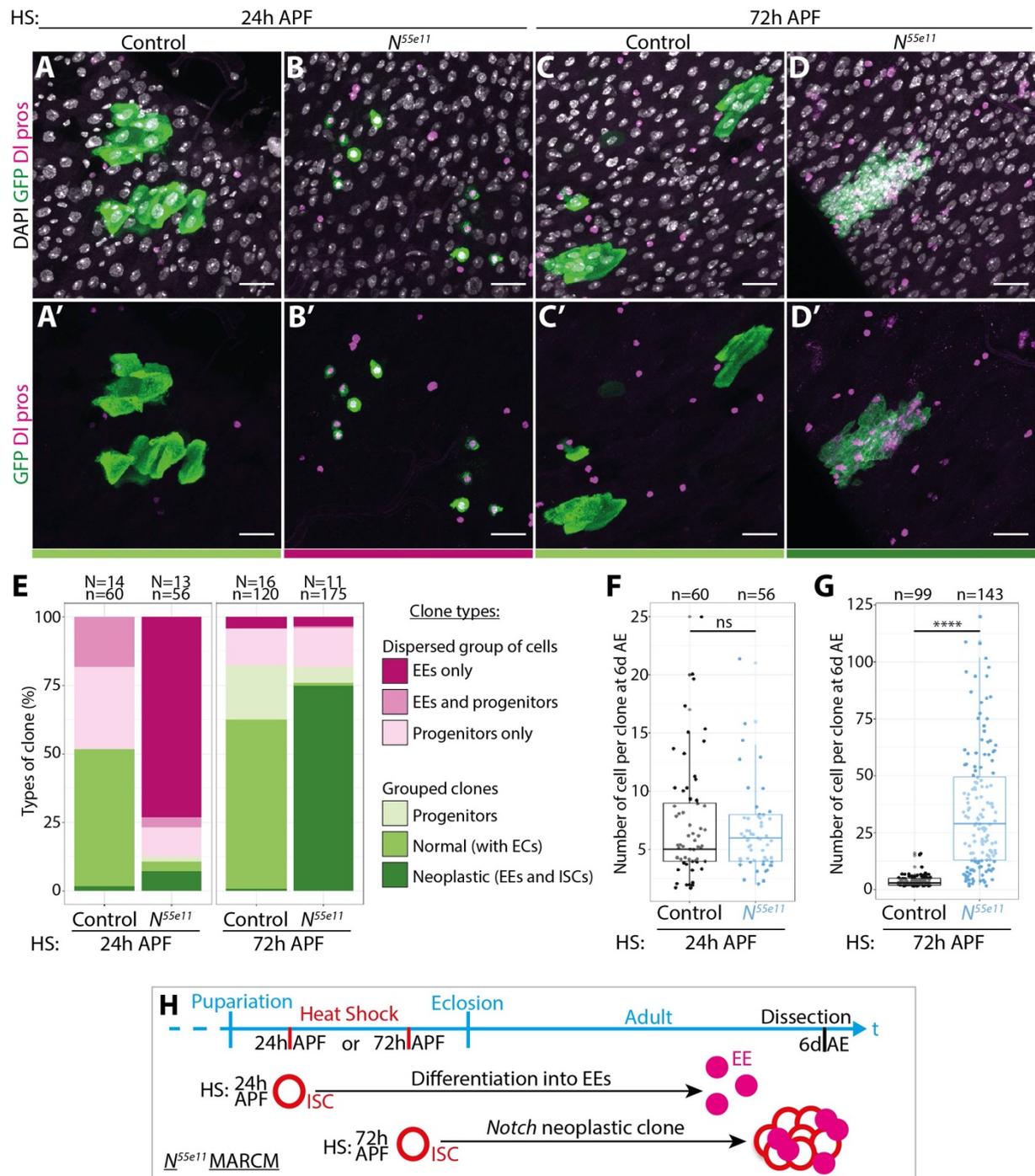
## II. Results

### *Notch inactivation occurs in adult stem cells to drive neoplasia*

In order to define factors driving stem cell DNA damage we first wanted to determine the timing of mutation of adult stem cells. Previous work from the lab showed that spontaneous neoplasia are detected in 10–15% of the aged male midguts and they arise due to the inactivation of the X-linked tumor suppressor gene *Notch*. Neoplasia were undetected in young flies and their onset increased with aging, with 0.3% at 2 weeks, 4% at 4 weeks, and 12% at 6 weeks<sup>5</sup>. Because of this observation we suspected that *Notch* inactivation happened in adult stem cells. However, we could not exclude that *Notch* was inactivated prior to adult life during earlier developmental stages. For example, larval stem cell precursors could acquire mutations and produce adult stem cells that remain quiescent until later in adult life or divided too slowly to produce a detectable neoplasia in young adults.

Therefore, to define when the *Notch* mutations that we detect could have happened, we followed the growth of GFP marked *Notch* mutant clones induced at early or late pupal stages 24h or 72h after pupal formation (APF) and examined the outcome in adult guts 6 days after adult eclosion (Fig.1 A–D). *Notch* inactivation would likely drive neoplasia growth early on, leading to detectable ISC accumulation at the dissection time. We first induced clones in the early pupal

precursors of ISCs at 24 hours after pupal formation (APF). Control clones were composed mostly of groups of polyploid enterocytes and progenitor cells (Fig. 1A–A', E). In contrast, *Notch* mutant clones appeared mostly as clusters of



**Figure 1. Notch inactivation occurs in adult stem cells to drive neoplasia. A–D'.** Representative pictures of adult female midgut 6 days after eclosion with GFP MARCM clones of Control (A–A', C–C') and *N<sup>55e11</sup>* (B–B', D–D') for clones induced 24h after pupal formation (APF, A–B') or 72h APF (C–D'). Stained with DAPI, GFP, Delta (ISC) and Pros(EEs), scale bar: 20µm. The color indicated below the pictures refers to the type of clones defined in E. **E.** Proportion of the types of clones obtained in Control and *N<sup>55e11</sup>* flies, for clones induced 24h or 72h APF. **F.** Clone size for clones induced 24h APF. **G.** Clone size for clones induced 72h APF representing only the main category of clones as defined in E, i.e. “Normal” clones in the Control and Neoplastic clones in *N<sup>55e11</sup>*. **H.** Model of the outcome of the experiment. Statistics: Welch test. \*\*\*\*:  $p < 0.0001$  N=number of guts, n=number of clones.

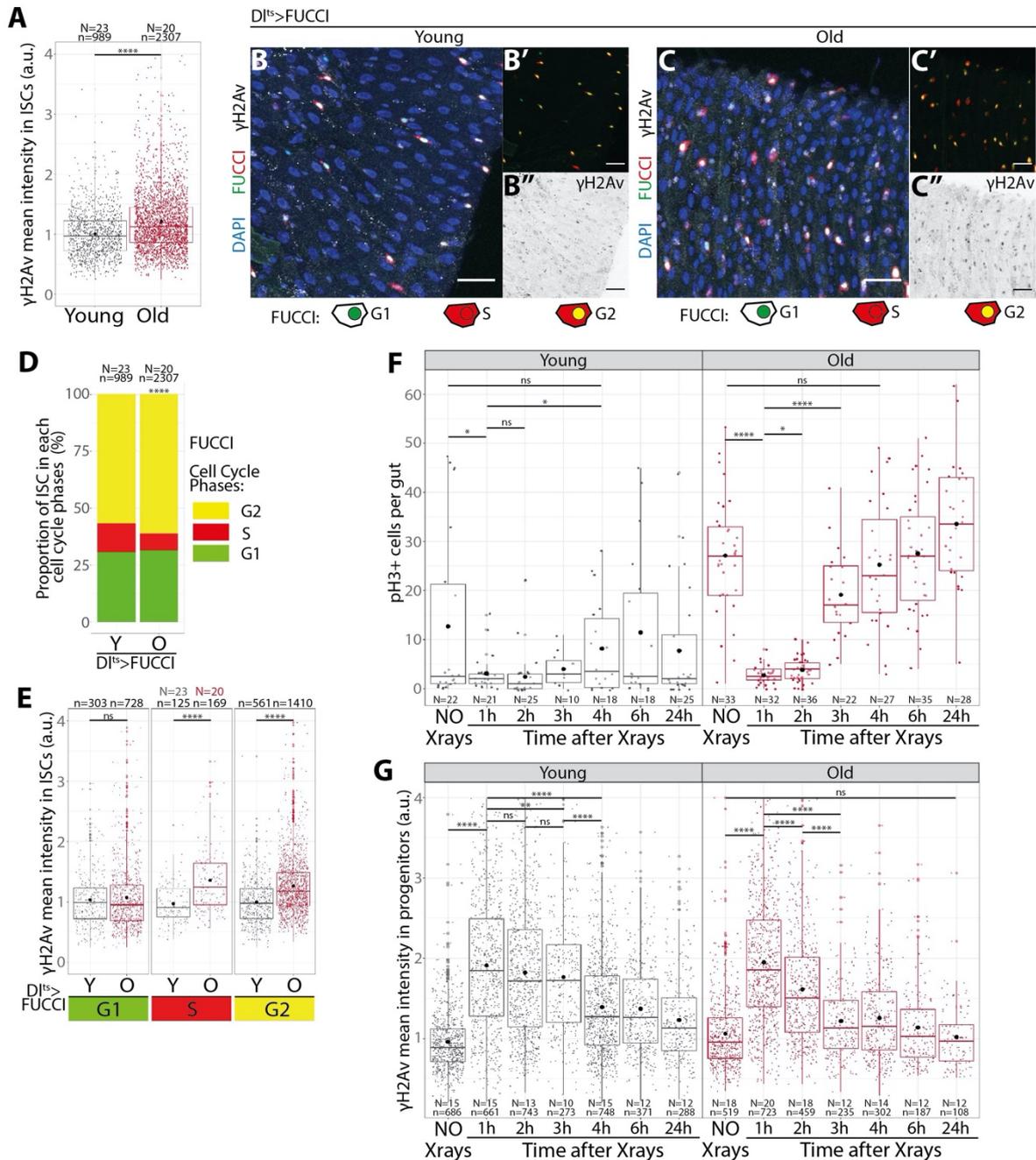
dispersed Enteroendocrine cells marked by Prospero (Fig. 1B–B', E). The clones had around 6 EE cells (Fig. 1F, H), likely indicating ISC mutant cells divided a few times prior to differentiation. These observations are consistent with previous studies<sup>9–11,11</sup> and indicate that when *Notch* mutations are induced in early pupal stages, the resulting cells differentiate into EE cells by adulthood. Therefore, these data argue against adult neoplasia being caused from DNA damage and mutation during early pupal stages.

We next wanted to test whether mutations in *Notch* giving rise to adult neoplasia might arise during late pupal stages that could give rise to adult neoplasia. Control clones induced at 72 hours APF and dissected 6d after eclosion, were composed mostly of grouped polyploid enterocytes and progenitor cells (Fig. 1C–C', E). However, *Notch* mutant clones displayed accumulation of ISC and EE cells, forming neoplasia (Fig. 1D–D', E). Furthermore, the mutant clones had 5 times more cells than control clones at the time of dissection (Fig. 1G). Most of the *Notch* mutant clones were comprised of >20 tightly grouped cells, these induced neoplasia were easily detectable and identifiable in the young adult midgut (Fig. 1H). Since we did not detect large spontaneous neoplasia prior to 2 weeks of adult age (Siudeja et al., 2015), we conclude that somatic mutation of *Notch* is unlikely to arise in late pupal stages as indeed we believe these would have produced easily detectible neoplasia. Altogether, these data indicates that somatic mutation of *Notch* causing neoplasia formation occur in adult stem cells. While we cannot rule out that other spontaneous mutations elsewhere in the genome arise during larval stages, our findings suggest that somatic mutations accumulate in adulthood and drive neoplasia formation. We therefore next wanted to understand what factors in adult stages might promote DNA damage or limit DNA repair.

### **Absence of age-related decline in the ISC in response to DNA damage**

Our previous data and the findings above indicated that somatic mutations and neoplasia occurrence increased with age during adult life. In addition,

previous studies reported an increase of DNA damage marks in ISCs upon aging<sup>12,13</sup>, which we also observed:  $\gamma$ H2Av was detected at slightly higher mean



**Figure 2. Absence of age-related decline in the ISC in response to DNA damage.** **A.**  $\gamma$ H2Av mean intensity in ISC nuclei of young and old flies (5 weeks old), the intensity is normalized to the mean value of the young group. **B-C'**. Representative images of young (B-B'') or old (C-C'') female midgut expressing the FUCCI reporter in the ISC with the *Delta<sup>ts</sup>* driver. With DAPI, RFP, GFP and  $\gamma$ H2Av staining. Pictures showing the FUCCI alone (B',C') and the  $\gamma$ H2Av staining alone (B'',C''). Scale bar: 25 $\mu$ m. **D.** Cell cycle proportion in ISC of young and 5 weeks old flies based on the FUCCI reporter, chisquare test. **E.**  $\gamma$ H2Av intensity in nuclei for each cell cycle phase in young and 5 weeks old flies, normalized to the mean value in the young flies. Black dot = mean. Welch test. **F.** PH3+ cells per gut in young and aged flies, without irradiation or 1-24h after X-rays. Welch test. Black dot = mean. **G.**  $\gamma$ H2Av intensity in progenitor cell nuclei in young and aged flies, without irradiation or 1-24h after X-rays. Normalized to the mean intensity in the young non-irradiated flies. The progenitor cells were determined by their small nuclear size and absence of Pros staining. It thus include ISC and EBs. Welch test. \*:  $p < 0.05$ ; \*\*:  $p < 0.01$ ; \*\*\*:  $p < 0.001$ ; \*\*\*\*:  $p < 0.0001$ . N = number of guts/flies, n=number of cells.

intensity levels in aged compared to young ISCs (Fig. 2A). We therefore explored whether there may be an age-related change in DNA repair of ISCs, which could account for an increased presence of DNA damage-associated marks and might give insight into the underlying drivers of DNA damage.

We first analyzed the cell cycle properties of young versus aged stem cells using the FUCCI system (Fig. 1B–C’), reasoning that accumulation of DNA damage may alter check point activation and cell cycle properties. We found that the proportion of stem cells in G2 increases with aging, at the expense of S-phase cells (Fig. 2D). In addition, we observed that increased marks of DNA damage were mostly affecting S and G2 phase cells (Fig. 2E), which could be indicative of replication related DNA damage. The observed increased proportion of stem cells in G2 in aged guts could be an indication of checkpoint activation in response to DNA damage. Alternatively, this could be due to an increased number of cycling cells in the gut, as previous studies have demonstrated that during aging, stem cell proliferation rates increase<sup>14</sup>. Our data suggest that upon aging, there are more stem cells with DNA damage and a higher proportion in G2.

The increase in  $\gamma$ H2Av marks in stem cells with aging could come from higher levels of DNA damage in aging stem cells or to the longer persistence of DNA damage marks after the lesions because of ineffective  $\gamma$ H2Av mark removal<sup>15</sup> or slower DNA repair. An age-related decline in DNA repair efficiency was observed in other stem cell models in *Drosophila*<sup>16</sup> and mammals<sup>15</sup>. Consequently, we decided to assess aged stem cell response to induced DNA damage and follow the kinetics of DNA repair. As reported previously, we found that aged midguts had significantly higher levels of proliferation (Fig. 2F)<sup>14</sup>. In response to X-ray induced DNA damage, the number of ISCs in mitosis, marked by phospho-Histone H3, was decreased in both young and aged flies 1 hour after irradiation. The numbers of mitotic ISCs returned to uninduced levels 3 hours after irradiation for aged flies and 4 hours after irradiation for young flies (Fig. 2F). These data suggest intestinal stem cells responded to DNA damage by activating mitotic checkpoints, which are released 3 hours after irradiation. Consistent with DNA

repair occurring rapidly after irradiation in both young and old ISCs,  $\gamma$ H2Av staining in progenitors was doubled 1h after irradiation, then progressively decreased with recovery time (Fig. 2G), as previously reported<sup>17</sup>. We found that  $\gamma$ H2Av decreased fast between 1h and 3h after irradiation in aged flies but slightly later in young flies. This kinetics also correlated with the observed proliferative activity of the gut (Fig. 2F). Thus, aged ISCs do not show any apparent decline in DNA repair kinetics or in checkpoint activation compared to young ISCs.

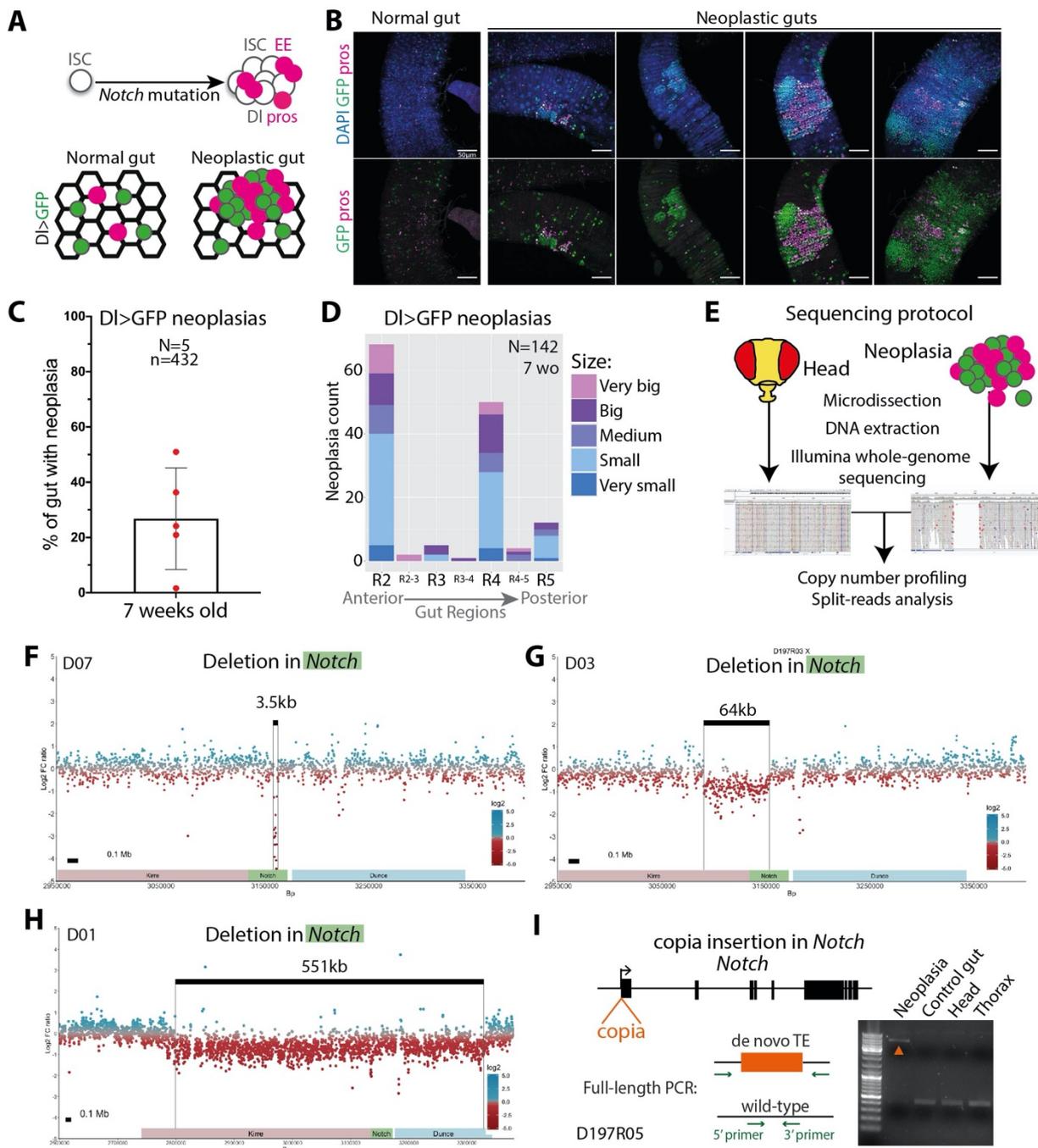
This led us to investigate the possible causes of DNA damage driving mutation accumulation and *Notch* inactivation in the intestinal stem cells using genomic sequencing approaches.

### *Neoplasia marked by ISC-specific GFP expression are caused by Notch deletions or transposable elements insertions*

Previous work from the lab had already demonstrated the frequent occurrence of somatic mutation in intestinal stem cells<sup>5</sup>. To gain insight into the mechanism driving mutagenesis in the ISC, we aimed at characterizing the mutational profile of aged ISC. We first applied whole-genome sequencing methods to better understand potential drivers of ISC mutation. K. Siudeja in our lab developed a methodology for micro-dissection, DNA extraction and whole-genome sequencing of neoplasia from male flies<sup>5</sup>. Using the *ProsperoGal4>UAS-2XGFP* fly lines, she could detect neoplasia specific accumulation of EE cells in live tissues, and subsequently dissect the estimated neoplastic mass. Whole-genome sequencing of neoplasia allowed the identification of somatic mutations through bioinformatic pipelines developed by N. Riddiford (see Riddiford et al. for more details).

To reinforce our analysis and avoid any bias that could emerge from a peculiar fly line, here we aimed to sequence neoplasia from another genetic background. We reasoned that we should be able to detect neoplasia using ISC-specific driver *DeltaGal4* in combination with *UAS-nlsGFP* from male flies in which spontaneous inactivation of the X-linked gene *Notch* occurs (Fig. 3A). We first

assessed neoplasia formation in this new genetic background and, as expected, could detect neoplasia as GFP marked ISC's interspersed with EE cells (Fig. 3B). We found neoplasia in more than 25% of 7 weeks old male guts, although with high variability depending on the experiment (Fig. 3C). Similar to different genetic background examined before<sup>5</sup>, neoplasia were found throughout the gut, but mostly in the R2 and R4 regions, which are the largest and most proliferative compartments of the gut (Fig. 3D)<sup>18,19</sup>. Importantly, a significant proportion of



**Figure 3. Neoplasia marked by ISC-specific GFP expression.** **A.** Model of neoplasia identification in the gut using the *Delta>GFP* lines. **B.** 7 weeks old guts from male *Delta>GFP* flies without or with neoplasia of different size. Neoplasia are detected by accumulation of GFP+ ISC and Pros+ EEs. Scale bar: 50µm. **C.** Neoplasia frequency in 7 weeks old *Delta>GFP* male from 5 different aging cages, with mean and standard deviation. N= number of experiments, n = number of guts **D.** Region-specific distribution of neoplasia size in 7 weeks old *Delta>GFP* male. N=number of guts analyzed. **E.** Sequencing methodology and bioinformatic pipeline to identify somatic mutations comparing neoplastic genome with the genome of the head of the same flies. It displays examples of IGV screenshots obtained from read-alignment after sequencing. **F-H.** Log2 ratio of the read coverage between neoplasia and the head. A drop in the ratio is a deletion, as it shows a decrease in read copy number in the neoplasia compared to the head. In these *Delta>GFP* male samples, Notch was inactivated by small (D07, F.), medium (D03, G) or large (D01, H) deletions. **I.** In the sample D05, Notch was inactivated by copia TE insertion, the specific PCR amplification of the TE is found in the neoplasia but not in the other tissues of the same fly, PCR by K. Siudeja.

neoplasia reached a large size big enough for microdissection and whole-genome sequencing analysis (Fig. 3B, 3D).

We microdissected neoplasia, performed DNA extraction and sequenced 4 neoplastic samples in parallel with the head from the same fly as a control (Fig. 3E, samples D01, D03, D05 and D07). We then characterized the somatic mutations in *Notch* and genome-wide from neoplasia developing in the *Delta>nlsGFP* background. The bioinformatic pipeline developed by N. Riddiford uses both well-established approaches and newly developed tools to identify structural variants in the genome. The comparison of read-depth between the neoplasia and the head allows for copy numbers variants detection and the identification of read-mapping abnormalities is indicative of structural variant breakpoints. These datasets have also been described in Riddiford, et al, in which genome-wide structural variant and point mutation analysis was performed and in Siudeja, et al, in which genome-wide transposon mobility was characterized. Here, I will describe important findings to come out of this analysis, focusing primarily on the mutations inactivating *Notch*.

Structural variant analysis revealed that *Notch* was inactivated by small (Fig. 3F sample D07, 3.5kb), medium (Fig. 3G, sample D03, 64kb) or large deletions (Fig. 3H, sample D01, 551kb). One sample (D05) did not harbor evidence for a deletion in *Notch*, but the read-mapping analysis revealed a full-length *de novo* insertion of a *copia* retrotransposon in the first codon of *Notch*, likely responsible for the gene inactivation (Fig. 3I, this example is also described in Siudeja et al.).

Overall, the neoplasia-causing *Notch* mutations in the *Delta>nlsGFP* background had the same characteristics as the ones identified in the 31 samples from the *Pros>2XGFP* male flies (samples P1–P65). The results of the analysis of both genotypes were published in Riddiford et al., and Siudeja et al., 2021<sup>6,7</sup>.

### **Neoplasia in females are induced by various types of mutations**

We have now demonstrated extensively the type of mutations that can lead to *Notch* inactivation and neoplasia in males. *Notch* is located on the X chromosome, therefore, a single hit is sufficient to induce *Notch* loss of function in males. However, neoplasia were not observed in wild-type females carrying two wild-type alleles of *Notch* (Siudeja et al., 2015, n=519). While performing microdissection in *Delta>nlsGFP* flies, we realized that some aged female guts also exhibited neoplastic growth reminiscent of *Notch* inactivation. To better understand the causes of neoplasia development in females, we microdissected 4 female samples (D09, D11, D13 and D15). The whole genome-sequencing revealed 3 different mechanisms explaining *Notch*-like ISC accumulation detailed below.

In two samples, we identified structural variants in *Delta* (*DI*, Fig. 3 Supplementary 1A). *Delta* is a receptor of *Notch*, thus *Delta* inactivation in ISC drives neoplastic growth similar to *Notch* loss. As our Gal4 driver line contains a lethal P-element insertion in *Delta*, the structural variants likely inactivated the second functional copy, leading to loss-of-heterozygosity (LOH). In the sample D11, the structural variant was a 38.5kb duplication in the *Delta* coding sequences, likely disrupting gene function as it only amplified half of the gene. In the sample D15, *Delta* contained a 10.8kb deletion. Thus, LOH in females can arise by deletions or other structural variants, a frequent source of inactivation of *Notch* in males.

In another sample, we found evidence for spontaneous mitotic recombination occurring on the chromosome harboring the *Delta<sup>Gal4</sup>*. This is predicted to lead to a complete loss of *Delta* activity as the neoplasia inherited 2

copies of the null *Delta*<sup>Gal4</sup> allele. LOH was detected in sample D13 due to shift in parental SNP frequency from heterozygous (~0.5) to homozygous (<0.25 or >0.75) along the chromosome arm, starting at a position centromeric to *Delta* (Fig. 3 Supplementary 1B; these data are also presented in L. Alzouabi, et al, in prep). LOH events in different *Notch* pathway components have been detected in the lab before<sup>5,20</sup> and a detailed analysis of the mechanisms of LOH through mitotic recombination will be presented elsewhere (L. Alzouabi, et al manuscript in prep).

Finally, one sample showed evidence for the bi-allelic inactivation of *Notch* (sample D09; Fig. 3 Supplementary 1C). First, we could detect loss of heterozygosity (LOH) of the whole X chromosome, as detected by the shift in parental SNPs frequency from heterozygous to homozygous. This was associated with a drop in the coverage of the whole chromosome in the neoplasia compared to the head, suggesting that one X chromosome was lost entirely leading to monosomic aneuploidy. In addition of the loss of one X chromosome, a small 3.1kb deletion at the beginning of *Notch* likely inactivated the other allele. Altogether, this would be predicted to lead to a bi-allelic loss of *Notch*, likely driving neoplasia development. We expect such bi-allelic inactivation to be a very rare event as we had never seen neoplasia in females carrying two wild-type alleles of *Notch* before (n=519, Siudeja et al. 2015), although we had already sequenced a male neoplasia exhibiting a bi-allelic inactivation of the Notch pathway gene *kuzbanian* (Riddiford et al.). Altogether, these data indicate that numerous types of mutational processes can affect the stem cell genome during aging, but deletions and structural variants are predominant mechanisms of gene inactivation.

### **Microhomologies and short insertions are found at structural variants breakpoints**

In order to investigate further the potential underlying sources of DNA damage and erroneous DNA repair mechanisms in the adult gut, we analyzed the genomic features and breakpoints associated with structural variants at the *Notch*

locus and genome-wide (these data are presented in detail in Riddiford et al.). To summarize, several lines of evidence suggest a potential role for replication stress in *Notch* inactivation. First of all, complex rearrangements were identified and are reminiscent of the products of Microhomology-Mediated Break Induced Replication (MMBIR) or Fork Stalling and Template Switching (FoSTeS) mechanisms<sup>5,7,21-23</sup>. These are erroneous DNA repair mechanisms generating

#### IN *NOTCH*

SAMPLE	Type	Length (kb)	Position	Allele Freq.	Microhomology	Insertion
P3	COMPLEX	30	X:3169736-3199425	0.95	-	GTATACATGCAATATGATATGATAACACA
P31	COMPLEX	1.3	X:3125694-3126976	0.48	-	CCGTGTG, TGTC
P41	DEL	101.8	X:3033491-3135340	0.78	-	ACACTCC
P5	COMPLEX	235.1	X:3129531-3364620	0.28	C	CTTTG
P63	COMPLEX	4	X:3195652-3199676	0.24	-	TATGAAATCCTA
P9	DEL	3.8	X:3135326-3139096	0.74	A	GA

#### NON-*NOTCH*

SAMPLE	Type	Length (kb)	Position	Allele Freq.	Microhomology	Insertion
P1	DEL	2.5	3L:9892365-9894889	0.35	TA	-
P13	DEL	1.7	2R:21382619-21384283	0.28	ACTGCAAT	-
P17	DEL	0.4	2R:6153193-6153589	0.05	GCCTCCGT	-
P25	DEL	2.3	3L:11261338-11263598	0.31	T	-
P33	DEL	0.5	3R:24856498-24856996	0.26	TC	-
P41	DEL	0.1	3L:14059217-14059363	0.21	-	AAATCT
P41	DEL	1466.7	3R:18794445-20261187	0.22	-	ACGTAAGT
P49	DEL	2400.1	3R:28657288-31057415	0.21	-	AGGTTGCCCTTTTTTTTT
P53	DEL	1	3L:20444578-20445608	0.46	A	CT
P57	DEL	206.4	3L:326321-532735	0.32	TTT	-
P9	DEL	0.1	3R:11562988-11563037	0.35	A	-
P35	TANDUP	30.4	3R:31043462-31073816	0.16	T	-
P11	BND	11161.3	2L:5103277-16264626	0.03	C	-
P11	BND	10201.9	3R:11288316-21490211	0.03	ATAGA	-
P13	BND	0.8	2L:205135-205916	0.22	GCATCCAGC	-
P17	BND	0.1	2R:13505400-13505470	0.04	CGCAA	-
P17	BND	0.1	X:9302146-9302259	0.16	GCTGCAGCTGGCC	-
P17	BND	0.1	X:10498521-10498657	0.04	GTGGGCGGCA	-
P21	BND	0.1	2L:2860962-2861103	0.3	GTGAGCT	-
P21	BND	0.3	2L:3040831-3041097	0.23	GCC	-
P21	BND	0.3	2L:7263746-7264071	0.2	AATGCAGAC	-
P21	BND	0.2	3R:11480462-11480630	0.41	CGCGCGG	-
P33	BND	1852.9	2R:5306714-7159623	0.13	TATGTGAT	GT
P33	BND	0.4	X:3550817-3551238	0.06	TTTTTTT	-
P37	BND	10896.3	2L:3314020-14210317	0.02	GGATG	-
P37	BND	1371.8	2R:16371025-17742832	0.07	GT	-
P43	BND	0.1	2L:5318061-5318184	0.04	GCTGCGGCTGC	-
P53	BND	0.1	2L:16655322-16655452	0.05	TGCCCGCTGC	-
P7	BND	2448.4	2L:6755669-9204071	0.16	GAGATCAAA	-
P7	BND	5292.5	2L:7116469-12408973	0.08	ATTTCCCACTC	-
P7	BND	99.9	2L:18249668-18349554	0.08	GCTCAAG	-
P7	BND	712.4	2L:18885742-19598141	0.1	GCCA	-
P7	BND	8013.4	2R:7680308-15693666	0.06	CCA	-
P7	BND	11744	2R:8694440-20438390	0.08	TCGCC	-
P9	BND	9908	2L:2380873-12288901	0.06	TTTACGAA	-
P9	BND	3989.7	2L:2615845-6605512	0.04	GCC	-
P9	BND	721.5	2R:9418566-10140104	0.03	GCACT	-
P9	BND	4402.5	3L:3943272-8345790	0.03	TGCG	-
P9	BND	3040.8	3L:9039625-12080473	0.03	CTGCTGC	-
P9	BND	20652.2	3R:9861512-30513691	0.03	T	-
P9	BND	4194.5	3R:16378760-20573305	0.03	TTGGC	-
P9	BND	375.6	X:4697763-5073382	0.09	GATGGAAA	-
P13	COMPLEX	9184.5	2L:7193295 3R:16377837	0.27	GTCCATGG	-
P31	COMPLEX	25.4	X:17286052-17311472	0.27	C, CT	-
P37	COMPLEX	46.6	2L:13540579-13587180	0.1	-	AATTGCATGCATGGGTAAGGTTGC

Table 1. Structural Variants with Microhomologies and/or insertions at the breakpoints in male neoplasia.

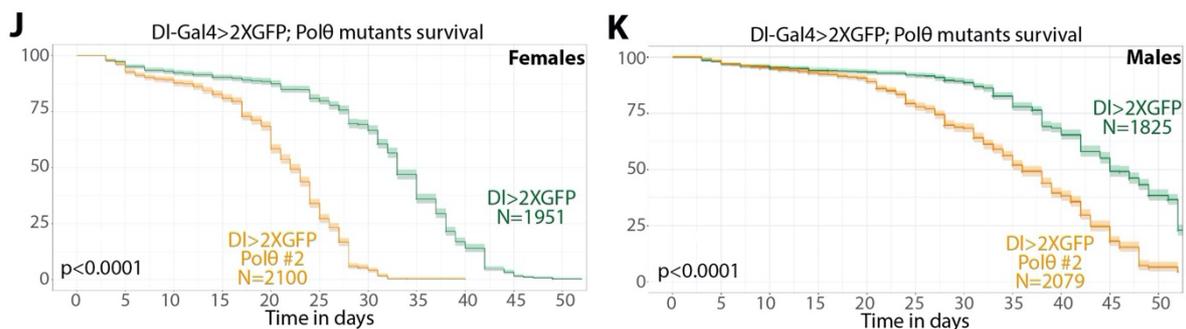
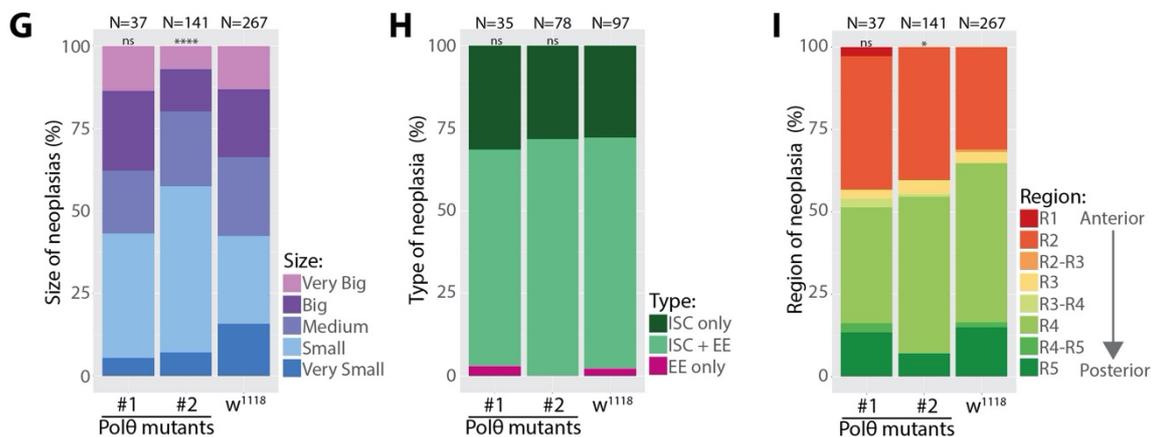
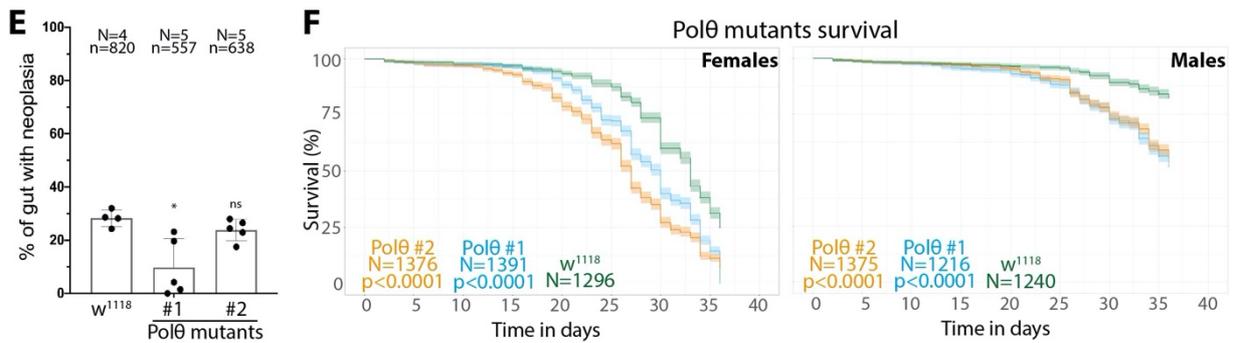
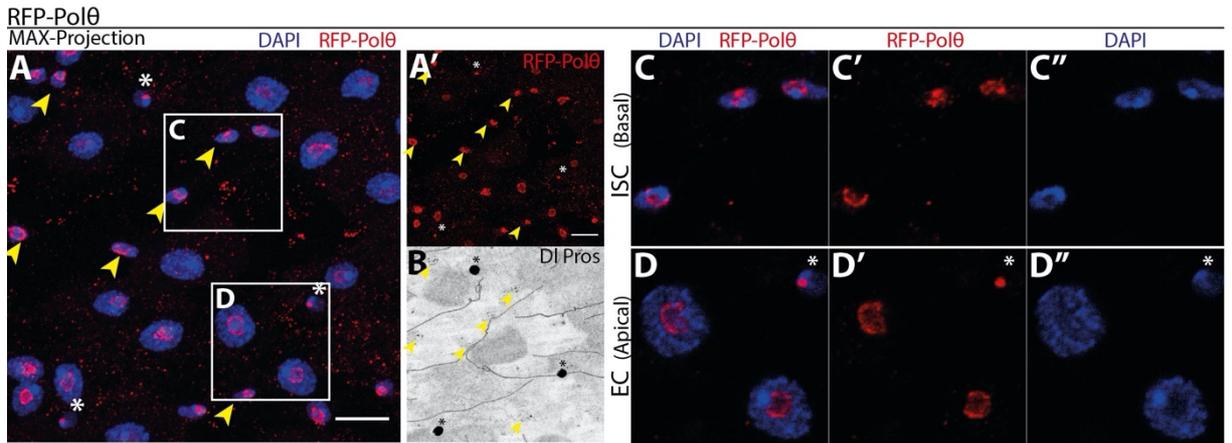
complex structural variants following replication fork collapse. Secondly, a *de novo* motif discovery approach uncovered an enrichment for poly(dA:dT) sequences at the breakpoints around *Notch*<sup>7</sup>. Similar sequences were shown to be preferential sites of replication fork collapse in condition of nucleotide unbalance<sup>24</sup>. Finally, the careful examination of the structural variants in *Notch* and genome-wide revealed the presence of microhomologies ranging from 2 to 14bp and short insertions at the breakpoints of deletions, complex events, inversion and translocations in male neoplasia (see a subset in Table 1 and supplementary Table 1). The presence of insertions indicates a DNA repair mechanism relying on DNA synthesis, and together with microhomologies it suggests a role for the Alternative Non-Homologous End Joining (Alt-EJ) pathway.

In higher eukaryotes, Alt-EJ is dependent on DNA Polymerase Theta (Pol $\theta$ ), which joins DNA double strand-breaks together using small (4–10 bp) stretches of microhomology (Microhomology-Mediated End Joining; MMEJ)<sup>25</sup>. Pol $\theta$  is a conserved DNA repair polymerase in higher eukaryotes encoded by the gene *POLQ* in mammals and *mus308* in *Drosophila*. The role of Pol $\theta$  in microhomology-mediated end joining was first demonstrated in *Drosophila*<sup>25,26</sup>. Pol $\theta$  uses short micro-homologies combined with DNA synthesis, which often leads to templated and non-templated insertions at deletion breakpoints. In addition, Pol $\theta$  could be involved in DNA repair after replication fork collapse<sup>27,28</sup>. This led us to investigate the role of Pol $\theta$  in structural variants formation and point mutations in *Drosophila* intestinal stem cells.

### **Pol $\theta$ is expressed in intestinal stem cells but is not required for neoplasia formation**

To understand the contribution of Pol $\theta$  to DNA repair and somatic mutation accumulation in *Drosophila* intestinal stem cells, we first surveyed available expression data. Comparable to most DNA repair genes, *mus308* was expressed at low levels in every cell types of the gut (rpkm = 0.167 in ISC and 0.411 in EC), and its transcripts levels were increased upon tissue injury induced by pathogenic

bacteria (rpkm = 2.5 in ISC and 5.4 in EC)<sup>19</sup>. However, transcript levels do not necessarily reflect the protein load of Polθ in the gut and its contribution to DNA repair. In addition, low levels of DNA repair gene expression are often sufficient for preserving genome stability and repair DNA damage. To verify Polθ expression,



**Figure 4. *Polθ* is not required for neoplasia formation. A-D''.** RFP-*Polθ* male midgut. Max-projection or DAPI and RFP -*Polθ* (A) or RFP-*Polθ* alone (A'), scale bar: 10μm. Pros and Delta staining (B) marking respectively EE cells (asterisks) and ISCs (yellow arrowheads). Basal plan of the epithelium with ISCs (C'-C'') and apical plan with ECs (D-D''). **E.** Neoplasia frequency in 5 weeks old *w<sup>1118</sup>* and *Polθ* transheterozygotes mutants for 4 or 5 aging experiments, mean and standard deviation. Welch test, N = number of experiment, n=number of dissected guts. **F.** Survival plots of *w<sup>1118</sup>* and *Polθ* transheterozygotes mutants. **G.** Neoplasia size in 5 weeks old *w<sup>1118</sup>* and *Polθ* transheterozygotes mutants, chisquare test **H.** Neoplasia types in 5 weeks old *w<sup>1118</sup>* and *Polθ* transheterozygotes mutants, chisquare test. **I.** Gut regions with neoplasia in 5 weeks old *w<sup>1118</sup>* and *Polθ* transheterozygotes mutants, chisquare test. **J-K.** Survival plots of *Delta>2XGFP* and *Delta>2XGFP Polθ* transheterozygotes mutants. N= number of flies or dissected guts. \*: p < 0.05; \*\*: p < 0.01; \*\*\*: p<0.001; \*\*\*\*: p<0.0001.

we generated through CRISPR mediated insertion, a line expressing a RFP-*Polθ* fusion protein at the endogenous promoter. We found RFP-*Polθ* to be localized in the nucleus of all the cell types in the midgut (Fig. 4A-B). RFP-*Polθ* intensity appears stronger in the low DAPI density regions of the nucleus, likely the nucleolus (Fig. 4C-D''). The nucleolus is known to participate in the regulation of DNA replication, DNA repair and stress response by segregating various factors of these pathways and limiting their diffusion and interaction with other partners<sup>29,30</sup>.

We then investigated the contribution of *Polθ* to neoplasia formation by quantifying the neoplasia incidence in *Polθ* mutant flies. *Polθ* mutants are homozygous viable, and several mutants have been characterized before, among which *mus308<sup>2003</sup>* is a nonsense mutation resulting in a truncated protein lacking the polymerase domain<sup>25,31</sup> and *mus308<sup>4</sup>* is a 14kb deletion of *mus308* and its neighboring gene *Men<sup>32</sup>*. To avoid a biased effect of a specific background, we isogenized mutants in the *w<sup>1118</sup>* background and used them as transheterozygotes. From the isogenization we kept several lines for each mutant and used them in two different transheterozygote combination ( $\frac{mus308^{2003}}{mus308^4}$  #1 and #2). Similar to the control flies (*w<sup>1118</sup>*), *Polθ* mutant males spontaneously acquired neoplasia with aging (Fig. 4E), although with different frequency depending on the combination of the isogenized lines combined. We also found that both females and males of the mutants had a shorter lifespan (Fig. 4F). We then analyzed the size, type and regions specificity of neoplasia (Fig. 4G-I). Overall, the nature of neoplasia in the mutants and the control were similar. Thus, *Polθ* is not required for neoplasia

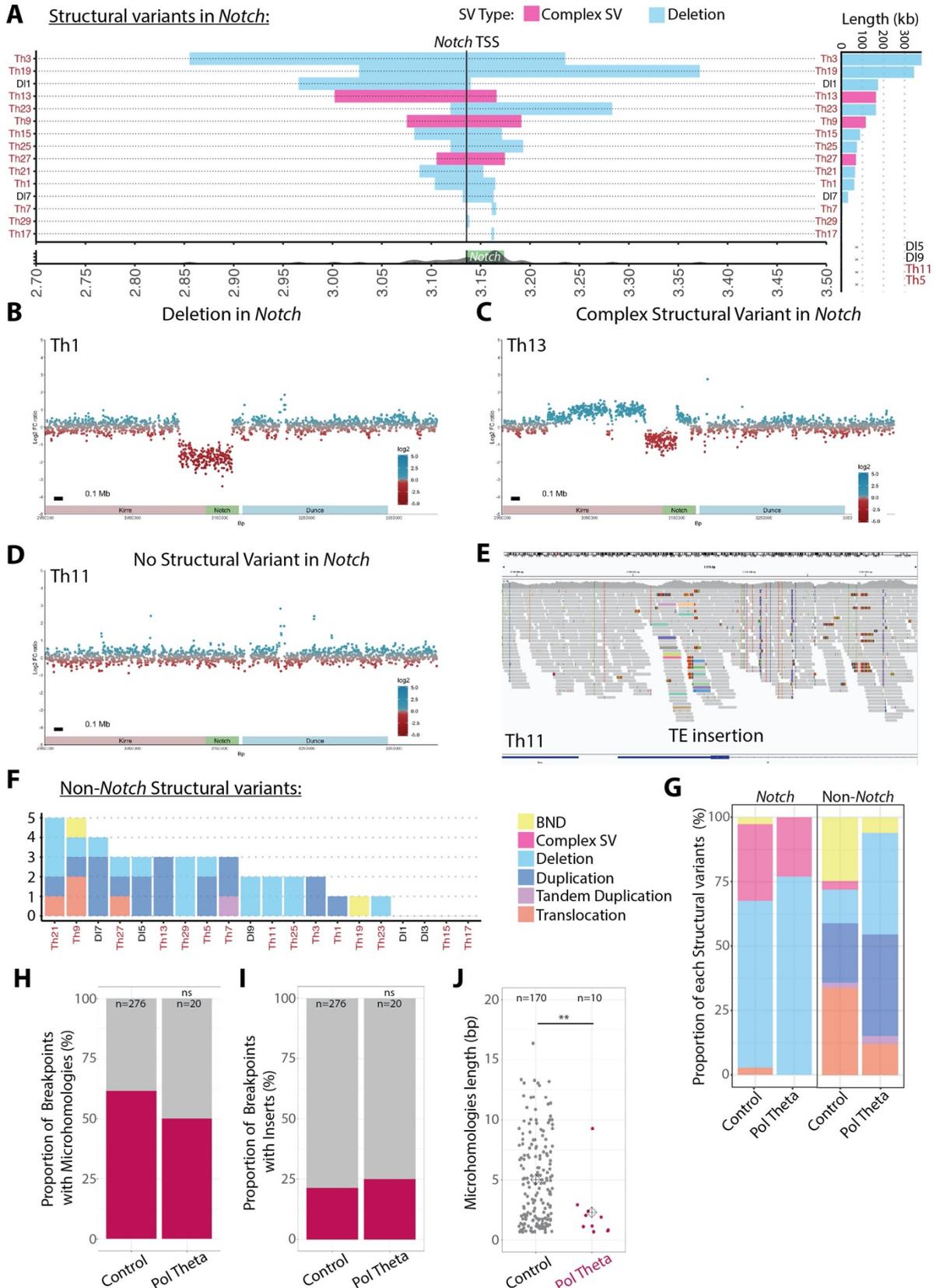
formation in the midgut. However, we hypothesized that other mechanisms would repair the DNA damage in absence of *Polθ*, which could still inactivate *Notch* through deletion. If this was the case, we would predict that there may be different mutational signatures in *Polθ* as compared to those of control flies.

To better characterize the involvement of Alt-EJ in intestinal stem cell mutagenesis, we sequenced neoplasia from the *Polθ* mutant background and compared the mutational profile of mutant and control flies. To achieve this, we recombined *mus308<sup>2003</sup>* with *DeltaGal4*, and combined *mus308<sup>Δ</sup>* with *UAS-2XGFP* to generate mutant flies expressing GFP in the stem cells. As a control, we also sequenced isogenized *Delta>2XGFP* flies. As seen before, the transheterozygote mutants displayed a shorter lifespan than the control flies (Fig. 4J-K). We performed microdissection and whole genome sequencing from 5 new control samples (DI1-9) and 15 *Polθ* samples (Th1-29). Importantly, both *mus308<sup>Δ</sup>* and *mus308<sup>2003</sup>* mutations could be identified in genomic sequencing in all *Polθ* samples and none of the controls (Fig. 4 Supplementary 1). Subsequently, we characterized the mutational profile in *Notch* and genome-wide.

### *Polθ likely contributes to structural variant formation in the intestine*

We first examined *Notch* inactivation events responsible for neoplasia formation. In *Polθ* mutants, *Notch* was mostly inactivated by deletions or complex structural variants ranging from 2kb to 400kb (Fig. 5A-C, Fig. 5 Supplementary 1). Deletions were found in 10 samples and complex events in 3, consequently representing 23.1% of the structural variants. Another sample, did not exhibit evidence for copy number variants in the *Notch* locus (Fig. 5D), however, the read mapping analysis identified *de novo* transposable elements insertions inside the *Notch* coding sequence (Fig. 5E, Fig.5 Supplementary 2, 1/15 = 6.67%). In the newly sequenced non-mutant control samples (DI1-9), *Notch* was inactivated by deletions (DI1,7) or transposon insertion (DI5,9). There is no evidence for *Notch* mutation in the DI3 sample however the read-depth was low in DI3 consequently complicating structural variants analysis. Altogether, the events explaining

neoplasia development through inactivation of *Notch* in *Polθ* mutants flies were globally similar to those we previously found in different wild-type genetic backgrounds<sup>6,7</sup>.



**Figure 5. *Polθ* likely contributes to structural variant formation in the intestine.** **A.** Structural variants type and size at the Notch locus from *Polθ* mutants (Th1-29) and non-mutants (Dl1-9) neoplasia. **B.** Deletion in Notch, as seen by the drop in Log2 ratio of the read coverage between the neoplasia and the head. **C.** Complex rearrangement in Notch, seen by the succession of increase (duplication) and decrease of read copy number ratio. **D.** Absence of structural variant in Notch. **E.** De novo transposable insertion in Notch 5'UTR, the sequencing reads with paired-read mapping to the transposon are in pastel colors, while the paired-read mapping to the *Drosophila* genome are in grey. **F.** Genome-wide, non-Notch structural variants in *Polθ* mutant (Th1-29) and non-mutant (Dl1-9, black) neoplasia from the new sequencing. **G.** Comparison in structural variant types in the Notch locus and outside of Notch between *Polθ* mutants and a panel of control non-mutant neoplasia. **H-I.** Proportion of identified structural variants breakpoints with microhomologies (H) or inserts (I) between *Polθ* mutants and a panel of control non-mutant neoplasia. Fisher's test. **J.** Microhomologies length between *Polθ* mutants and a panel of control non-mutant neoplasia. n = number of structural variants considered. Welch's test. \*: p < 0.05; \*\*: p < 0.01; \*\*\*: p < 0.001; \*\*\*\*: p < 0.0001.

We then assessed structural variants at the genome-wide level in controls versus *Polθ* mutants (Fig. 5F). The number of events were quite low and we did not detect differences in frequencies between the *Polθ* mutants and the newly-sequenced control neoplasia. We therefore used the samples previously analyzed in Riddiford et al., to provide a larger dataset of control samples in order to compare with *Polθ* mutants. Consistent with our data above, we find that the *Notch* inactivation events are similar in type between the mutant and non-mutant (control) background (Fig. 5G). Genome-wide, the types of structural variants were more diverse in the non-mutant control background, likely because it incorporates a larger dataset (Fig 5G).

Finally, we examined genome-wide features related to the putative role of *Polθ* in structural variants formation. We found that an equal proportion of structural variants breakpoints exhibit microhomologies and insertions between the mutant and non-mutant control neoplasia (Fig. 5H-I). However, there was a difference in the length of microhomologies, with shorter microhomologies detected at the breakpoints of the *Polθ* mutant neoplasia. Microhomologies were shorter than 4bp in the *Polθ* mutant neoplasia while ranging from 1 to 14bp in the control neoplasia (Fig. 5J). This is consistent with previous studies supporting the idea that MMEJ works better with >4bp microhomologies<sup>25</sup>, although *Polθ* can also mediate the repair with shorter microhomologies. In the mutants flies, breakpoints with microhomologies <4bp are likely products of the classical Non-Homologous End Joining pathway which can accommodate breaks with 4bp overhangs<sup>33</sup>.

Additionally, we detected a small quantity of breakpoints displaying insertions in the *Polθ* mutants flies which made a size comparison difficult. Lastly, we found that SVs affecting *Notch* mostly had no microhomologies or microhomologies shorter than 4bp (not shown) in both mutant and non-mutant flies. However, at the *Notch* locus, breakpoints often displayed short insertions (Table 1) in the control flies, while none were detected in the *Polθ* mutants, suggesting that *Polθ* could also contribute to neoplasia initiating *Notch* mutations (not shown).

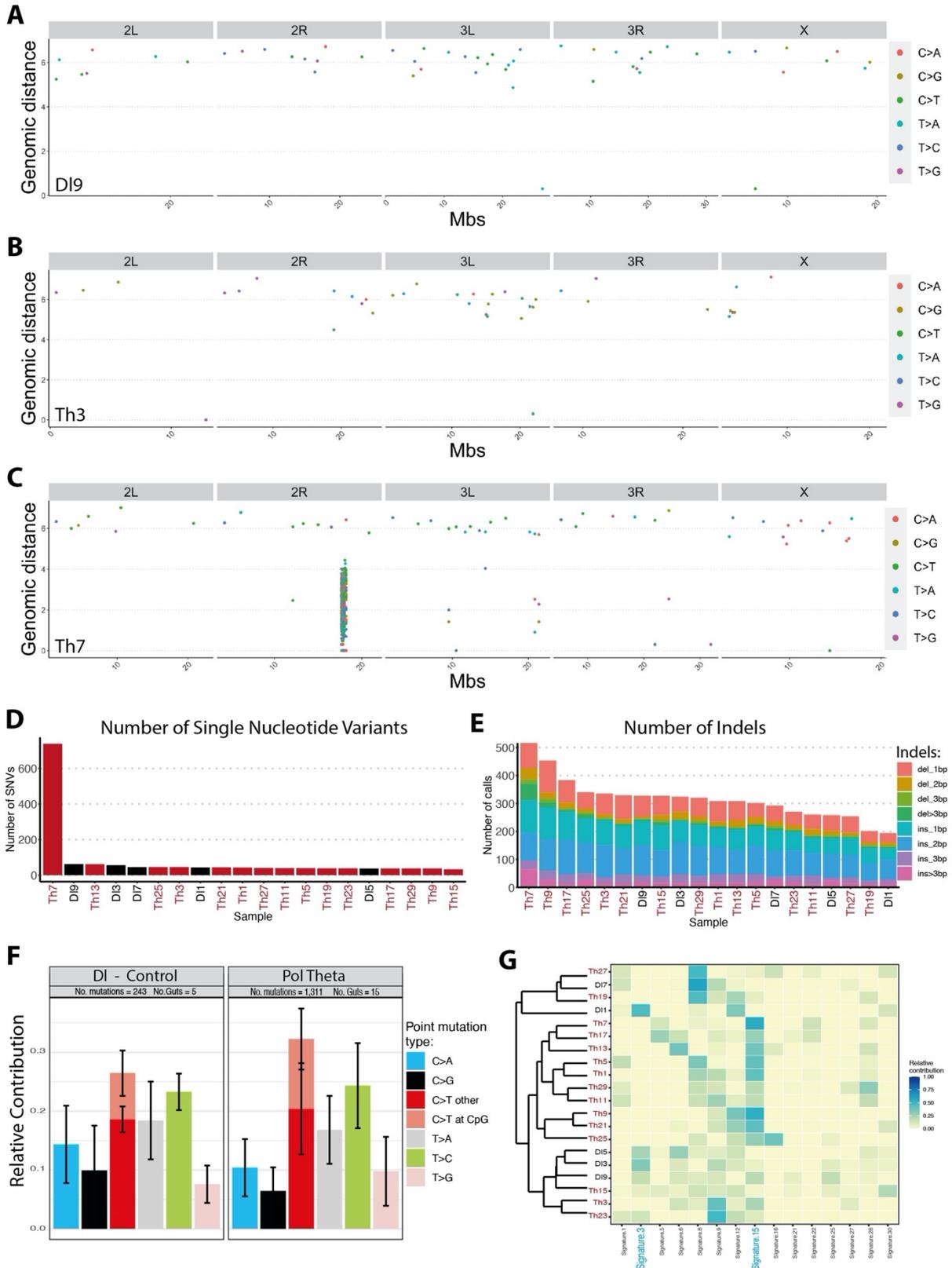
Thus, we found differences in the mutational profile of SV in *Polθ* mutants flies, suggesting that *Polθ* contributes in part to structural variants formation in neoplasia arising from ISC in the drosophila intestine<sup>25</sup>. However, the number of mutant samples analyzed and the small number of detected structural variants per sample likely weakens the conclusions established here.

#### ***Polθ likely contributes to the specific point mutation signature of ISC neoplasia***

In our previous study, we described the genome-wide point mutation landscape of ISC neoplasia (Riddiford et al). Since *Polθ* is an error prone polymerase that can generate point mutations during DNA repair<sup>34</sup>, we investigated the role of *Polθ* in single nucleotide variants (SNVs) and indel accumulation in intestinal stem cell neoplasia. Less than 50 SNVs were detected genome-wide for most neoplastic samples, in agreement to our previous analysis (Riddiford et al.). We detected no obvious difference in SNV number per sample between the *Pol theta* mutants and the newly-sequenced control flies (Fig. 6A–B,D; Fig. 6 Supplementary 1). We also examined the frequency of indels and found between 200 and 400 indels per samples genome wide, with no difference between the mutants and the new control flies (Fig. 6E).

However, one mutant sample showed signs of local hypermutation: Th7 displayed more than 700 SNVs and 500 indels, most of which were found inside a 560kb region on chromosome 2R (672 SNVs and 292 indels, Fig. 6C–E). With

an average of 1720 somatic mutation per megabase in this region, it is 1230 times higher than the average mutation rate (1.4 per Mb) that we identified previously in neoplasia. This type of mutation hotspot is reminiscent of mutations accumulated around breakpoints during Break-Induced Replication (BIR), a synthesis



**Figure 6. *Polθ* likely contributes to the point mutation signature of ISC neoplasia.** **A-C.** Rainfall plots of single nucleotide variant in the genome from *Polθ* mutants (Th1-29) and non-mutants (DI1-9) neoplasia. The type of mutation, is color-coded. The genomic distance is a measure of the distance between the two closest single nucleotide variants. Most samples have a few SNVs dispersed in the genome (A-B). The sample Th7 (C) displays a mutational pile-up in chromosome 2R. **D.** Single nucleotide variant number per sample in the genome of *Polθ* mutants (Th1-29) and non-mutants (DI1-9) neoplasia. **E.** Indel number per sample in the genome of *Polθ* mutants (Th1-29) and non-mutants (DI1-9) neoplasia. **F.** Relative contribution of each point mutation type to the mutational landscape of *Polθ* mutant and control neoplasia. **G.** Relative contribution of different COSMIC mutational signature to the mutational landscape of ISC neoplasia, this takes into account the tri-nucleotide context of point mutations to assign the most likely Signatures to explain the mutation profile for each sample.

dependent homology-based DNA double strand break repair mechanism with low fidelity of DNA synthesis. In BIR, the mutation rate has been shown to be 900 to 2800 times higher than spontaneous events across the genome<sup>35,36</sup>. It is, so far, the only sample in which we detected a localized accumulation of point mutations in intestinal stem cells and so we cannot conclude on a role for *Polθ* in limiting such events, however it is of note that we have never detected this in other wild-type samples.

The types of SNVs identified in samples did not vary between control and *Polθ* mutant samples (Fig. 6F). However, extending the analysis to consider the trinucleotide sequence context of the SNV allows for the detection of mutational signatures classified on the COSMIC database. These mutational signatures were generated from human cancer genomes datasets to represent specific mutational patterns related to particular DNA damage or repair mechanisms, or associated with specific tumor types. For each sample, we identified the mutational profiles and the relative contribution of each known COSMIC signatures to the profile (Fig. 6G). Our previous study had uncovered several signatures that contributed strongly in multiple samples (Riddiford et al.), among which, Signature 3 is related to failure of DNA double strand break repair by homologous recombination<sup>37</sup>. This signature strongly associated with homologous recombination deficient tumors in humans, such as in the context of germline or somatic BRCA1 and BRCA2 mutations. Here, in these new control samples (DI1, 3, 5, 9), we also found that Signature 3 contributed largely to the mutational profile but not to that in the *Polθ* mutant neoplasia. Of note, in cancers, homologous recombination-deficient tumors upregulate *Polθ* expression and depend on Alt-EJ for survival as it

promotes the recovery of stalled forks and repair of DNA double strand breaks in the absence of homologous repair<sup>27,28</sup>. In addition, in breast cancer<sup>37</sup> and in cell culture<sup>38</sup>, Signature 3 was also shown to be associated with the production of large deletions with overlapping microhomology at the breakpoints, a feature of *Polθ* dependent Alt-EJ. Since our data here suggest that loss of *Polθ* suppresses the signature 3 contribution to mutagenesis in ISC, *Polθ* may contribute to mutation signature 3.

Another mutational signature that stood out is Signature 15, since its contribution to the mutational profile seems higher in *Polθ* deficient neoplasia (Fig. 6G). Signature 15 is associated with defective DNA mismatch repair (MMR) in human cancer. Mismatch repair (MMR) proficiency is linked to *Polθ* inhibition<sup>39</sup> but how *Polθ* influences MMR has not been investigated, though synthetic lethality was found between *Polθ* and several MMR genes in mouse embryonic fibroblasts<sup>40</sup>. Our findings here suggest that *Polθ* might influence MMR efficiency in the intestinal stem cells. *Polθ* has also been involved in base excision repair in several species<sup>41-43</sup> but the mutational signatures associated with base excision repair (i.e. Signature 30) were not strongly represented in either group of samples.

To conclude, we identified several point mutation signatures that could be linked to *Polθ* activity in the ISCs, although the number of samples analyzed and the small amount of point mutations per sample invite for caution in the interpretation.

### III. DISCUSSION

Mutations arising in stem cells have important repercussions on adult tissues as they can drive tumor initiation, lead to lineage defects, or result in tissue functional decay during aging. Here, we clarify the timing and nature of DNA mutations arising spontaneously in the adult fly intestine.

Our analyses suggest that adult intestinal neoplasia are caused by inactivating mutations in *Notch* or other genes encoding Notch signaling components that arise in the adult stages and not during development. We cannot currently rule out that other mutations, not affecting the Notch pathway may occur during larval stages and contribute to genetic mosaicism in the adult gut. However, previous work of others showed that inactivation of *Notch* during larval development in adult stem cell precursors results in complete differentiation towards enteroendocrine cells, without formation of adult neoplasia<sup>10,11</sup>. Our analysis here at different timepoints during pupal development demonstrate that the inactivation of *Notch* in late pupal stages leads to medium-sized neoplasia in young adults. However, we never detect neoplastic growths spontaneously in young guts.

We can envisage two potential explanations as to why we do not detect neoplasia in young adults arising from mutations in pupal stages; either mutant cells are effectively eliminated or there are very low levels of DNA damage at this stage. Regarding cell elimination, it is possible that mutations occurring in pupal stages (but not in young adults) are culled by cell competition or through cell-intrinsic responses to DNA damage. Consistent with this hypothesis, recent findings upon X-ray induced DNA damage support the notion that cells undergoing loss of heterozygosity are eliminated during pupal stages through a mechanism independent of pro-apoptotic gene expression levels but dependent on *Serpent*<sup>44</sup>. As *Serpent* is involved in hemocyte development, the authors suggested that hemocytes might remove mutant cells. In this study however, spontaneous mutations were detected at very low frequencies and their mechanisms of removal were not assessed. Therefore, whether a similar process could apply to spontaneous mutations having much lower DNA damage than X-ray induced mutations is not currently clear. A second possible reason we do not detect neoplasia in young adults arising from mutations in pupal stages could be due to low DNA damage in this developmental stage. Alternatively, DNA repair may be highly efficient. It is possible that in adults, changes to the microbiome, food

consumption, or other physiological changes associated with adult life might increase DNA damage. Consistent with this, we here and others<sup>12,13</sup> have observed that marks of DNA damage increase in ISCs during aging. However, it should be noted that while the number of neoplasia in male flies increases during aging, this may be simply a linear increase and accumulation during aging and not necessarily an increased rate of mutation. In addition, we found no evidence of age-related changes in cell cycle arrest or DNA repair kinetics, which were similar in young and old ISCs.

Our analyses here and also published in Riddiford et al bioRxiv; Siudeja, et al, 2021 of a large cohort of spontaneously occurring gut tumors demonstrated that a multitude of genetic changes occur in ISCs during aging including large deletions, complex variants, somatic transposable element insertion, mitotic recombination and chromosome loss and that these can have dramatic effects on the genome. Ongoing studies in our lab aim to understand better these processes, how they are influenced by the environment, and to what extent other tissues are impacted by similar somatic genomic alterations. Here, we further provide evidence that erroneous repair mechanisms such as Pol $\theta$ -dependent MMEJ are likely involved in the generation of deletions in ISCs.

Through whole-genome sequencing, we find evidence for novel genomic signatures associated with the inactivation of *Pol $\theta$* . While the number of structural variants that we analyzed in *Pol $\theta$*  mutants is only 47, our data suggest that shorter inserts of DNA are found at breakpoints of structural variants than in our wild-type cohort of samples. This is consistent with recent finding of others in the *Drosophila* male germline<sup>25,26,45</sup>, in hematopoietic cells in contexts of clonal hematopoiesis<sup>46</sup>, and in cancer genomes<sup>47,48</sup>. Interestingly, while analyzing SNVs, we only observed a mutational pile-up in one *Pol $\theta$*  sample (out of 15). While this is only 1 sample, we have never observed a mutational pile-up in any of our previous wild-type samples (n= $\sim$ 45 neoplasia), suggesting it could be due to the inactivation of *Pol $\theta$* . Mutational pile-ups like these are associated with BIR and this could therefore implicate Pol $\theta$  in preventing BIR, which has not previously been suggested.

Consistent with this possibility, a recent study showed that *Polθ* blocks mitotic recombination in the male germline<sup>45</sup>. The authors assumed that the mitotic recombination resulted from cross-over of homologous chromosomes. However, we believe an alternative model is the mitotic recombination instead occurred through BIR, whereby the homologous chromosome is copied and not exchanged. Further studies will be necessary to distinguish between these models. Our study highlights the utility of the *Drosophila* midgut in understanding spontaneously arising stem cell mutations.

#### IV. MATERIALS AND METHODS

##### Fly Stocks

The experiments presented in this manuscript used the following fly lines.

From the Bloomington stock center: *w[\*]*; *P{w[+mC]=UAS-2xEGFP}AH2* (BL6874); *w[1118]*; *P{w[+mC]=UAS-GFP.nls}8* (BL4776);

The following stocks were gifts: *w<sup>1118</sup>*, *mus308<sup>2003</sup>* and *mus308<sup>Δ</sup>* (gift from M. McVey)<sup>25,32</sup>; *FUCCI* (*w1118*; *QUAS-GFP-E2F1<sub>1-230</sub><sup>#2</sup>*, *QUAS-mRFP1-NLS-CycB<sub>1-266</sub><sup>#3B</sup>/CyO* *wg-lacZ*; *MKRS/TM6B*, B. Edgar)<sup>49</sup>; *DeltaGal4/TM6TbHu* (S. Hou); *w N[55e11] P[neo, ry, FRT]19A/FM7 GFP* (F. Schweisguth); *y[1]*, *w[1118]*, *P{ry[+] neo FRT}19A* and *p{ry[+] neoFR}19A*, *P{w[+] tubP-GAL80}L1*, *P{ry[+] hsFLP}1*, *w*; *CyO/P{w[+] UAS-CD8:GFP}LL5*; *TM6, Tb, Hu / P{w[+] tubP-GAL4}LL7* (MARCM19A, A.Gould)

The *Delta<sup>ts</sup>* line was generated by recombination of the *DeltaGal4/TM6TbHu* (S. Hou), and the *tub-Gal80<sup>ts</sup>* from another line. It was used to generate: *FUCCI;Dl<sup>ts</sup>/TM6TbHu*

The *RFP-Polθ* flies were generated via CRISPR insertion to produce a protein with RFP fused at the N-terminal end of *Polθ* by *inDroso@ functional genomics* (indroso.com).

##### MARCM clone induction

For the clone induction in pupae, white newly formed pupae were selected every hour and kept at 25°C to insure a correct developmental timing. They were heat-shocked at 37°C for 45mins 24 or 72 hours after pupal formation. The guts were dissected 6 days after adult eclosion.

### *Aging and survival analysis*

For the X ray experiment, flies were aged at 25°C in tubes with 20–25 females + 5 males per tubes, and flipped to new tubes with fresh yeasted food every 3–4 days.

For neoplasia experiments, flies were aged as previously described<sup>5</sup>. Briefly, the flies were crossed in standard tubes (15 females + 3–4 males per tube), the newly eclosed progeny were collected for 2–3 days and aged together in plastic cages (10cm diameter, 942mL) with ~300 males and 300 females per cages. Fresh yeasted food was provided every 2–3 days and cages were changed every week. Dead flies were scored in the food plates at every food change to measure the survival rate presented in Figure 4.

### *Xray irradiation*

Flies were irradiated 8 mins and 45 secs in their tube with a X-rays generator Philips 320 kV. With 13.1 mA, 320kV for a total dose of 10Gy (+/- 15%).

### *Immunostaining*

The tissues were fixed and stained following a previously published protocol<sup>50</sup>. Briefly, the flies were dissected in PBS and the tissues were fixed at room temperature in 4% paraformaldehyde for 2hrs (guts). They were then washed in PBS+ 0.1% Triton X-100 (PBT). To remove the content of the gut lumen, the fixed gut were equilibrated 30mins in PBS 50% Glycerol and then washed in PBT. The primary antibody staining was performed overnight at 4°C. After washing with PBT 20 mins 3 times, the secondary antibody was incubated 3–4hrs at room

temperature. After washing with PBT, DAPI staining was done in PBT for 30 mins (1 $\mu$ g/ml).

The antibodies used were: Chicken anti-GFP (1:2000, Invitrogen #A10262), Rabbit anti-RFP (1:1000, Clontech #632496), Rabbit anti- $\gamma$ H2Av (1:2000, Rockland 600-401-914), Mouse anti- $\gamma$ H2Av (1:200, DSHB UNC93-5.2.1), mouse anti-Delta (1:1000, DSHB C594.9B), mouse anti-Pros (1:2000, DSHB MR1A-c), Goat anti-Ph3 (ser28, 1:200, Santa Cruz sc-12927 lot#C0172).

The secondary antibodies were all from Jackson Laboratory, raised in Donkey with different fluorophores (DyLight 488, 549, 649)

### **Microscopy and Image Analysis**

Pictures showed here were taken using the Zeiss confocal microscope LSM700, LSM800 or LSM900 with 40x oil objective and 63x objective for RFP-Pol $\theta$ . Quantifications for neoplasia detection and PH3 counts were done with a Leica Epifluorescence microscope.

The images were processed and the quantification performed using Fiji or Imaris (X-ray experiment), with manual detection of cells or automatic detection of cells with a macro that we designed.

The quantification for the Fucci experiments and X-ray experiment were performed blind.

### **Data Processing and Statistics**

Data Analysis was performed on Microsoft Excel, Prism (version 9.0.1) or R (version 4.0.2) using the Rstudio interface (version 1.3.1073) and ggplot2 (version 3.3.3), survminer (version 0.4.8) and survival (version 3.2-7) packages. ggplot2 for data visualization; survminer and survival for plotting and statistics of survival curves. Initial training for the use of the ggplot2 package for data visualization was obtained by the U900 Bioinformatics unit of the Institute.

### **Sequencing of *Drosophila* neoplasia and controls**

A more detailed methodology can be found in (Siudeja et al. 2021). In brief, for selection of neoplastic tissue, clustered ISC cells were selected by expression of DI-Gal4 driven UAS-nlsGFP or UAS-2xGFP. 5–7 weeks-old DI>GFP males or females were used to visually identify midguts containing neoplasia based on clonal accumulation of GFP positive cells. Neoplasia were manually dissected together with the head as a direct comparison. For the DI>nlsGFP flies (Fig. 3, supp Fig.3), genomic DNA was isolated with the QIAamp DNA MicroKit (Qiagen) according to the manufacturer's protocol.

For the Pol $\theta$  mutants and control neoplasia of DI>2XGFP (Fig. 5–6), genomic DNA was extracted with the Zymo Research Quick-DNA™ Microprep Plus Kit (D4074), following the instruction manual

DNA concentration was measured with the Qubit dsDNA High Sensitivity Assay Kit.

Because of the small size of the dissected tissue and our inability to completely isolate the neoplasia from the surrounding non-neoplastic midgut tissue, we estimate that the dissected neoplasia contained 40%–80% neoplastic cells (GFP+), which represents more than 80% of DNA from neoplastic cells (see Riddiford et al).

### ***DNA sequencing***

Genomic DNA libraries were prepared with the Nextera XT protocol (Illumina)

For the Pol $\theta$  experiment, the libraries were prepared with Roche's KAPA Hyper Plus Library, using 1.25 ng of extracted DNA per sample.

Whole-genome 2 x 150 bps paired-end sequencing was performed on HiSeq2500 or NovaSeq (Illumina SP – PE150).

### ***Bioinformatics***

The Bioinformatic analysis was performed entirely by N. Riddiford, a detail of the bioinformatics tools and pipeline can be found in Riddiford et al. and

<https://github.com/nriddiford>

One sample (DI3) was excluded from the structural variant analysis because of low read-depth.

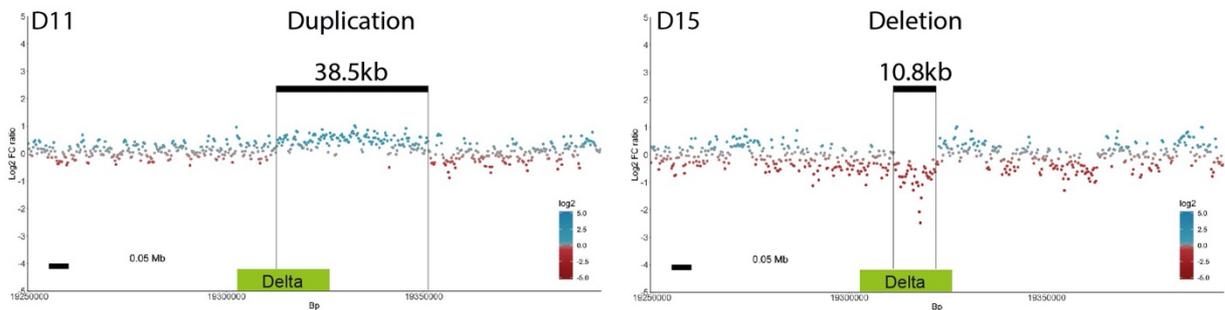
The IGV software was used to visualize the reads from the sequencing and validate the detection of mutations.

### ***Author contributions***

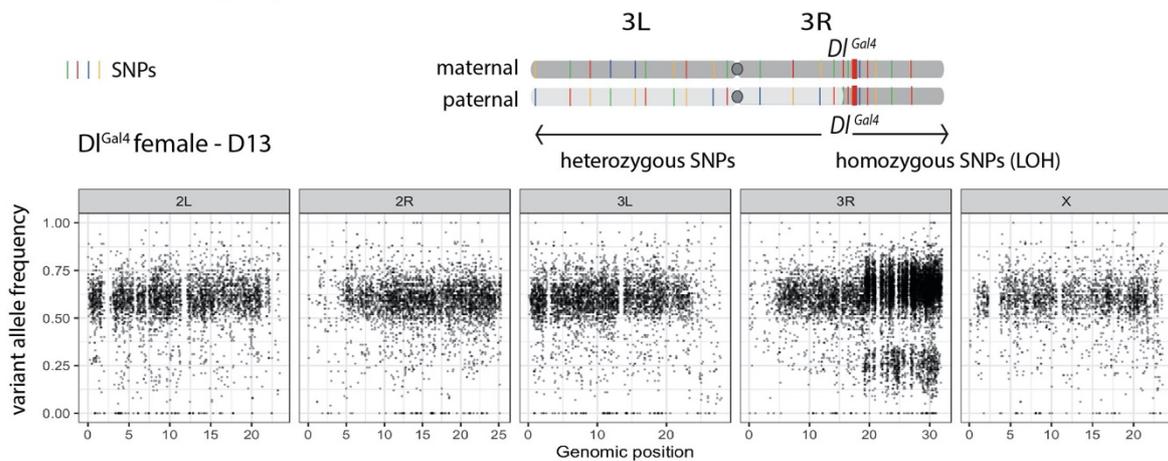
**BB** designed the study, performed all the experiments and quantification, and wrote the manuscript. **NR** conducted the bioinformatic analysis. **AB** designed the study and wrote the manuscript.

## V. SUPPLEMENTARY FIGURES

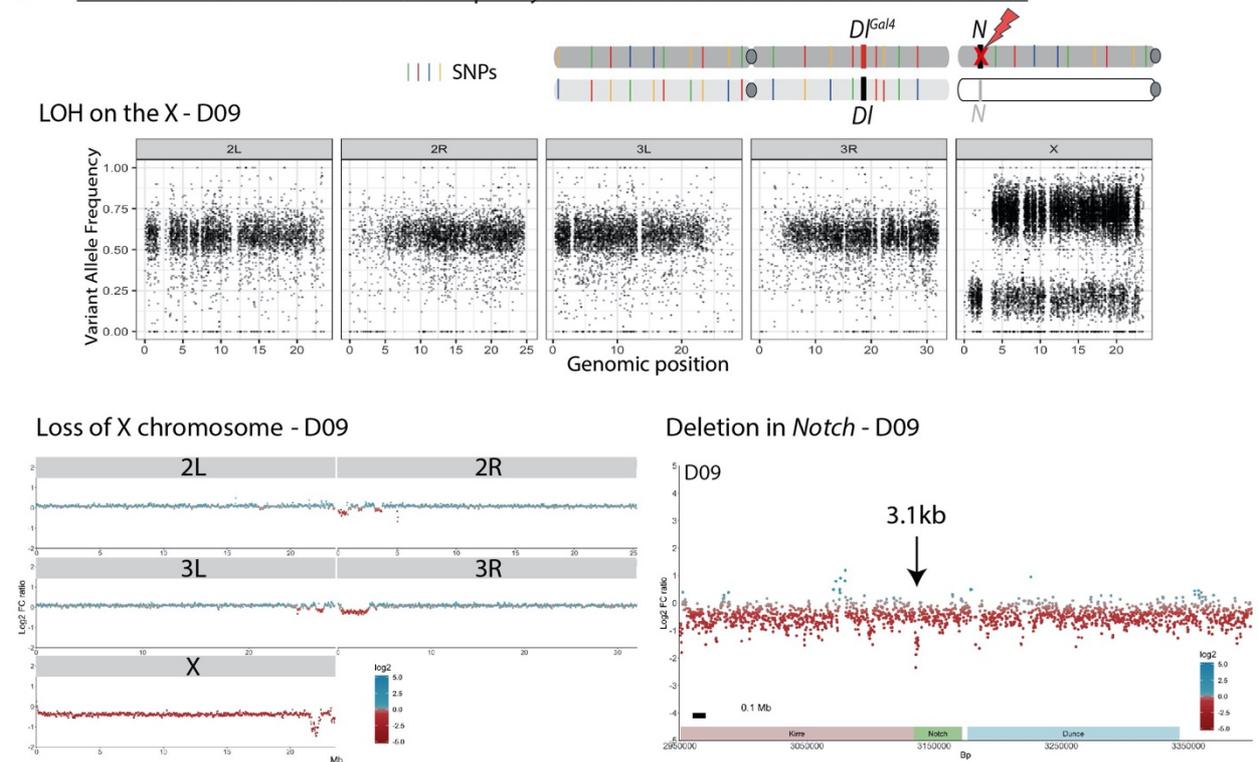
### A Structural variants in Delta



### B Loss of Heterozygosity via mitotic recombination on chromosome 3R

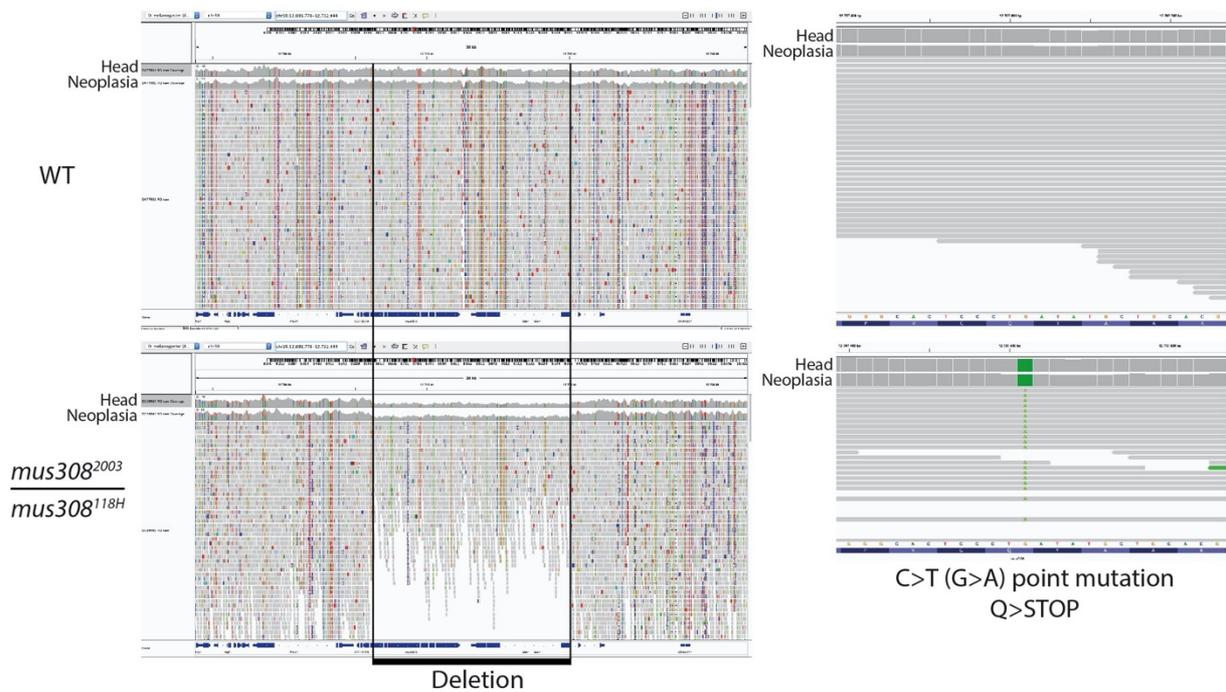


### C Bi-Allelic inactivation of Notch: Aneuploidy of the X chromosome and Deletion in Notch



**Figure 3 Supplementary 1. Neoplasia growth in female is driven by different mutagenic events.** **A.** Structural variants in *Delta* as seen by Log2 ratio changes in read copy-number. **B.** Loss of Heterozygosity of *Delta* via mitotic recombination seen by the shift in parental variant allele frequency from homozygous to heterozygous. **C.** Bi-Allelic inactivation of *Notch*: the shift to homozygosity in parental allele frequency in the X-chromosome is associated with a drop in the read copy-number ratio, indicative of loss of the X. And a deletion in *Notch* responsible for the second allele inactivation.

Trans-heterozygotes *mus308* mutants

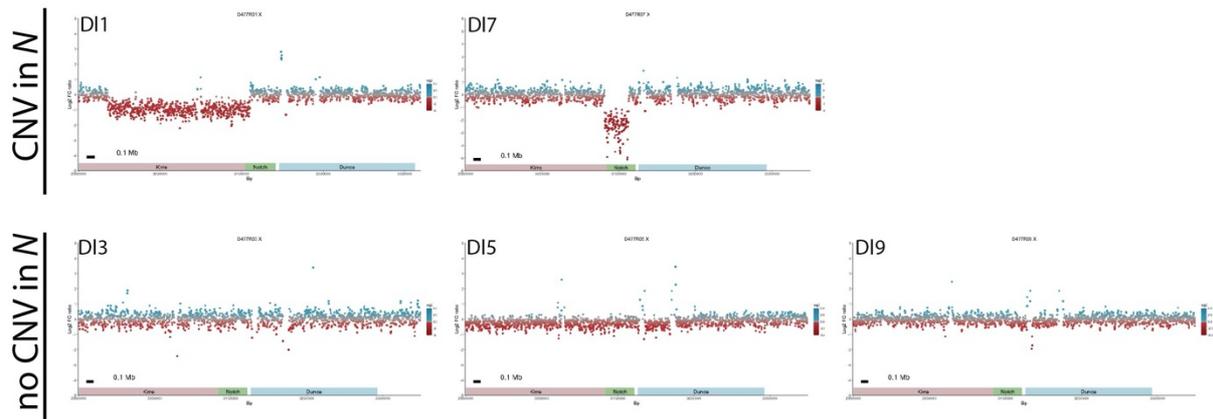


**Figure 4 Supplementary 1. *Polθ* mutants validation.** The sequenced transheterozygotes mutants have a 14kb deletion spanning *Polθ* and *Men* (*mus308*<sup>118H</sup>), and C>T point mutation (green) (*mus308*<sup>2003</sup>) in both the head and the neoplasia. The point mutation appears homozygote in the sequencing since it is in trans with the deletion.

SAMPLE	TYPE	LENGTH (KB)	POSITION	ALLELE FREQ.	MICROHOMOLOGY	INSERTION
P1	TRA	15162.5	2L:16310775 4:1148246	0.1	CGGTCAC	-
P11	TRA	3690.5	3R:14990196 2R:18680738	0.04	AAGGACATTGA	-
P13	TRA	8571.1	2L:3093585 3L:11664665	0.29	ACCAAAT	-
P13	TRA	2930.8	2L:4138262 3L:7069104	0.29	GCA	-
P13	TRA	8102.6	2L:6084142 2R:14186700	0.19	CG	-
P13	TRA	7471.9	2L:14068330 3L:21540255	0.21	-	GTAGATG
P13	TRA	3426	2L:14395906 3L:10969893	0.22	TTAGA	-
P19	TRA	10587.7	2L:3115500 3R:17303238	0.07	-	TTAC
P19	TRA	7341	2L:5313800 X:12654753	0.06	GGAAAAGGAAGGC	-
P19	TRA	6074.7	2L:6110188 Tome:35464	0.11	AT	-
P19	TRA	19771.1	2L:7747929 3R:27519078	0.06	GTTC	-
P19	TRA	8696.4	2L:10991843 X:19688243	0.05	-	TGCGCCTAAGTGT
P33	TRA	21801	2L:21808001 Tomelloso:7008	0.09	G	-
P33	TRA	29160.5	3L:120417 3R:29280935	0.08	-	G
P33	TRA	24626.4	Tome:20114 2R:24646497	0.1	AC	-
P5	TRA	3464.4	3L:9960263 2R:13424638	0.12	ATT	-
P63	TRA	718	2L:7888676 2R:7170648	0.04	GCCAGC	-
P63	TRA	4114.7	2R:16168568 3L:20283281	0.04	GCCA	-
P63	TRA	7127	2R:23886797 3R:31013782	0.04	GGCAC	-
P65	TRA	16004.7	3L:3775419 X:19780138	0.1	ACGA	-
P7	TRA	301.4	2L:603468 3L:302053	0.12	T	-
P7	TRA	4723.6	2L:1621944 3R:6345586	0.1	TGCCGC	-
P7	TRA	14894.8	2L:2208673 3L:17103423	0.08	CATCT	-
P7	TRA	22155.1	2L:2716106 2R:24871228	0.14	TAAGCTGCTGT	-
P7	TRA	20846.4	2L:2837516 3R:23683953	0.15	GCG	-
P7	TRA	1228.1	2L:3225319 X:4453399	0.19	CC	-
P7	TRA	103	2L:3399658 3L:3296613	0.09	TTGACTGGC	-
P7	TRA	18408.7	2L:4687951 3L:23096651	0.13	AGCA	-
P7	TRA	22452.9	2L:5077640 3R:27530506	0.13	-	CTGCTTGTG
P7	TRA	11594.3	2L:6469039 X:18063310	0.14	CGGGC	-
P7	TRA	12666.8	2L:6567403 3R:19234202	0.11	GTG	-
P7	TRA	2493.5	2L:7096540 3L:4603018	0.12	TGCGAAC	-
P7	TRA	9548.8	2L:7458082 3R:17006857	0.14	CTGAA	-
P7	TRA	349.9	2L:7807464 2R:8157412	0.1	GGCAG	-
P7	TRA	11511.7	2L:8839310 3L:20350987	0.09	-	GATAG, CTGTC
P7	TRA	6157.4	2L:9252716 2R:15410147	0.09	CGAGCAGCTG	-
P7	TRA	3018.7	2L:11635382 X:14654037	0.14	-	CTGAACAA
P7	TRA	3861	2L:11956048 X:15817056	0.12	GAACGA	-
P7	TRA	1051.6	2L:13724341 X:14775902	0.25	-	GACAG, TGTT
P7	TRA	12364.3	2L:14178344 X:1814038	0.14	ACGAAAC	-
P7	TRA	10930.7	2L:17692544 3L:6761859	0.16	T	-
P7	TRA	1827	2L:18949106 2R:17122080	0.15	-	ACCCTAAC
P7	TRA	4310.5	2L:19515803 2R:15205313	0.11	CGT	-
P7	TRA	7408	2L:19792664 3L:12384685	0.17	T	AGACAG, CTGTCT
P7	TRA	3705.6	2L:20424190 2R:24129750	0.11	GCG	-
P7	TRA	15256.1	2L:20495067 3L:5238975	0.22	A	-
P7	TRA	6565.2	2L:22041503 X:15476281	0.1	TGATGG	-
P7	TRA	325.4	2L:22997150 X:22671759	0.19	TCTAC	-
P7	TRA	1077.5	2R:3524816 X:4602283	0.22	CTGG	-
P7	TRA	8431.2	2R:7317301 X:15748502	0.11	-	GGAAGTAAAAGTT
P7	TRA	866.1	2R:8397888 3L:7531805	0.09	-	T
P9	TRA	10947.9	2L:2565270 2R:13513152	0.06	T	-
P9	TRA	14666.5	2L:3829232 X:18495746	0.08	GGGATTTTC	-
P9	TRA	6768.1	2L:5445521 3R:12213592	0.06	CAAATTA	-
P9	TRA	13847.4	2L:6096637 2R:19944064	0.05	C	-
P9	TRA	14751.1	2R:6532200 2L:21283344	0.04	T	-
P9	TRA	2991.3	2R:9334321 3L:12325601	0.07	CCAAAAC	-
P9	TRA	790.7	2R:10667716 3L:9876984	0.21	CAA	-
P9	TRA	4171.4	2R:12455384 3L:8284008	0.05	G	-
P9	TRA	7851.9	2R:21549103 3R:13697225	0.05	G	-
P9	TRA	1776.8	2R:23141212 3R:21364381	0.04	CCCTT	-
P9	TRA	16706.9	2R:23480790 3R:6773848	0.05	GTAG	-
P9	TRA	8838.6	3L:4762617 2L:13601230	0.06	C	-
P9	TRA	2149.8	3L:7116471 2R:4966685	0.27	A	-
P9	TRA	1918.9	3L:12130274 3R:10211411	0.06	GTGCCA	-
P9	TRA	14834.6	3L:16590696 3R:31425288	0.05	C	-
P9	TRA	18449.3	3L:22127698 X:3678445	0.06	TTACCCC	-
P9	TRA	13250.6	3R:11181613 2R:24432165	0.05	AG	-
P9	TRA	12878.3	3R:25330986 X:12452678	0.05	TCAATTTGTAGG	-
P9	TRA	12619.4	4:418098 2L:13037524	0.04	AGACA	-
P9	TRA	8946.8	2R:23186121 3L:14239313	0.04	CTTGGCCAT	-

**Supplementary Table 2. Structural Variants with Microhomologies and/or insertions at the breakpoints of translocations in male neoplasias.**

**WT**



**Pol Theta mutants**

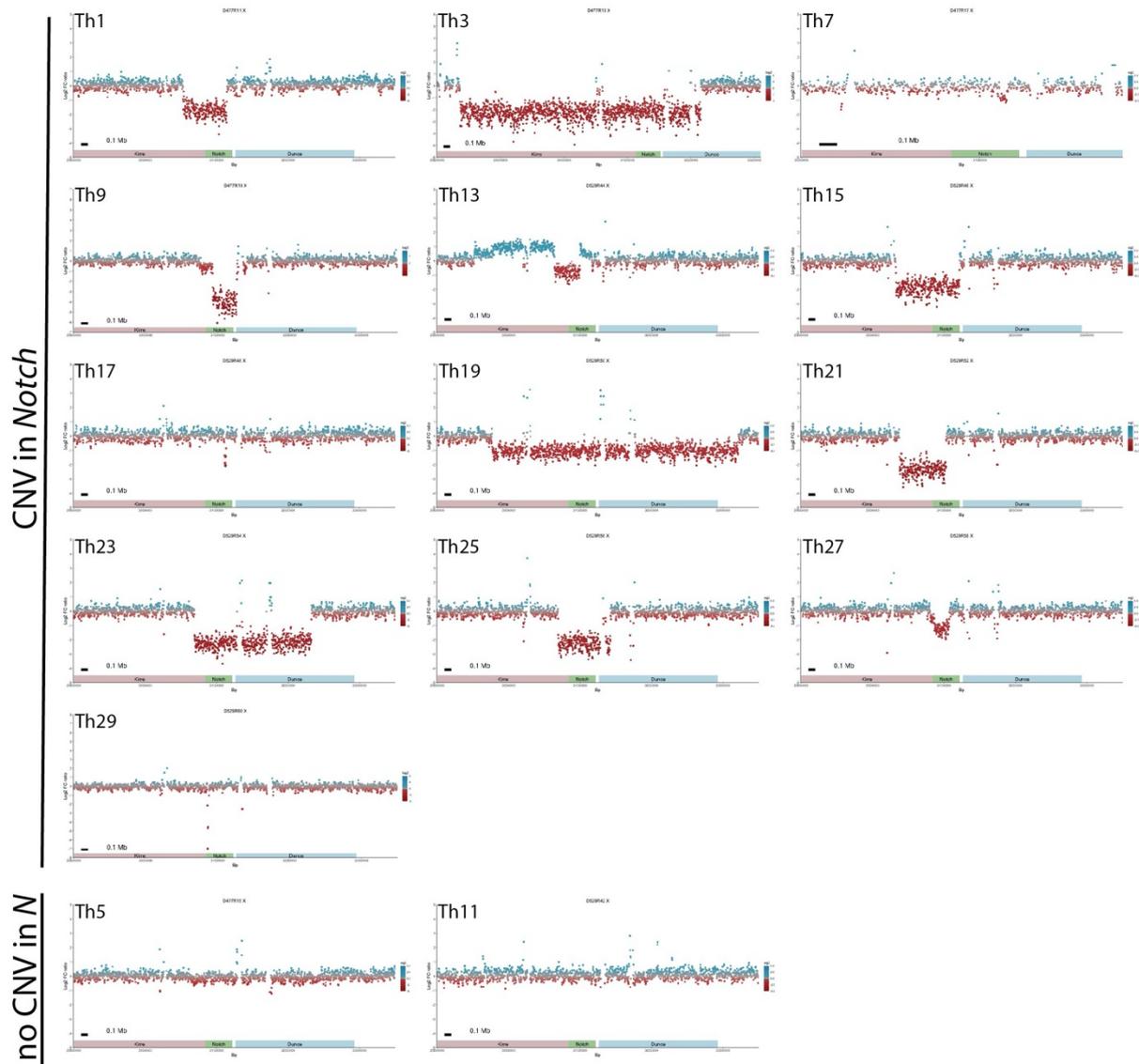
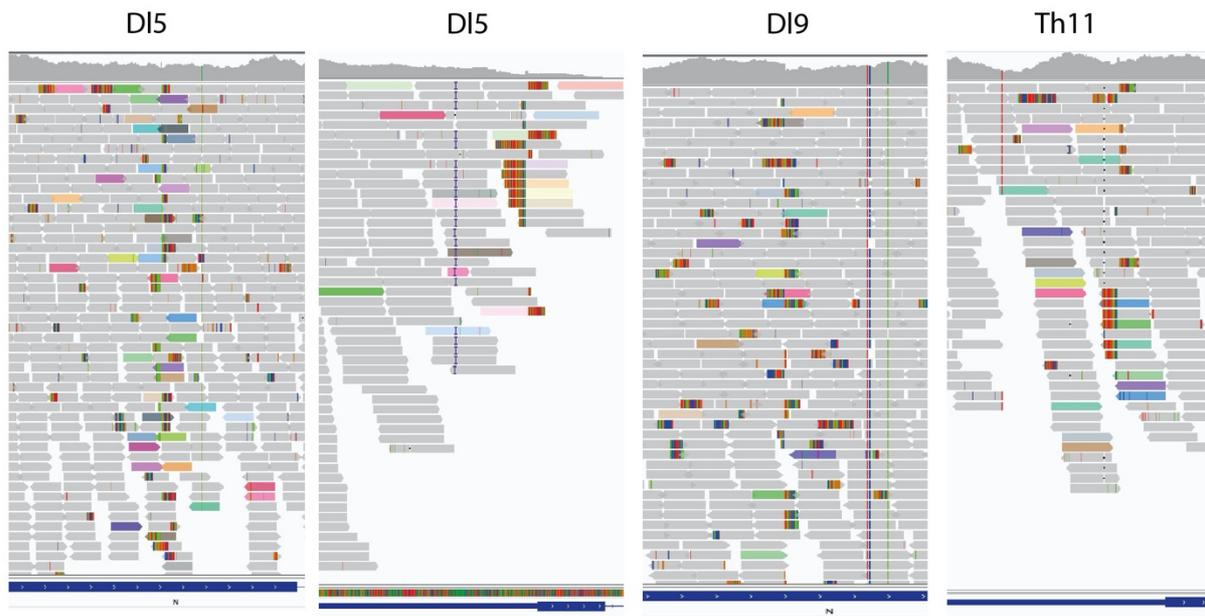
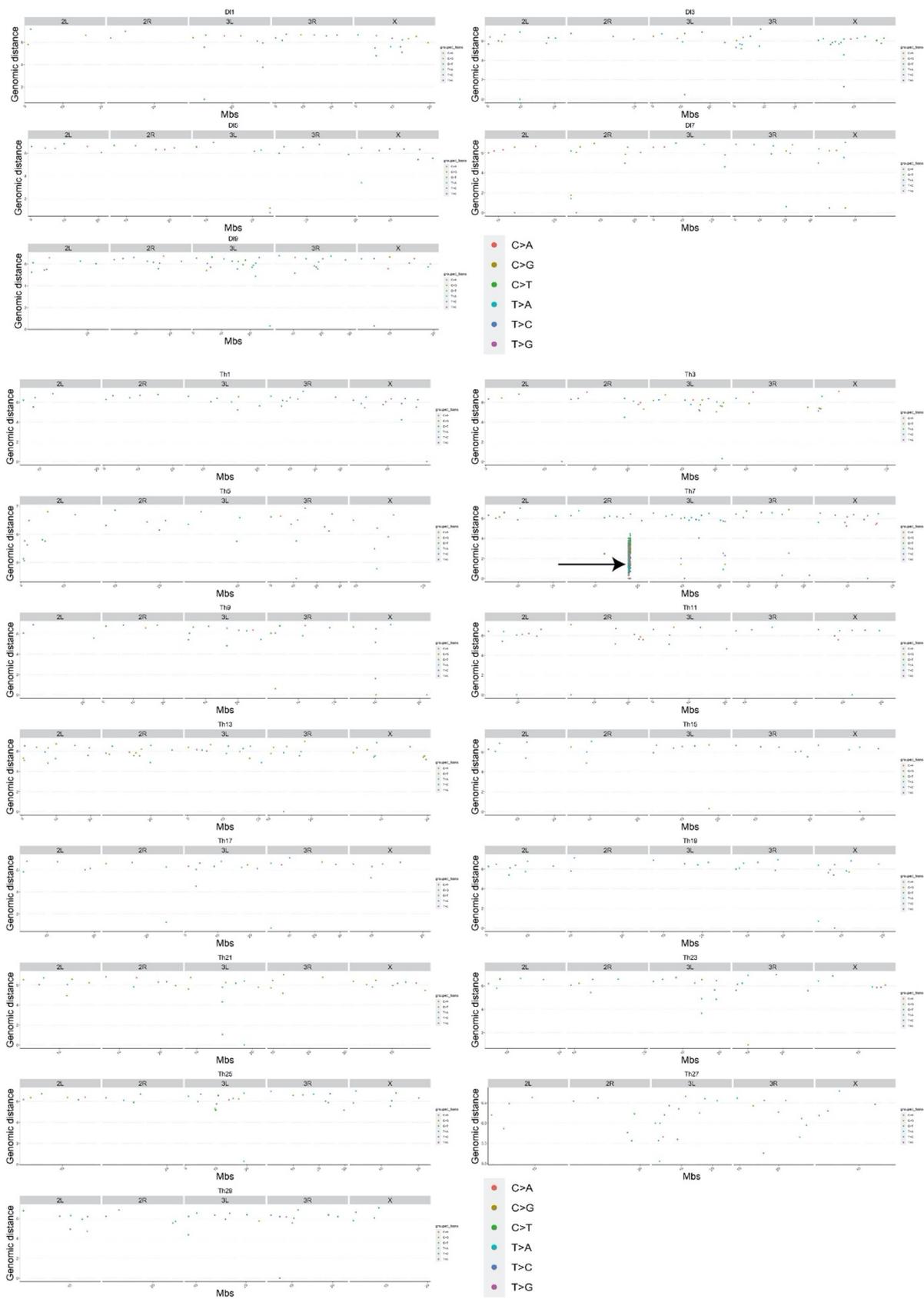


Figure 5 Supplementary 1. Copy-number variant detection at Notch in *Polθ* mutants and isogenized non-mutant neoplasia.

## Transposable elements insertions in *Notch*



**Figure 5 Supplementary 2. Transposable element insertion detection at *Notch* in *Polθ* mutants and isogenized non-mutant neoplasia.** IGV screenshots from the sequencing reads alignment of the neoplasia. Reads with paired-read mapping to the transposon are in pastel colors



**Figure 6 Supplementary 1. Single nucleotide variant detection genome-wide in *Polθ* mutants and isogenized non-mutant neoplasia.**

## VI. REFERENCES:

1. Hanahan, D. & Weinberg, R. A. Hallmarks of Cancer: The Next Generation. *Cell* **144**, 646–674 (2011).
2. Al zouabi, L. & Bardin, A. J. Stem Cell DNA Damage and Genome Mutation in the Context of Aging and Cancer Initiation. *Cold Spring Harb. Perspect. Biol.* a036210 (2020) doi:10.1101/cshperspect.a036210.
3. Micchelli, C. A. & Perrimon, N. Evidence that stem cells reside in the adult *Drosophila* midgut epithelium. *Nature* **439**, 475–479 (2006).
4. Ohlstein, B. & Spradling, A. The adult *Drosophila* posterior midgut is maintained by pluripotent stem cells. *Nature* **439**, 470–474 (2006).
5. Siudeja, K. *et al.* Frequent Somatic Mutation in Adult Intestinal Stem Cells Drives Neoplasia and Genetic Mosaicism during Aging. *Cell Stem Cell* **17**, 663–674 (2015).
6. Siudeja, K., Wurmser, A., Stefanutti, M., Lameiras, S. & Bardin, A. J. Unraveling the features of somatic transposition in the *Drosophila* intestine. *EMBO J.* 19 (2021).
7. Riddiford, N., Siudeja, K., van den Beek, M., Boumard, B. & Bardin, A. J. Evolution and genomic signatures of spontaneous somatic mutation in *Drosophila* intestinal stem cells. *BioRxiv* (2020) doi:10.1101/2020.07.20.188979.
8. Beagan, K. & McVey, M. Linking DNA polymerase theta structure and function in health and disease. *Cell. Mol. Life Sci.* **73**, 603–615 (2016).
9. Jiang, H. & Edgar, B. A. EGFR signaling regulates the proliferation of *Drosophila* adult midgut progenitors. *Development* **136**, 483–493 (2009).
10. Guo, Z. & Ohlstein, B. Bidirectional Notch signaling regulates *Drosophila* intestinal stem cell multipotency. *Science* **350**, aab0988–aab0988 (2015).
11. Takashima, S. *et al.* Development of the *Drosophila* entero–endocrine lineage and its specification by the Notch signaling pathway. *Dev. Biol.* **353**, 161–172 (2011).
12. Park, J.–S. *et al.* Age– and oxidative stress–induced DNA damage in *Drosophila* intestinal stem cells as marked by Gamma–H2AX. *Exp. Gerontol.* **47**, 401–405 (2012).
13. Sousa–Victor, P. *et al.* Piwi Is Required to Limit Exhaustion of Aging Somatic Stem Cells. *Cell Rep.* **20**, 2527–2537 (2017).
14. Biteau, B., Hochmuth, C. E. & Jasper, H. JNK Activity in Somatic Stem Cells Causes Loss of Tissue Homeostasis in the Aging *Drosophila* Gut. *Cell Stem Cell* **3**, 442–455 (2008).
15. Flach, J. *et al.* Replication stress is a potent driver of functional decline in ageing haematopoietic stem cells. *Nature* **512**, 198–202 (2014).
16. Delabaere, L. *et al.* Aging impairs double–strand break repair by homologous recombination in *Drosophila* germ cells. *Aging Cell* (2016) doi:10.1111/accel.12556.
17. Pyo, J.–H. *et al.* Functional Modification of *Drosophila* Intestinal Stem Cells by Ionizing Radiation. *Radiat. Res.* **181**, 376–386 (2014).

18. Buchon, N. *et al.* Morphological and Molecular Characterization of Adult Midgut Compartmentalization in *Drosophila*. *Cell Rep.* **3**, 1725–1738 (2013).
19. Dutta, D. *et al.* Regional Cell-Specific Transcriptome Mapping Reveals Regulatory Complexity in the Adult *Drosophila* Midgut. *Cell Rep.* **12**, 346–358 (2015).
20. Siudeja, K. & Bardin, A. J. Somatic recombination in adult tissues: What is there to learn? *Fly (Austin)* 1–8 (2016) doi:10.1080/19336934.2016.1249073.
21. Rodgers, K. & McVey, M. Error-Prone Repair of DNA Double-Strand Breaks: ERROR-PRONE REPAIR OF DNA DSBs. *J. Cell. Physiol.* **231**, 15–24 (2016).
22. Lee, J. A., Carvalho, C. M. B. & Lupski, J. R. A DNA Replication Mechanism for Generating Nonrecurrent Rearrangements Associated with Genomic Disorders. *Cell* **131**, 1235–1247 (2007).
23. Hastings, P. J., Ira, G. & Lupski, J. R. A Microhomology-Mediated Break-Induced Replication Model for the Origin of Human Copy Number Variation. *PLoS Genet.* **5**, e1000327 (2009).
24. Tubbs, A. *et al.* Dual Roles of Poly(dA:dT) Tracts in Replication Initiation and Fork Collapse. *Cell* **174**, 1127–1142.e19 (2018).
25. Chan, S. H., Yu, A. M. & McVey, M. Dual Roles for DNA Polymerase Theta in Alternative End-Joining Repair of Double-Strand Breaks in *Drosophila*. *PLoS Genet.* **6**, e1001005 (2010).
26. Yu, A. M. & McVey, M. Synthesis-dependent microhomology-mediated end joining accounts for multiple types of repair junctions. *Nucleic Acids Res.* **38**, 5706–5717 (2010).
27. Ceccaldi, R. *et al.* Homologous-recombination-deficient tumours are dependent on Pol $\theta$ -mediated repair. *Nature* **518**, 258–262 (2015).
28. Kais, Z. *et al.* FANCD2 Maintains Fork Stability in BRCA1/2-Deficient Tumors and Promotes Alternative End-Joining DNA Repair. *Cell Rep.* **15**, 2488–2499 (2016).
29. Boulon, S., Westman, B. J., Hutten, S., Boisvert, F.-M. & Lamond, A. I. The Nucleolus under Stress. *Mol. Cell* **40**, 216–227 (2010).
30. Audas, T. E., Jacob, M. D. & Lee, S. Immobilization of proteins in the nucleolus by ribosomal intergenic spacer noncoding RNA. *Mol. Cell* **45**, 147–157 (2012).
31. Boyd, J. B., Sakaguchi, K. & Harris, P. V. mus308 mutants of *Drosophila* exhibit hypersensitivity to DNA cross-linking agents and are defective in a deoxyribonuclease. *Genetics* **125**, 813–819 (1990).
32. Beagan, K. *et al.* *Drosophila* DNA polymerase theta utilizes both helicase-like and polymerase domains during microhomology-mediated end joining and interstrand crosslink repair. *PLOS Genet.* **13**, e1006813 (2017).
33. Yousefzadeh, M. J. *et al.* Mechanism of Suppression of Chromosomal Instability by DNA Polymerase POLQ. *PLoS Genet.* **10**, e1004654 (2014).
34. Arana, M. E., Seki, M., Wood, R. D., Rogozin, I. B. & Kunkel, T. A. Low-fidelity DNA synthesis by human DNA polymerase theta. *Nucleic Acids Res.* **36**,

- 3847–3856 (2008).
35. Sakofsky, C. J. *et al.* Break-Induced Replication Is a Source of Mutation Clusters Underlying Kataegis. *Cell Rep.* **7**, 1640–1648 (2014).
  36. Deem, A. *et al.* Break-Induced Replication Is Highly Inaccurate. *PLOS Biol.* **9**, e1000594 (2011).
  37. Nik-Zainal, S. *et al.* Mutational Processes Molding the Genomes of 21 Breast Cancers. *Cell* **149**, 979–993 (2012).
  38. Zámboreszky, J. *et al.* Loss of BRCA1 or BRCA2 markedly increases the rate of base substitution mutagenesis and has distinct effects on genomic deletions. *Oncogene* **36**, 746–755 (2017).
  39. Fujimori, H. *et al.* Mismatch repair dependence of replication stress-associated DSB recognition and repair. *Heliyon* **5**, (2019).
  40. Feng, W. *et al.* Genetic determinants of cellular addiction to DNA polymerase theta. *Nat. Commun.* **10**, (2019).
  41. Yoshimura, M. *et al.* Vertebrate POLQ and POL $\beta$  Cooperate in Base Excision Repair of Oxidative DNA Damage. *Mol. Cell* **24**, 115–125 (2006).
  42. Prasad, R. *et al.* Human DNA polymerase  $\theta$  possesses 5'-dRP lyase activity and functions in single-nucleotide base excision repair in vitro. *Nucleic Acids Res.* **37**, 1868–1877 (2009).
  43. Asagoshi, K. *et al.* Single-nucleotide base excision repair DNA polymerase activity in *C. elegans* in the absence of DNA polymerase  $\beta$ . *Nucleic Acids Res.* **40**, 670–681 (2012).
  44. Brown, J., Bush, I., Bozon, J. & Su, T. T. Cells with loss-of-heterozygosity after exposure to ionizing radiation in *Drosophila* are culled by p53-dependent and p53-independent mechanisms. *PLoS Genet.* **16**, e1009056 (2020).
  45. Carvajal-Garcia, J., Crown, K. N., Ramsden, D. A. & Sekelsky, J. DNA polymerase theta suppresses mitotic crossing over. *PLOS Genet.* **17**, e1009267 (2021).
  46. Feldman, T. *et al.* Recurrent deletions in clonal hematopoiesis are driven by microhomology-mediated end joining. *Nat. Commun.* **12**, 2455 (2021).
  47. Hwang, T. *et al.* Defining the mutation signatures of DNA polymerase  $\theta$  in cancer genomes. *NAR Cancer* **2**, (2020).
  48. Kamp, J. A., van Schendel, R., Dilweg, I. W. & Tijsterman, M. BRCA1-associated structural variations are a consequence of polymerase theta-mediated end-joining. *Nat. Commun.* **11**, 3615 (2020).
  49. Zielke, N. *et al.* Fly-FUCCI: A Versatile Tool for Studying Cell Proliferation in Complex Tissues. *Cell Rep.* **7**, 588–598 (2014).
  50. Bardin, A. J., Perdigoto, C. N., Southall, T. D., Brand, A. H. & Schweisguth, F. Transcriptional control of stem cell maintenance in the *Drosophila* intestine. *Development* **137**, 705–714 (2010).

# CHAPTER 3 : A TISSUE-SPECIFIC BUFFERING MECHANISM OF REPLICATION STRESS

# Closing the GAP on replication stress: Tissue-specific buffering of nucleotide levels protects from replication stress

Benjamin Boumard<sup>1</sup>, Marwa El-Hajj<sup>1</sup>, Allison Bardin<sup>1</sup>

<sup>1</sup> Institut Curie, PSL Research University, CNRS UMR 3215, INSERM U934, Stem Cells and Tissue Homeostasis Group, Paris, France.

## I. INTRODUCTION

Tissue-specific stem cells are essential to maintain tissue homeostasis: they have the unique ability to self-renew and differentiate into the different cell types of the tissue and their regenerative potential is tightly regulated to adapt to the organ needs. Altered stem cell function, for example as a consequence of genome instability<sup>1</sup>, can lead to stem cell loss or stem cell hyperproliferation and tumorigenesis. Thus, genome integrity in long-lived tissue stem cells is essential to maintain tissue function and prevent cancer initiation. In parallel, an age-related increase in tissue specific somatic mutations was observed in various organs, linking the stem cell mutational landscape with tissue specific tumorigenesis<sup>2</sup>. How different stem cells cope with DNA lesions determines their mutation rate, susceptibility to cancer, and likely age-related functional decline<sup>2</sup>. However, the sensitivity and responses of different tissue- and cancer-stem cells to DNA damage remain to be fully understood.

Differences in stem cell exposure to various exogenous mutagens have been identified<sup>3</sup>, but the *in vivo* consequences of endogenous sources of DNA damage, such as replication stress, and their contribution to mutagenesis have not yet been thoroughly investigated. Replication stress driven by oncogene expression is an important driver of genome instability in cancer<sup>4,5</sup>, as replication fork stalling and collapse participates in structural variant formation, chromosome rearrangements and aneuploidy. Many factors can endanger replication initiation

and replication fork progression: (1) availability of deoxyribonucleotides and replication factors like MCM helicases; (2) existing lesions on the DNA; (3) repetitive DNA and DNA secondary structures; (4) conflicts with transcription. One mechanism by which oncogene expression drives replication stress is by accelerating the cell cycle and bypassing cell cycle checkpoints, resulting in nucleotide depletion<sup>6-9</sup> and conflicts between DNA and RNA polymerases<sup>10,11</sup>. In hematopoietic stem cells, an increase of DNA damage and replication stress marks during aging was linked to lowered expression of MCM4 and MCM6<sup>12</sup> likely causing an age-related functional decline, as rapidly dividing hematopoietic progenitors are particularly sensitive to low levels of MCM complex proteins<sup>13</sup>. However, the factors contributing to replication stress, as well as the consequences of replication stress and its contribution to genome instability in adult stem cells remain unclear. Particularly, whether adult stem cells and cancer stem cells show tissue-specific susceptibility to replication stress has not been investigated.

Here, we use the *Drosophila* adult intestine and developing wing disc to characterize the consequences of replication stress on adult stem cells, and to uncover tissue-specific responses to replication stress. The *Drosophila* intestine as proven to be a useful model to study stem cell regulation and genome instability. Homeostasis of the adult *Drosophila* midgut is maintained by multipotent intestinal stem cells (ISCs) that can self-renew and differentiate into absorptive Enterocytes (ECs) and hormone-secreting enteroendocrine cells (EEs)<sup>14,15</sup>. Importantly, extensive work over the last 15 years has led to an understanding of many of the factors regulating stem cell proliferation and differentiation<sup>16</sup>. In addition, age-related functional changes occur in the intestine and the consequences of aging on stem cell function have also been partially described<sup>17</sup>. In particular, we previously demonstrated that somatic mutations arise frequently in aging intestinal stem cells (see Chapter 2 and <sup>18,19</sup>). Furthermore, somatic mutation events displayed hallmarks of replication stress, suggesting a role for replication stress in stem cell-specific DNA damage and genome instability<sup>20</sup>.

Here we investigate the consequences of induced replication stress on stem and progenitor cells. We find that intestinal stem cells are highly sensitive to replication stress. HU feeding and depletion of *RnrL* induced high levels of DNA damage in the stem cells, and was associated with stem cell loss in the tissue. However, *RnrL* depletion did not cause DNA damage in the developing wing disc. We identified a novel tissue-specific buffering mechanism, depending on gap junctions, that can limit stem and progenitor cells exposure to replication stress. Our data demonstrate that differences in cell-cell junctions and communication impact DNA damage susceptibility having important implications for understanding replication stress in healthy tissues and cancers.

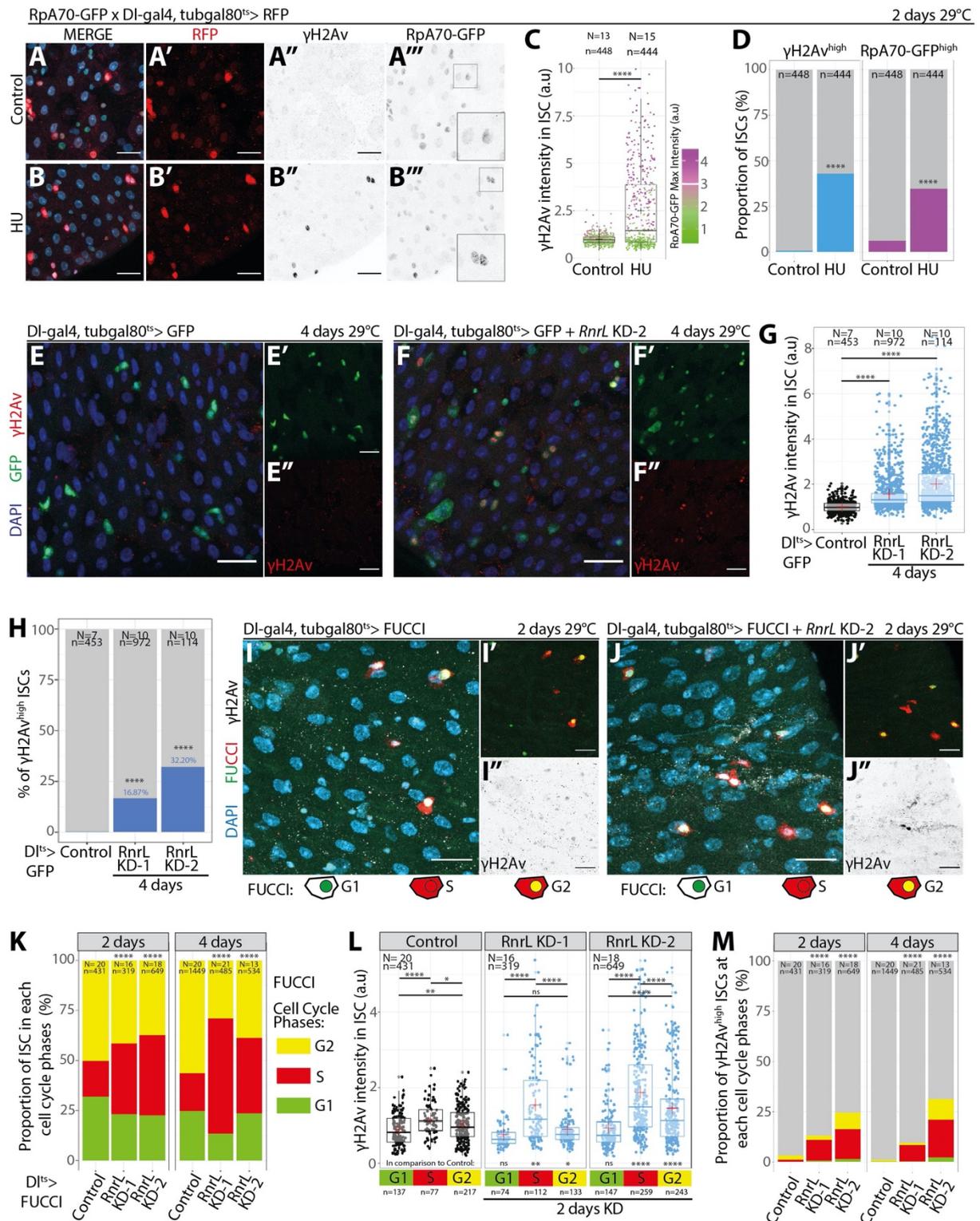
## II. RESULTS

### *Stem cell-specific induction of replication stress by RnrL knockdown*

Our previous data suggested a potential role for replication defects in stem cell mutagenesis<sup>18,20</sup>; we therefore wanted to examine how replication stress affects adult stem cell viability and DNA damage *in vivo*. To address these questions, we aimed to induce replication stress in *Drosophila* intestinal stem cells. As nucleotide depletion is a main cause of replication stress in cancer, we chose to induce replication stress by feeding flies with Hydroxyurea (HU). HU is a well-known inhibitor of the ribonucleotide reductase (Rnr), an enzyme of the nucleotide *de novo* synthesis pathway that reduces ribonucleotides into deoxyribonucleotides (dNTPs), the building blocks of DNA. Consequently, in cell culture and *in vivo*, HU limits the pool of deoxyribonucleotides available for DNA replication leading to replication fork stalling.

Feeding adult female flies with HU efficiently induced replication stress in ISCs, as seen by the accumulation of both  $\gamma$ H2Av and RpA-70 foci (Fig. 1A-C).  $\gamma$ H2Av is the *Drosophila* homologue for  $\gamma$ H2AX, a histone modification deposited at sites of DNA damage<sup>21,22</sup>, while RpA-70 is part of the Replication Protein A (RPA) complex that binds to single-stranded DNA, notably present at sites of

replication stress where replication forks are stalled and ssDNA is exposed<sup>23</sup>. HU feeding increased  $\gamma$ H2Av in ~40% of the intestinal stem cells (Fig. 1D). Stem cells with strong  $\gamma$ H2Av staining had brighter RPA70 foci (Fig. 1C). HU also decreased ISC proliferation as observed by a fewer phospho-Histone 3 (PH3) positive ISCs per guts (Fig. 1 supp 1A). Thus, HU can efficiently induce replication stress in



**Figure 1. Stem cell-specific induction of replication stress by *RnrL* knockdown. A-B''.** *Delta<sup>ts</sup>>RFP, RpA70-GFP* female midguts fed with control (A-A''') or HU-containing food (B-B''') for 2 days and stained for DAPI, RFP,  $\gamma$ H2Av, and GFP. Scale bar: 20 $\mu$ m. **C.**  $\gamma$ H2Av mean and RpA70 max intensity in each ISC (RFP+) nuclei in Control or HU condition. Normalized to the mean value in the control. Welch's test. **D.** Proportion of ISC (RFP+) with high levels of  $\gamma$ H2Av (>2) or RpA70 (>3). Fisher's test. **E-F''.** *Delta<sup>ts</sup>>GFP* midguts expressing no RNAi or *RnrL* RNAi-2, stained with DAPI, GFP and  $\gamma$ H2Av. Scale bar: 25 $\mu$ m. **G.**  $\gamma$ H2Av mean intensity in stem cell nuclei (GFP+) without or with *RnrL* RNAi, two different RNAi lines were tested, normalized to control mean. Red cross = mean of all cells. Welch's test. **H.** Proportion of ISC (GFP+) with high levels of  $\gamma$ H2Av (>2). Fisher's test. **I-J''.** *FUCCI;Delta<sup>ts</sup>* with or without *RnrL* RNAi. **K.** Proportion of cells in each cell cycle phase from *FUCCI;Delta<sup>ts</sup>* with or without *RnrL* RNAi. Chisquare test. **L.**  $\gamma$ H2Av mean intensity in each cell cycle phase after 2 days of *RnrL* RNAi or control in *FUCCI;Delta<sup>ts</sup>*. Normalized to mean of the control. Welch's test. Red cross = mean. **M.** Proportion of cells with high levels of  $\gamma$ H2Av staining (>2) and their cell cycle phase. Fisher's test. \*: p < 0.05; \*\*: p < 0.01; \*\*\*: p < 0.001; \*\*\*\*: p < 0.0001. N = guts, n = cells

ISCs. However, the response to HU feeding was highly variable, with some guts showing no apparent impact on DNA damage (Fig. 1 supp 1B), likely due to differences in feeding between the flies and HU incorporation in the ISC. In addition, HU also targets other replicating cells in the gut such as endocycling enterocytes (Fig. 1 supp 1C, supp 2), and potentially other proliferating organs like the ovaries. For these reasons, we explored additional methods to induce replication stress in a stem cell-specific manner.

Since HU inhibits Rnr, we reasoned that depleting RNA encoding components of the Rnr complex would phenocopy the effect of HU. Therefore, we depleted *RnrL*, the large subunit of the Rnr complex, in a stem cell-specific manner by driving RNA interference (RNAi) expression using the *DeltaGal4* driver. After 4 days of *RnrL* knockdown, ISC showed increased levels of DNA damage as measured by  $\gamma$ H2Av (Fig. 1E-G). This effect was validated by two independent RNAi lines (Fig. 1H). Comparable to HU feeding, *RnrL* RNAi also induced RpA-70 foci formation in the ISCs (Fig. 1 supp 1D-E''). Thus, *RnrL* RNAi efficiently induced DNA damage in the intestinal stem cells. In addition, it was also more consistent than HU feeding (Fig. 1 supp 1B).

To investigate the influence of *RnrL* depletion on the cell cycle, we combined *RnrL* RNAi with the FUCCI system (Fig. 1I-J''). The FUCCI transgenes<sup>24,25</sup> allow cell cycle phase determination of the cell in which they are expressed: GFP<sup>+</sup> alone marks cells in G1, RFP<sup>+</sup> alone marks cell in S phase, and GFP<sup>+</sup>/RFP<sup>+</sup> cells indicates cells in G2 or mitosis. ISC-specific expression of *RnrL*

RNAi along with FUCCI using the *DeltaGal4* driver led to an increased proportion of intestinal stem cells in S phase and a reduction of G1 and G2 phases (Fig. 1K). Specifically, the proportion of S-phase ISCs increased from 17.9% in the controls to 35.1% and 39.9% after 2 days of *RnrL* RNAi. This increase was even stronger after 4 days of knockdown. The increased proportion of cells in S-phase is likely due to a longer S-phase because replication is slowed down or blocked, as has been described upon HU treatment of cells in culture<sup>26</sup>.

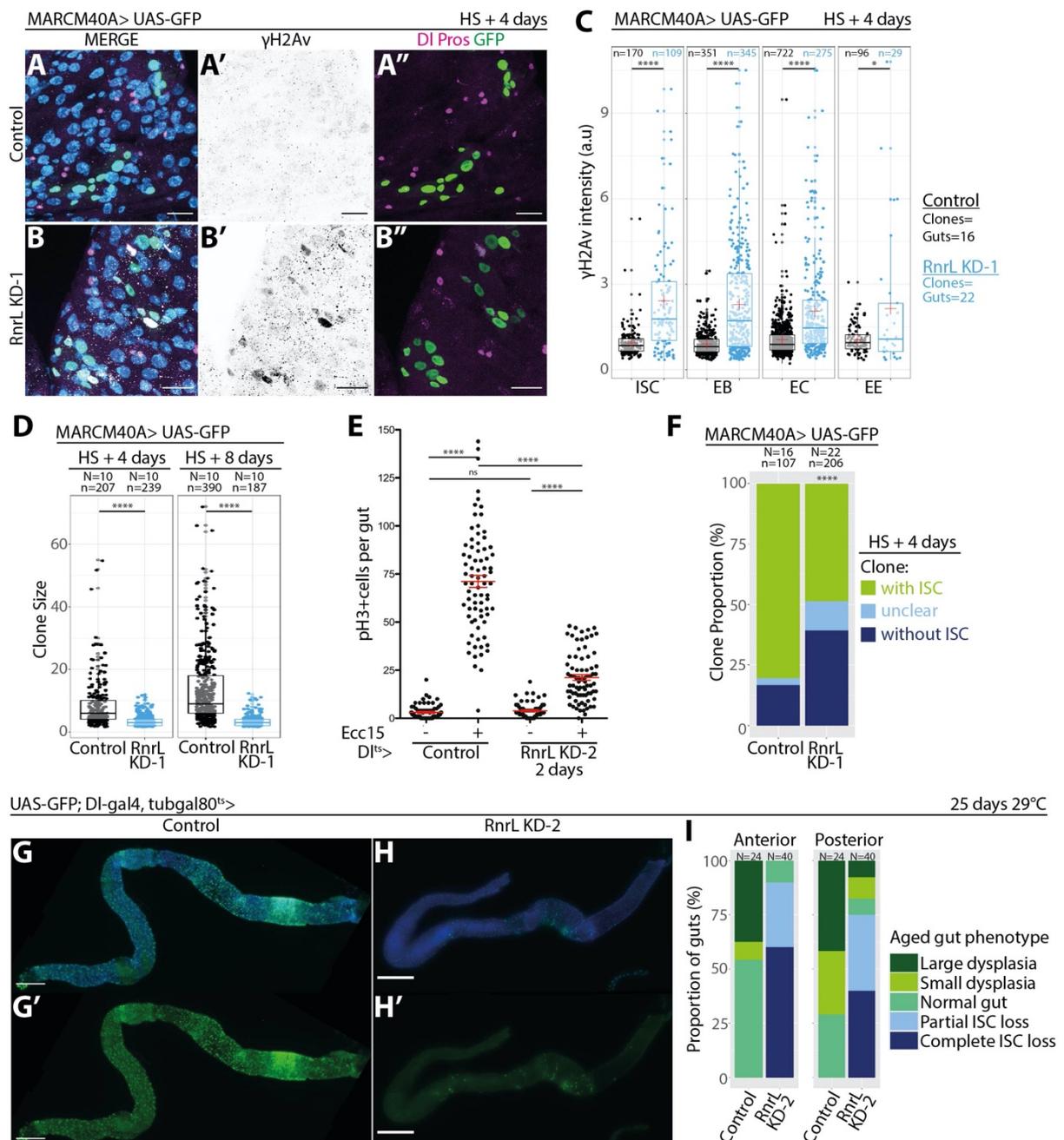
We then assessed the level of DNA damage in each cell cycle phase. In the control conditions, stem cells in each cell cycle phase had low levels of  $\gamma$ H2Av. As observed before, *RnrL* RNAi induced DNA damage in stem cells (Fig. 1I-J"). The increased levels of  $\gamma$ H2Av were stronger in S phase cells (Fig. 1L, Fig. 1 supp 1H) and the majority of cells with high levels of DNA damage were in S phase (Fig. 1M), suggesting that DNA damage is induced in S-phase. A smaller proportion of cells in G2 also showed an increase in  $\gamma$ H2Av levels (Fig. 1L, Fig. 1 supp 1H), which could be S-phase induced DNA damage inherited in G2. Alternatively, DNA damage in G2 could also be induced during DNA synthesis in G2 as in conditions of replication stress, DNA replication is often prolonged after S phase into G2 and even in mitosis. Thus, *RnrL* RNAi induced longer S phase and S phase-linked DNA damage in the stem cells. Altogether these data imply that reduction of *RnrL* activity induces replication stress, likely through nucleotide depletion, consistent with our knowledge of how HU affects cells. This genetic method to induce replication stress allows cell type-specific nucleotide depletion and is less variable than HU feeding.

### ***RnrL* knockdown drives stem cell loss and affects differentiating progenies**

Upon DNA damage, depending on its extent, stem cells can undergo transient cell-cycle arrest, differentiation, programmed cell death, or senescence<sup>1,2</sup>. We already observed that DNA damage induced by X-ray irradiation induced a transient blockage of mitosis (see Chapter II), likely through activation of the cell cycle checkpoints via activation of ATR and ATM, which are

also known to be required for stem cell proliferation<sup>27</sup>. However, the consequences of replication stress induced DNA damage on ISCs and tissue homeostasis are unknown. Thus, we asked what are the consequences of nucleotide depletion induced by *RnrL* RNAi on stem cell proliferation, maintenance and differentiation.

To understand the consequence of replication stress on the intestinal stem cells and lineage decisions, we generated GFP labelled stem cell clones using the Mosaic Analysis with a Repressible Cell Marker system (MARCM) (Fig. 2A-B'', Fig 2 supp 1). Cell types were identified based on specific markers and nuclear size



**Figure 2. *RnrL* knockdown drives stem cell loss and affects differentiating progenies. A-B''.** MARCM40A GFP clones 4 days after heat-shock with or without *RnrL* RNAi, stained for DAPI, GFP,  $\gamma$ H2Av, Delta (DI) and Pros. Scale bar: 20 $\mu$ m. **C.**  $\gamma$ H2Av mean intensity in each cell types in the clones, normalized to control mean. Red cross = mean. Welch's test. **D.** Clone size 4 days or 8 days after heat-shock in Control or *RnrL* RNAi conditions. Welch's test. **E.** PH3+ cells per gut in Control or *Delta<sup>ts</sup>>RnrL* RNAi flies without infection or 15hrs after Ecc15 bacterial infection by feeding. **F.** Proportion of clones with or without stem cells, 4 days after clone induction in Control and *RnrL* RNAi conditions. Stem cells were detected with DI staining (see A'',B''). The clones with cells for which the DI staining was not clearly positive or negative were classified as "unclear". Fisher's test. **G-H'.** *Delta<sup>ts</sup>>GFP* midguts aged for 25 days with or without *RnrL* RNAi. Stem cell are seen with GFP expression. **I.** Proportion of anterior and posterior midguts with dysplasia or stem cell loss.  
 \*:  $p < 0.05$ ; \*\*:  $p < 0.01$ ; \*\*\*:  $p < 0.001$ ; \*\*\*\*:  $p < 0.0001$ . N = guts, n = cells or clones (D,F).

(Fig. 2 supp 2A). Control clones showed low amounts of  $\gamma$ H2Av (Fig. 2A–A'', C, Fig. 2 supp 2B) whereas depletion of *RnrL* in clones resulted in high levels of DNA damage as shown by  $\gamma$ H2Av accumulation, which was especially prominent in ISCs and EBs (Fig. 2B–B'', C, Fig. 2 supp 2B), but also affected ECs and EEs. These findings are consistent with our above results using an ISC-specific Gal4 driver (Fig. 1).

To further examine how the different cell types are sensitive to nucleotide depletion-induced DNA damage, we expressed *RnrL* RNAi with the ubiquitous daughterless-Gal4-GeneSwitch driver (daGS) (Fig. 2 supp 2E–F''). After 4 days of induction, ISCs (Delta-LacZ expressing cell), and EBs (identified by their nuclear size), had extensive DNA damage (Fig. 2 supp 2G–J). While enteroendocrine cells did not have a significant difference in  $\gamma$ H2Av compared to the control (Fig. 2 supp 2K), ECs had increased  $\gamma$ H2Av staining (Fig. 2 supp 2L). ECs with the smaller nuclei, likely young ECs recently produced in the tissue, were more affected (Fig. 2 supp 2M). While in the MARCM experiments, we assess DNA damage in newly formed progeny in the clones, the analysis here also included cells that are terminally differentiated and no longer mitotically active, which could account for the differences observed regarding the effect of *RnrL* RNAi on the EE cells. Enteroendocrine cells differentiate from ISCs through an Enteroendocrine Precursor (EEP) that divide one last time to generate two EEs. In contrast, EB differentiation into ECs requires several endocycles, during which the cells replicate their DNA to become polyploid cells. Thus, both EC and EE lineages rely on DNA replication in the early steps of their differentiation and likely

the number of rounds of replication impacts the degree to which DNA damage is induced. Alternatively, DNA damage induced by *RnrL* RNAi in the differentiating progenies may have been inherited from the stem cell division.

To gain insight on the consequences of *RnrL* depletion on differentiation, we examined the composition of MARCM clones. *RnrL* RNAi led to a reduction of EEs and polyploid ECs proportions compared to the control (Fig. 2 supp 2C). The ECs in *RnrL* knockdown clones also had smaller nuclei (Fig. 2 supp 2D), suggesting an impairment of polyploidization. These observations are in agreement with a requirement of RnrL for proper differentiation of EEs and ECs, likely for EEP division and EC endocycles respectively. However, we cannot exclude that EE and EC proportion reduction were due to cell death in the clones.

Since we observed extensive DNA damage in the stem cells following *RnrL* RNAi and an increased proportion of ISCs in S phase, we hypothesized that ISC proliferation potential would be changed in this context. Measuring the clone size gives us insight on the cell division rate of ISCs. *RnrL* RNAi clones were smaller than control clones 4 days AHS, and did not grow more within 8 days (Fig. 2D). Thus, ISC proliferation is altered in condition of *RnrL* RNAi. We also analyzed the proliferation consequent to *DeltaGal4* driven *RnrL* depletion. The proliferation in the whole gut is low in young flies and short term *RnrL* RNAi expression did not decrease the amount of PH3 positive cells per gut (Fig. 2E). Tissue damage following *Ecc15* bacterial infection is known to increase ISC proliferation rates to promote tissue regeneration<sup>28</sup>. Consistent with this, untreated guts had a mean of 3 PH3+ cells per gut versus 71 PH3+ cells for *Ecc15* treated (Fig. 2E). 2 days after expression of *RnrL* RNAi, stem cell proliferation upon *Ecc15* treatment was still increased relative to *RnrL* RNAi controls and wild-type controls, but to a much lesser extent than that of wild-type controls (21 PH3+ cells per gut). Therefore, ISC proliferation in homeostasis and regeneration is blocked in conditions of nucleotide depletion induced by *RnrL* RNAi.

We then wanted to understand why loss of stem cell proliferation occurred upon accumulation of DNA damage. We presumed that it could result from cell

cycle arrest or stem cell loss. In the MARCM clones, 40% of the *RnrL* RNAi clones did not have ISC (Fig. 2F), a much higher proportion compared to the control. Furthermore, to assess stem cell maintenance in the tissue, we compare aged control guts to those in which *RnrL* was expressed in ISCs for 25 days. Previous studies showed that during normal midgut ageing, ISCs increase proliferation rates in response to dysbiosis of commensal bacteria leading to accumulation of stem cell and progenitor cells, so-called "dysplasia"<sup>29</sup>. Consistent with this occurring, we detected ~40% guts with large and ~30% guts with small dysplasia in the posterior midgut (Fig. 2G, 2I). However, after *RnrL* depletion guts rarely developed dysplasia. Conversely, ISCs were lost in almost all midguts in both anterior and posterior regions (Fig. 2H-I). Therefore, replication stress induced by *RnrL* depletion causes stem cell loss.

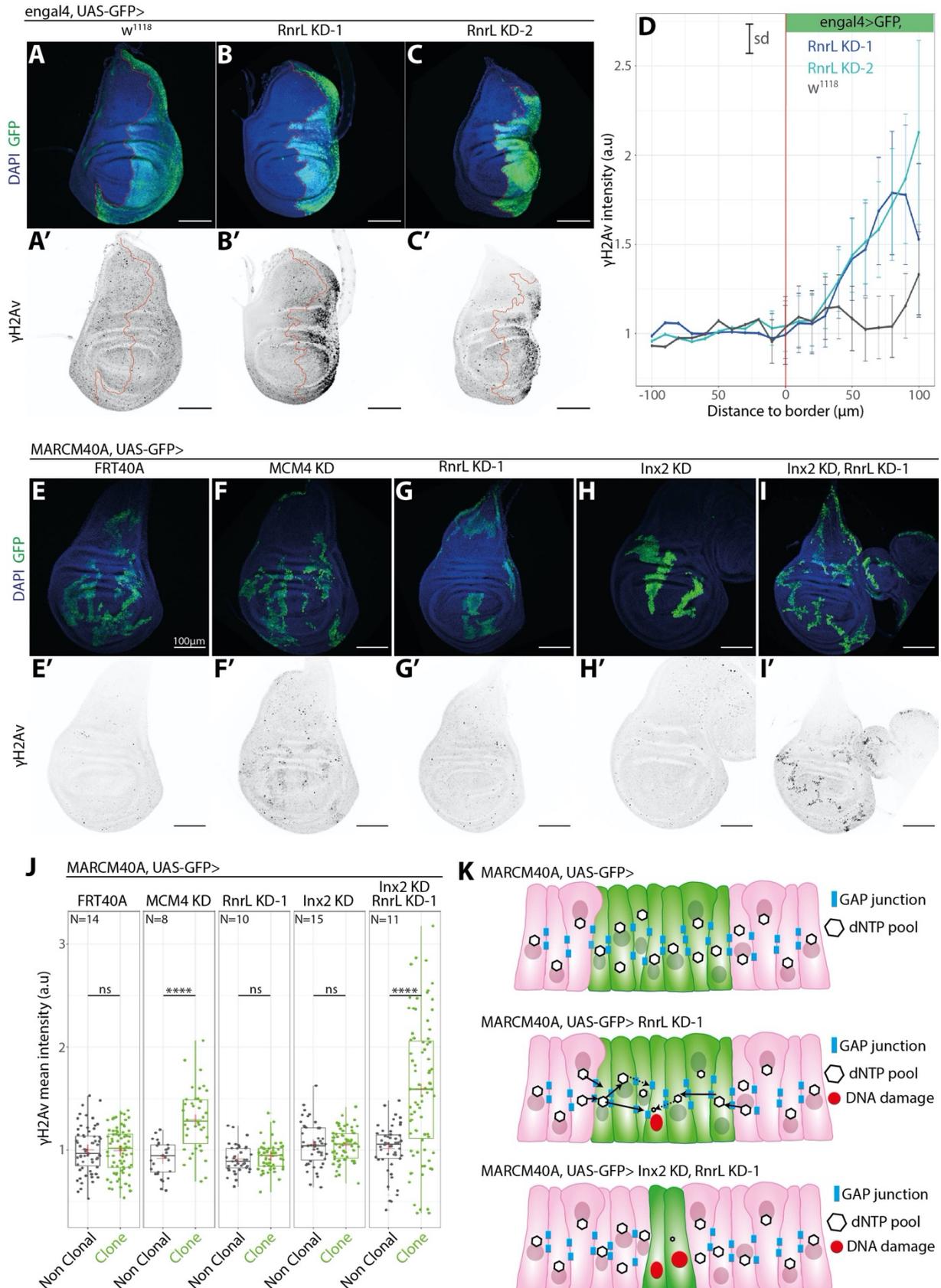
To conclude, we find that inactivation of *RnrL* limits proliferation in the gut by causing S phase delays and inducing stem cell loss after DNA damage accumulation. Moreover, differentiating progeny are also affected by the loss of *RnrL* in the gut, with ECs having small nuclei suggestive of defective endoreplication.

### *Non-autonomous rescue of DNA damage upon RnrL knockdown in the developing wing disc*

Our data thus far indicate that intestinal stem cells are highly sensitive to replication stress induced by cell-autonomous nucleotide depletion. However, it is not clear whether other proliferating cells would be similarly susceptible to *RnrL* depletion or if sensitivity to nucleotide depletion is a specific feature of the gut cells. Using the cell-type specific *RnrL* knock-down, assessed how progenitors are affected by nucleotide depletion another tissue, the developing wing disc.

In order to determine the effect of induced replication stress via inactivation of *RnrL* in the developing wing, we first used the *engrailed-Gal4* driver to knockdown *RnrL* in the posterior half of the wing disc in larvae. Both *RnrL* RNAi lines induced high levels of DNA damage in the wing disc (Fig. 3A-C'). However,

the cells accumulating DNA damage were only located at the posterior edge of the wing disc. The cells closer to the border of the engrailed domain, in the middle of the disc, did not display increased  $\gamma$ H2Av levels (Fig A–C'). The *RnrL* RNAi–



**Figure 3. Non-autonomous rescue of DNA damage by gap junctions in the wing disc.** A-C. *enGal4>GFP* control discs (A-A') or disc expressing two different *RnrL* RNAi (B-C'). The discs were dissected at late larval L3 stage and stained for  $\gamma$ H2Av. The border of the *engrailed* domain is showed with a red line. Scale bar: 100 $\mu$ m. D.  $\gamma$ H2Av intensity along the antero-posterior axis and relative to the engrailed domain border in control and *RnrL* expressing wing discs. For each disc, the  $\gamma$ H2Av level is normalized to the level outside the *engrailed* domain. E-I'. MARCM GFP+ clones 4 days after induction in the wing disc and stained for  $\gamma$ H2Av, with control clones (E-E'), MCM4 RNAi (F-F'), *RnrL* RNAi (G-G'), *Inx2* RNAi (H-H') and *RnrL/Inx2* RNAi (I-I'). Scale bar: 100 $\mu$ m. J. Quantification of  $\gamma$ H2Av levels inside and outside the clones, normalized to control non-clonal tissue mean intensity. N=Discs. Welch's test, red cross = mean. K. Model for the non-cell autonomous rescue of DNA damage in the disc through gap junctions by providing dNTPs.

\*:  $p < 0.05$ ; \*\*:  $p < 0.01$ ; \*\*\*:  $p < 0.001$ ; \*\*\*\*:  $p < 0.0001$ .

mediated DNA damage induction was only seen 50  $\mu$ m posteriorly away from the engrailed domain border (Fig. 3D). We speculated that this could be explained by a non-cell autonomous rescue of *RnrL* RNAi, possibly by nucleotide transfer from the anterior part of the disc.

To further test a potential rescue mechanism from non-autonomous wild-types tissue, we then used the MARCM system to induce clones in the developing wing disc. If, indeed, *RnrL* depletion is rescued non-cell autonomously by the neighboring tissue, clones expressing *RnrL* RNAi should have DNA damage restricted to the center of the clone, since the edge of the clone would benefit from rescue by the surrounding wild-type tissue. Clones were induced by heat-shock at the first larval stage, the discs were dissected 4 days later and stained for  $\gamma$ H2Av. As a positive control for cell-autonomous induction of replication stress, we induced MCM4 RNAi clones. MCM4, known as *dpa* in *Drosophila*, is part of the helicase complex important for replication origin firing and DNA unwinding during DNA synthesis. In *Drosophila*, loss of MCM4 is known to create mitotic abnormalities and defects in BrdU incorporation<sup>30,31</sup>. In budding yeast and in human cells, depletion of MCM4 creates a S-phase block<sup>32,33</sup>. MCM4 RNAi clones showed an increase of  $\gamma$ H2Av levels throughout the clone, consistent with a cell-autonomous induction of replication stress (Fig. 3E-F', 3J). In contrast, however, *RnrL* RNAi clones had no increase in  $\gamma$ H2Av staining compared to the wild type tissue and to the control clones (Fig. 3G-G', 3J), suggesting that *RnrL* depletion could be rescued by the wild-type tissue. High levels of DNA damage were occasionally observed in *RnrL* RNAi clones, but exclusively localized in spots at

the center of the clone, far from the wild-type tissue (Fig. 3 supp 1A–B). Altogether, this supports the idea that *RnrL* RNAi is non-cell autonomously rescued in the wing disc.

### *Gap junctions are required for non-autonomous rescue of RnrL inactivation*

Considering the relationship between the location of DNA damage and the proximity to wild-type tissue, we hypothesized that nucleotides or nucleotide precursors could be transferred from wild-type cells to *RnrL* knockdown cells, through intercellular junctions. Gap junctions are intercellular channels that allow the passage of small metabolites (<1kDa), such as ions, sugars and nucleotides<sup>34</sup> between cell cytoplasm. In *Drosophila*, eight Innexins proteins (Inx1–8) have been identified to contribute to gap junctions. They form homotypic or heterotypic transmembrane octameric hemichannels. The joining of two hemichannels from adjacent cells makes a gap junction connection<sup>35</sup>. Previous work in the wing disc suggested that a sugar, GDP–L–fucose, was transported between cells in a manner that relied on the gap junctions component Inx2<sup>36</sup>. In addition, the GDP–L–fucose transfer between cells was sufficient to rescue fucosylation activity in clones deficient for *Gmd*, an enzyme necessary for *de novo* GDP–L–fucose synthesis. We wondered if a similar process could explain the non-autonomous rescue of the *RnrL* phenotype. Could gap junctions allow the passage of nucleotides or nucleotide precursors between cells?

If gap junctions participate in non-autonomous rescue of DNA damage, the combined loss of *RnrL* and gap junctions should result in high levels of DNA damage in clones. As described previously, control clones and *RnrL* RNAi clones did not display higher levels of DNA damage (Fig. 3E–G', 3J). Importantly, *Inx2* RNAi alone did not induce DNA damage in the clones (Fig. 3H–H', 3J). However, clones with co-depletion of *RnrL* and *Inx2* had bright  $\gamma$ H2Av staining throughout and appeared smaller than control clones (Fig. 3I–I', 3J). To confirm these results, we also co-expressed *RnrL* RNAi with a RFP-tagged version of Inx2, acting as a dominant negative form<sup>37</sup>. Consistent with our *inx2* RNAi data, RFP–Inx2<sup>DN</sup> *RnrL*

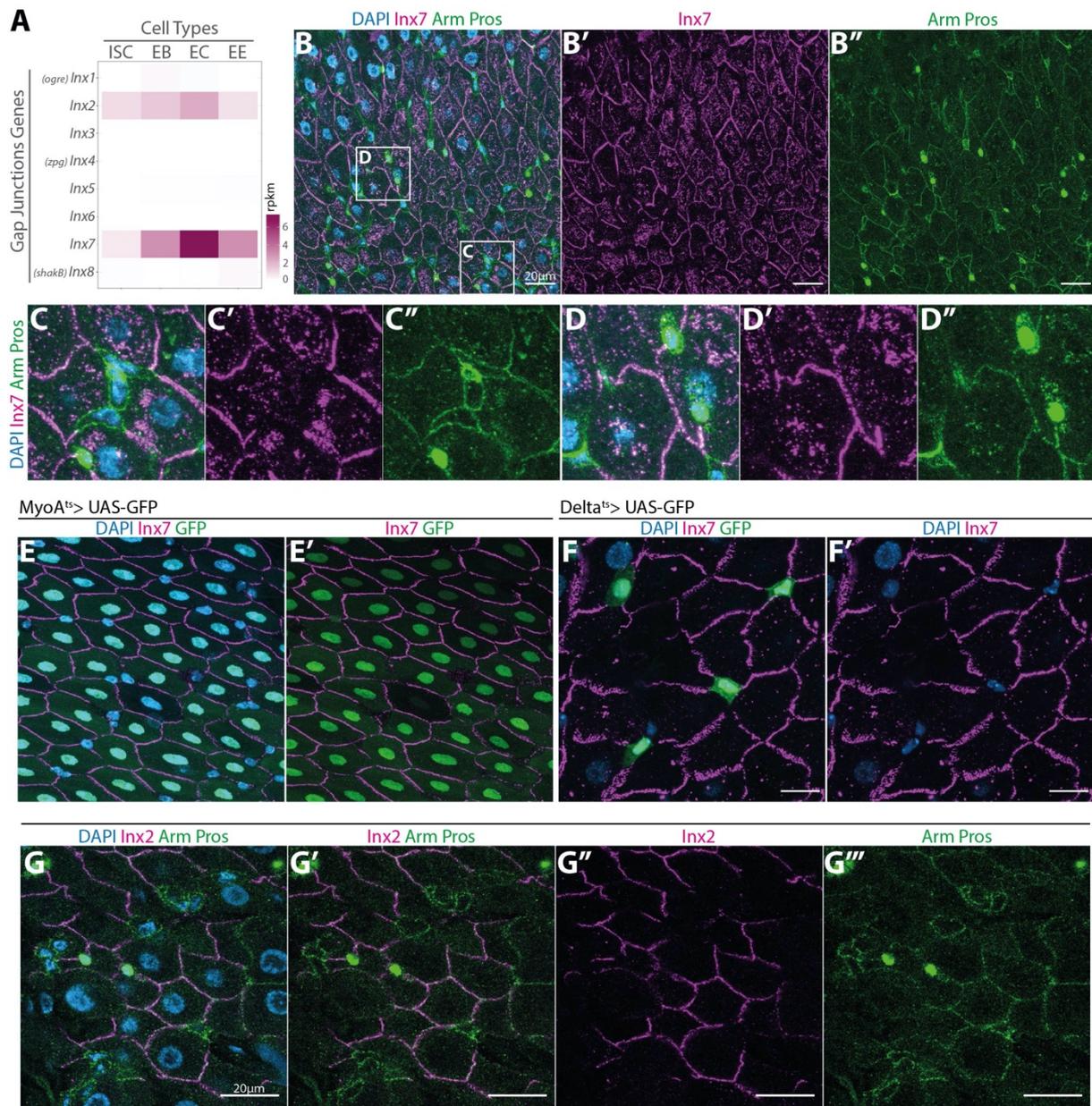
RNAi clones were small and had higher levels of DNA damage compared to the wild-type surrounding tissue (Fig. 3 supp 1C–G). To conclude, in the wing disc, gap junctions rescue DNA damage induced by *RnrL* RNAi, likely by allowing nucleotide transfer between cells for DNA replication.

***Inx2 and Inx7 are not expressed in ISCs and are restricted to enterocytes in the Drosophila midgut***

We showed before that intestinal stem cells were highly sensitive to *RnrL* depletion, but wing discs epithelial cells were not because of non-cell autonomous mechanism that depends on gap Junction. To understand better these different responses between ISCs and wing disc progenitor cells, we investigated gap junction proteins expression and localization in the adult midgut.

We first, surveyed the expression of gap junction encoding genes in the different cell types of the midgut using an available RNAseq dataset<sup>38</sup>. We found that only *Inx2* and *Inx7* were expressed in the midgut (Fig. 4A) with very low expression in ISCs, weak expression in EBs and highest level of expression in ECs.

We then determined the localization of gap junctions and their contribution to the epithelial tissue architecture of the intestine. We examined the localization of *Inx 7* and *Inx2* in the gut by immunofluorescence. First, we found a strong membrane staining for *Inx7* in the midgut epithelium (Fig. 4B–B'). In particular, *Inx7* staining was prominent at the membrane of large nucleated cells. In contrast, *Inx7* was absent from the membranes of small nuclei progenitor cells (ISCs and EBs) where the staining for Armadillo, an adherens junction component, was enriched (Fig. 4B–C"). Similarly, *Inx7* was absent at the membrane of the Prospero-expressing enteroendocrine cells (Fig. 4B–B", 4D–D"). Thus, *Inx7* localizes at EC membranes but is absent from EE, ISC and likely EB membranes.



**Figure 4. gap junctions in the gut are restricted to ECs.** **A.** Expression levels of the 8 Innexin genes in the intestinal cell types (data from Dutta et al. 2015). **B-D''.** *w<sup>1118</sup>* female midguts stained for DAPI, Innexin 7, Armadillo and Prospero. **E-E'.** *MyoA<sup>ts</sup>>GFP* midguts stained for DAPI, GFP and Inx7. *MyoAGal4* drives GFP expression in the ECs. **F-F'.** *Delta<sup>ts</sup>>GFP* midguts stained for DAPI, GFP and Inx7. *DeltaGal4* drives GFP expression in the ISCs. **G-G''.** Female midguts stained for DAPI, Innexin 2, Armadillo and Prospero. Dissection and staining done by M. El-Hajj.

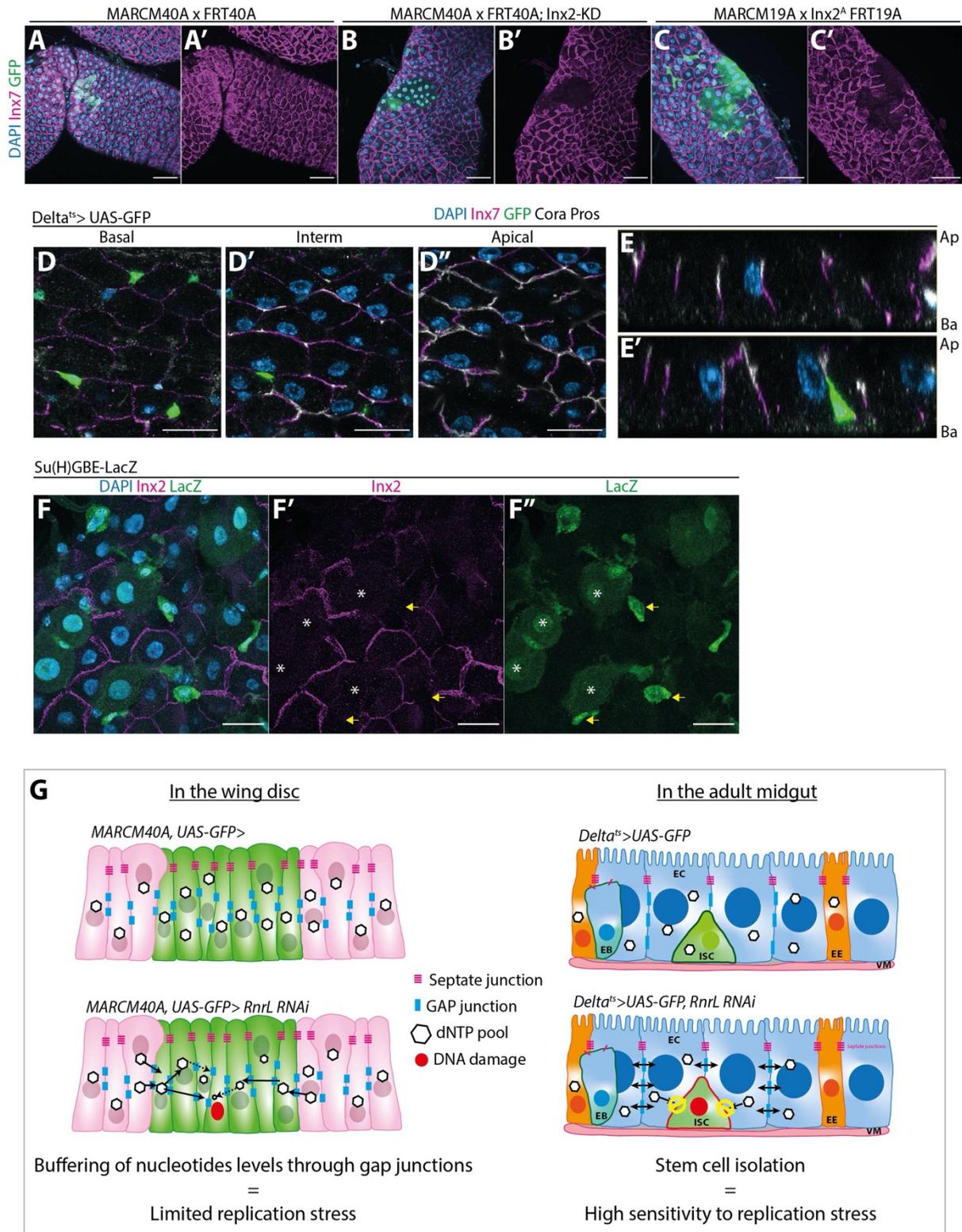
Confirming these results, Inx7 was detected at the cell membranes of ECs marked by *MyoA-Gal4*-driven GFP expression (Fig. 4E-E'), but was lacking from the cells with ISC-specific *DeltaGal4*-driven GFP expression (Fig. 4F-F'). Similar to Inx7 staining, the gap Junction protein, Inx2 was also found at the membrane of ECs but was lacking from the membranes of other cell types (Fig. 4G-G''').

Interestingly, thus far, the coexpression or colocalization of Inx7 and Inx2 has not been reported in any other tissues. In order to test whether Inx7 and Inx2

form heterotypic gap junctions, we assessed *Inx7* localization in the absence of *Inx2*. In clones expressing *Inx2* RNAi or an *Inx2* mutant protein (*Inx2<sup>A</sup>*), *Inx7* staining was lost (Fig. 5A–C'). Consequently, *Inx2* is required for *Inx7* localization and *Inx7* cannot form hemijunction without *Inx2*. *Inx7* is likely forming heterotypic gap junctions with *Inx2* between ECs. Although expressed in ECs, both *Inx2* and *Inx7* were lacking from EC membranes in contact with other cells that did not express them (Fig. 4, Fig.5A–C'). This suggests that to be stabilized at the membrane, gap junction hemichannels need to find a partner to connect with on the adjacent membrane.

We then investigated the relationship of gap junctions with intercellular junctions of the gut epithelium. We observed that gap junctions were localized in the basal section of the epithelium in the same plane as stem cells, but they were lost apically where septate junctions were found (Fig. 5D–E'). Thus, gap junctions are localized baso–laterally between ECs (Fig 5D–E'). Septate junction component expression increases in EBs and is required during EC differentiation for the integration in the epithelium<sup>39,40</sup>. We investigated, whether like septate junctions, gap junctions were established during EB to EC differentiation<sup>41</sup>. We used the EB reporter line *Su(H)GBE–LacZ* to address this question. *LacZ* was exclusively expressed in the EBs, but the *LacZ* protein persisted into newly–made ECs. We found that *Inx2* was not localized on EB membranes or in newly formed ECs as opposed to older ECs (Fig. 5F–F''). Thus, unlike septate junctions, gap junctions were likely formed in differentiated ECs and not required for integration in the epithelium of differentiating EBs.

To conclude, we find that gap junctions in the adult midgut are restricted to differentiated ECs and are absent from intestinal stem cells. This could explain the vulnerability of ISC to replication stress induced by nucleotide depletion compared to disc progenitor cells. Indeed, without gap junctions, ISCs may solely rely on cell–autonomous production of dNTP for DNA replication.



**Figure 5. *Inx2* is required for *Inx7* localization and gap junction are established in differentiated ECs. A-C'** MARCM GFP+ clones in midguts stained for *Inx7*. Control clone (A-A'), *Inx2* RNAi clone (B-B'), *Inx2*<sup>A</sup> mutant clone (C-C'). **D-E'** *Delta*<sup>ts</sup>>*GFP* midguts stained for *Inx7* for gap junctions and Coracle for septate junctions. Different views of the epithelium, basal (D), intermediate (D') and apical (D''), as well as orthogonal reconstitution from the Z-stack (E-E'). **F-F''** *Su(H)GBE-LacZ* midguts stained for *Inx2*. LacZ is expressed in EBs (yellow arrows) and persists in newly made ECs (asterisks). **H**. Model of non-cell autonomous rescue of DNA damage in the wing disc through the gap junction, while ISC are isolated from the ECS because they do not express gap junction proteins. Thus, ISCs do not have a nucleotide level buffering via gap junction in condition of *RnrL* RNAi.

### III. DISCUSSION

Our findings highlight important tissue-specific differences in regulating responses to replication stress and suggest a role for gap junctions in this process. By limiting dNTP production through genetic inactivation of the critical nucleotide metabolizing enzyme, *Ribonucleotide reductase*, we find that ISCs but not wing disc precursor cells acquire DNA damage. Our results indicate that wing disc precursors require Gap Junctions to prevent DNA damage. We further observe that gap junction proteins are lacking in ISCs, EBs, EEs, and young ECs. Altogether, these findings lead us to propose a model whereby transport of dNTPs or dNTP precursors through gap junctions allows tissue-buffering of replication stress: a lack of gap junctions in ISCs may make them exquisitely sensitive to changes in nucleotide levels (Fig 5G).

The wing disc was previously shown to have active gap junctions that can allow fucose transport between cells<sup>36</sup>, here we provide evidence that Gap Junctions are also involved in buffering the effects of nucleotide depletion. Two lines of evidence support this: First, upon knockdown of *RnrL*, we find that cells adjacent to wild-type cells do not acquire DNA damage. Upon co-inactivation of Gap Junction components, DNA damage was now detected throughout the entire region of *RnrL* knockdown, suggesting that Gap Junctions participate in non-autonomous rescue. It is formally possible that hemi-junctions of innexin 2 between wing disc cells and the hemolymph, as opposed to Gap Junctions between wing disc cells, could support dNTP uptake and rescue. However, this would not explain the rescue adjacent to wild-type cells that is seen at the junction of *RnrL* knockdown tissue. Therefore, we favor a model in which Gap Junctions allow passage of metabolites necessary for dNTP production helping to buffer changing nucleotide pools.

Our data show that the Gap Junction proteins *Inx2* and *Inx7* are co-expressed in the gut but only in mature ECs. They apparently form molecular complexes that localize to basolateral EC membranes as *Inx7* requires *Inx2* for both cell-autonomous as well as non-autonomous membrane localization. The

lack of gap junctions in ISCs and progenitor cells may have important physiological basis as gap junctions are known to transport small molecules such as reactive oxygen species (ROS) and calcium (Ca<sup>2+</sup>) as well as to electrically couple of cells such as neurons. Interestingly, both ROS and Ca<sup>2+</sup> have been previously shown to have essential function in regulating stem cell proliferation<sup>16</sup>. We speculate that the lack of gap junction is important for isolating stem cells from Ca<sup>2+</sup> or ROS production in neighboring EC cells and may be necessary to prevent erroneous activation of proliferation cascades. A negative consequence of the lack of gap junctions, however, may be the inability to ISCs to properly buffer nucleotide levels resulting in replication stress. Whether the lack of gap junctions in ISCs promotes replication stress contributing to endogenous DNA damage and spontaneous mutation needs to be further investigated. The lack of gap Junctions in young ECs is also intriguing. Mechanistically, it is possible that the EC lateral membranes require time to mature and can only then insert gap junctions. This difference in gap junctions plays a role in preventing ROS passage from damaged ECs undergoing turnover to adjacent, newly made ECs. Further studies will be required to test whether there are functional consequences of gap junction difference between young versus old ECs, such as control of EC turnover through ROS passage.

Oncogene-driven replication stress in human cancers is associated with low nucleotide pools, and is a significant driver of genome-instability<sup>6,42</sup>. Our findings that gap junctions may be important regulators of nucleotide pools has important implications for our understanding of replication stress *in vivo* in healthy tissues as well as in contexts of oncogene-driven tumor growth. In humans, 21 genes encode Connexins, the functional equivalent of Innexins in flies, which assemble into multimeric complexes and form intercellular gap junctions. We hypothesize that different normal as well as cancer cell types vary greatly in their ability to buffer dNTP pools through gap junctions. Buffering capacity could be dramatically different depending on the expression levels of connexins, gap junction activation state, and distinct multimeric complexes exhibiting different gating properties.

Consistent with the notion that cancer cells may have altered gap junctions, a classic study demonstrated that while healthy liver cells were electrically coupled, liver cancer cells lacked electrically coupling<sup>43</sup>. In addition, over the years, connexins have been reported to have both pro- and anti-tumorigenic properties<sup>44,45</sup>, though many of these studies are correlative. Our findings suggest that an important future direction is to gain a better understanding of the role of gap Junctions and their potential to buffer nucleotide levels and prevent replication stress in normal and cancer contexts.

#### IV. MATERIALS AND METHODS

##### Fly Stocks – Aging

The experiments presented in this manuscript used the following fly lines.

From the Bloomington stock center: *w; UAS-2XGFP* (BL6874); *w;UAS-RFP* (BL30556); *UAS-RnrL RNAi #1* (BL51418); *UAS-RnrL RNAi #2* (BL44022); *UAS-MCM4 RNAi* (BL28620); *Delta-LacZ* (BL11651); *MARCM40A* (BL5192); *MARCM82B* (BL5135); *FRT82B* (BL2050); *Inx2<sup>A</sup> FRT19A* (BL54481)

The following stocks were gifts: *w<sup>1118</sup>* (M. McVey); *RpA70-GFP* (E. Wieschaus)<sup>23</sup>; *FUCCI (w1118; QUAS-GFP-E2F1<sub>1-230</sub><sup>#2</sup>, QUAS-mRFP1-NLS-CycB<sub>1-266</sub><sup>#3B</sup>/CyO wg-lacZ; MKRS/TM6B*, B. Edgar)<sup>24</sup>; *En-Gal4* (Y. Bellaiche); *UAS-Inx2RNAi & UAS-RFPInx2<sup>DN</sup>* (P. Spéder)<sup>37</sup>; *MyoA<sup>ts</sup>* (J. Jiang); *Su(H)GBE-LacZ* (S. Bray); *daGS* (M. Rera); *MARCM19A* and *FRT19A* (A.Gould); *FRT40A* (A. Bardin);

The *Delta<sup>ts</sup>* line was generated by recombination of the *DeltaGal4/TM6TbHu* (S. Hou), and the *tub-Gal80<sup>ts</sup>* from another line. It was used to generate: *UAS-2xGFP;Delta<sup>ts</sup>/TM6TbHu*; *UAS-RFP;Delta<sup>ts</sup>/TM6TbHu*; *FUCCI;Delta<sup>ts</sup>/TM6TbHu*

Unless indicated the flies were kept at 25°C on standard medium with yeast.

For the temperature sensitive experiments, flies were crossed at 18°C and the progeny kept at 18°C until 2–3 days old adult were collected for experiments and transferred at 29°C for the time indicated on the figures before dissection.

For the aging experiment, the adult flies were aged in tubes with 15–20 females and 3–5 males and transferred to fresh medium every 3–4 days.

### *HU/RU feeding*

For the hydroxyurea feeding experiments, 3–4 days old flies were shifted in food supplemented with 10mg/mL of HU diluted in water or water alone.

For the GeneSwitch experiments, 3–4 days old flies were shifted on food containing 50µg/mL of RU486 diluted in EtOH, or EtOH alone.

### *MARCM Clone Induction*

For the clone induction in the gut, adult flies were heat-shocked (HS) at 36.5°C for 30mins at 2–3 days old and the kept at 25°C on yeasted food for the time of the experiment.

For the clone induction in the discs, flies were left on yeasted food to lay egg for 24hrs, and the progeny was HS 24hrs later, at larval stage 1, at 37°C for 45mins. The wing disc were dissected 4 days later from late L3 larvae.

### *Immunostaining*

The tissues were fixed and stained following a previously published protocol<sup>46</sup>. Briefly, the flies were dissected in PBS and the tissues were fixed at room temperature in 4% paraformaldehyde for 2hrs (female guts) or 30mins (discs). They were then washed in PBS+ 0.1% Triton X-100 (PBT). To remove the content of the gut lumen, the fixed gut were equilibrated 30mins in PBS 50% Glycerol and then washed in PBT. The primary antibody staining was performed overnight at 4°C. After washing with PBT 20 mins 3 times, the secondary antibody was incubated 3–4hrs at room temperature. After washing with PBT, DAPI staining was done in PBT for 30 mins (1µg/ml).

The antibodies used were: Chicken anti-GFP (1:2000, Invitrogen #A10262), Rabbit anti-RFP (1:1000, Clontech #632496), Rabbit anti- $\gamma$ H2Av (1:2000, Rockland 600-401-91), Mouse anti- $\gamma$ H2Av (1:200, DSHB UNC93-5.2.1), Rabbit anti-Inx7 (1:100, H. Bauer<sup>47</sup>), Guinea Pig anti-Inx2 (1:1000, G. Tanentzapf<sup>48</sup>), mouse anti-Delta (1:1000, DSHB C594.9B), mouse anti-Pros (1:2000, DSHB MR1A-c), Goat anti-LacZ (1:200, F. Schweisguth), mouse anti-Arm (1:200, DSHB N2 7A1), anti-Cora (???), rabbit anti-PH3 (ser10, 1:2000, Merck Millipore #06-570).

The secondary antibodies were all from Jackson Laboratory, raised in Donkey with different fluorophores (DyLight 488, 549, 649)

### **Microscopy and Image Analysis**

Pictures showed here were taken using the Zeiss confocal microscope LSM900 with 40x objective (most of the gut pictures) or 63x objective (gap junction close up); Zeiss confocal microscope LSM780 with 25x objective (discs pictures); Zeiss Apotome microscope 10X (whole-gut pictures)

The images were processed and the quantification performed using Fiji, with manual detection of cells or automatic detection of cells with a macro that we designed. Exceptionally Imaris(c) software was used for the quantification presented in Fig. 1 Supplementary 1D).

$\gamma$ H2Av quantification represents the mean intensity in  $\gamma$ H2Av staining for each nucleus. For the RpA70-GFP intensity, the maximum intensity for each nucleus is considered, it corresponds to the brighter foci of RpA70-GFP nuclei.

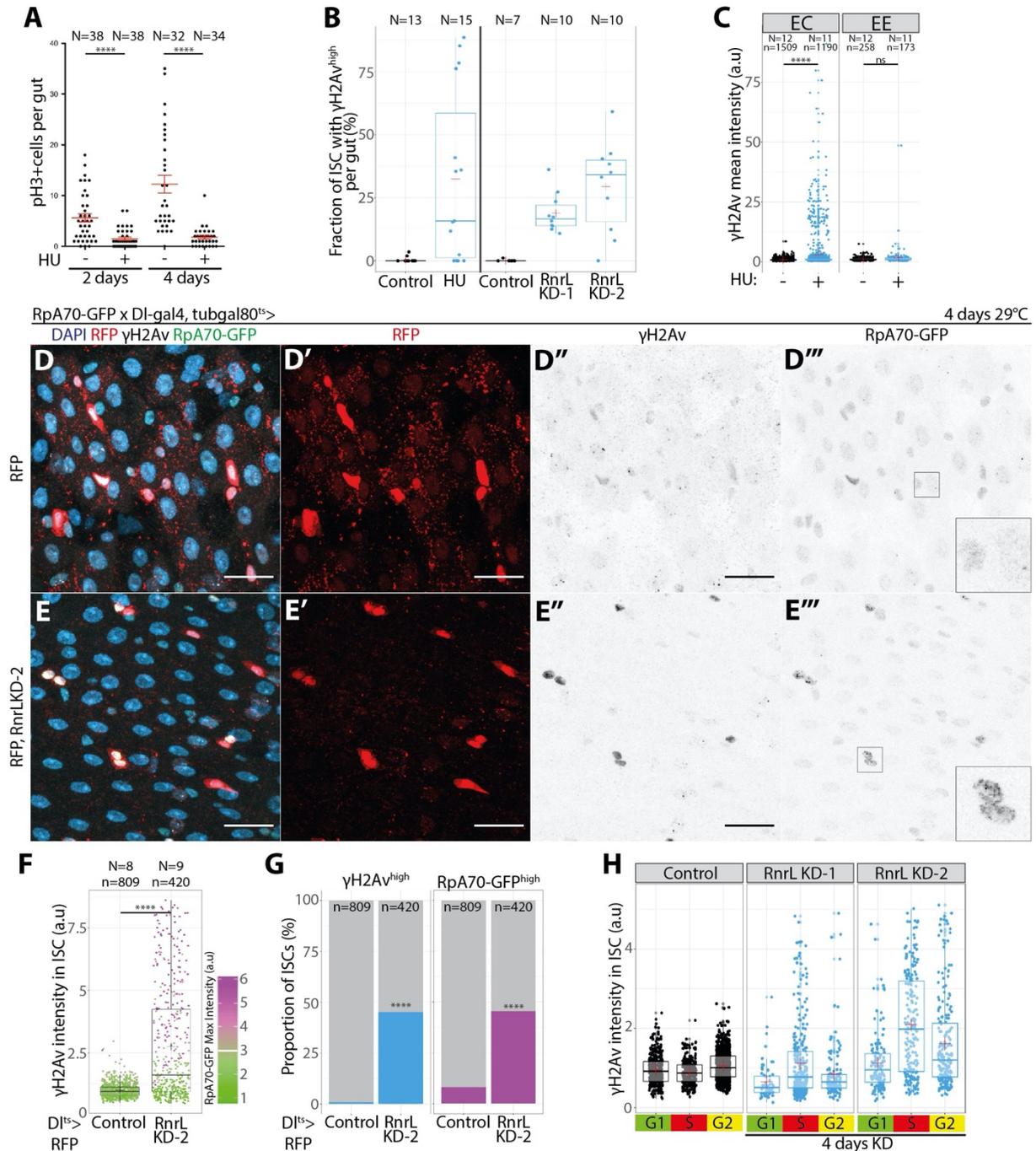
### **Data Processing and Statistics**

Data Analysis was performed on Excel, Prism (version 9.0.1) or R (version 4.0.2) using the Rstudio interface (version 1.3.1073) and ggplot2 package (version 3.3.3). Initial training for the use of the ggplot2 package for data visualization was obtained by the U900 Bioinformatics unit of the Institute.

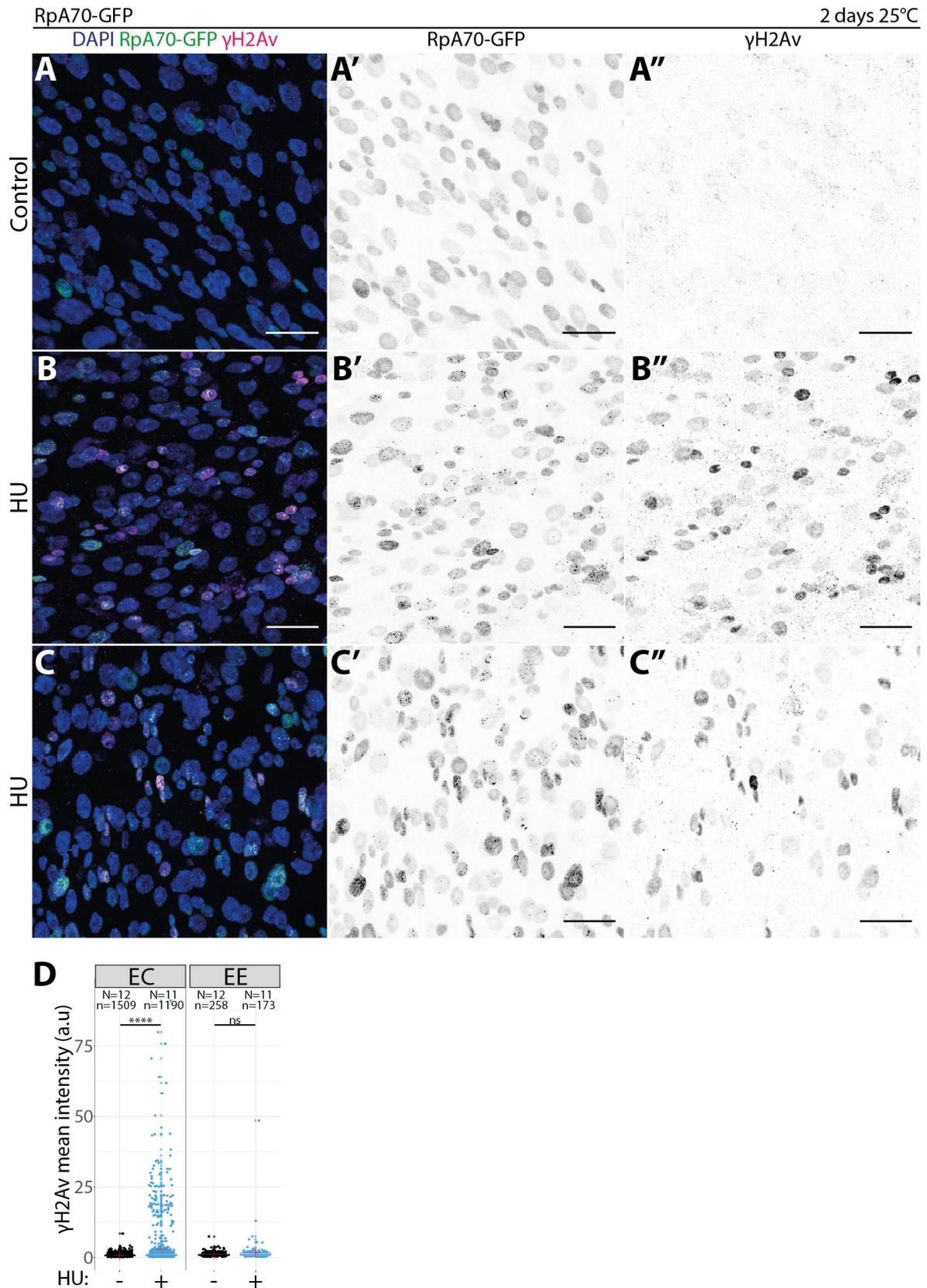
*Author Contribution*

**B.B.** designed the study, performed all the experiments and quantification, and wrote the manuscript. **M.E-H.** performed the Inx2 staining experiment. **A.B.** designed the study and wrote the manuscript.

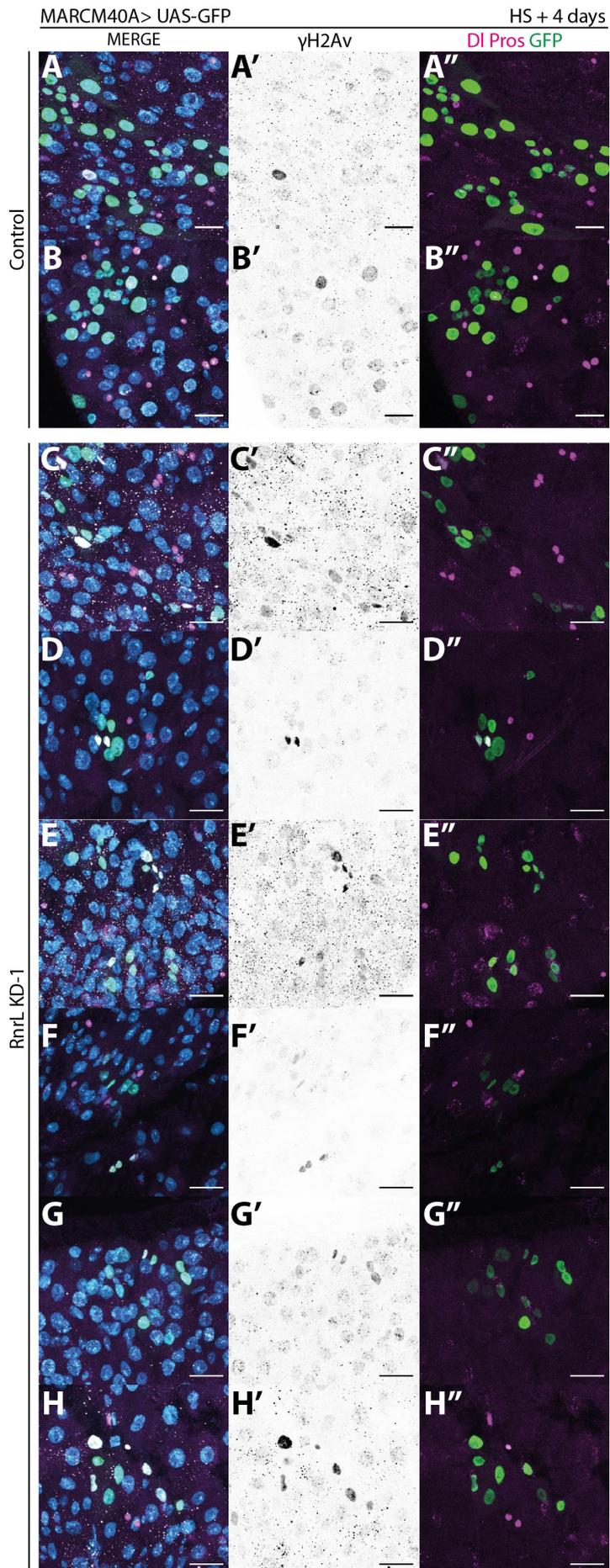
## V. SUPPLEMENTARY FIGURES

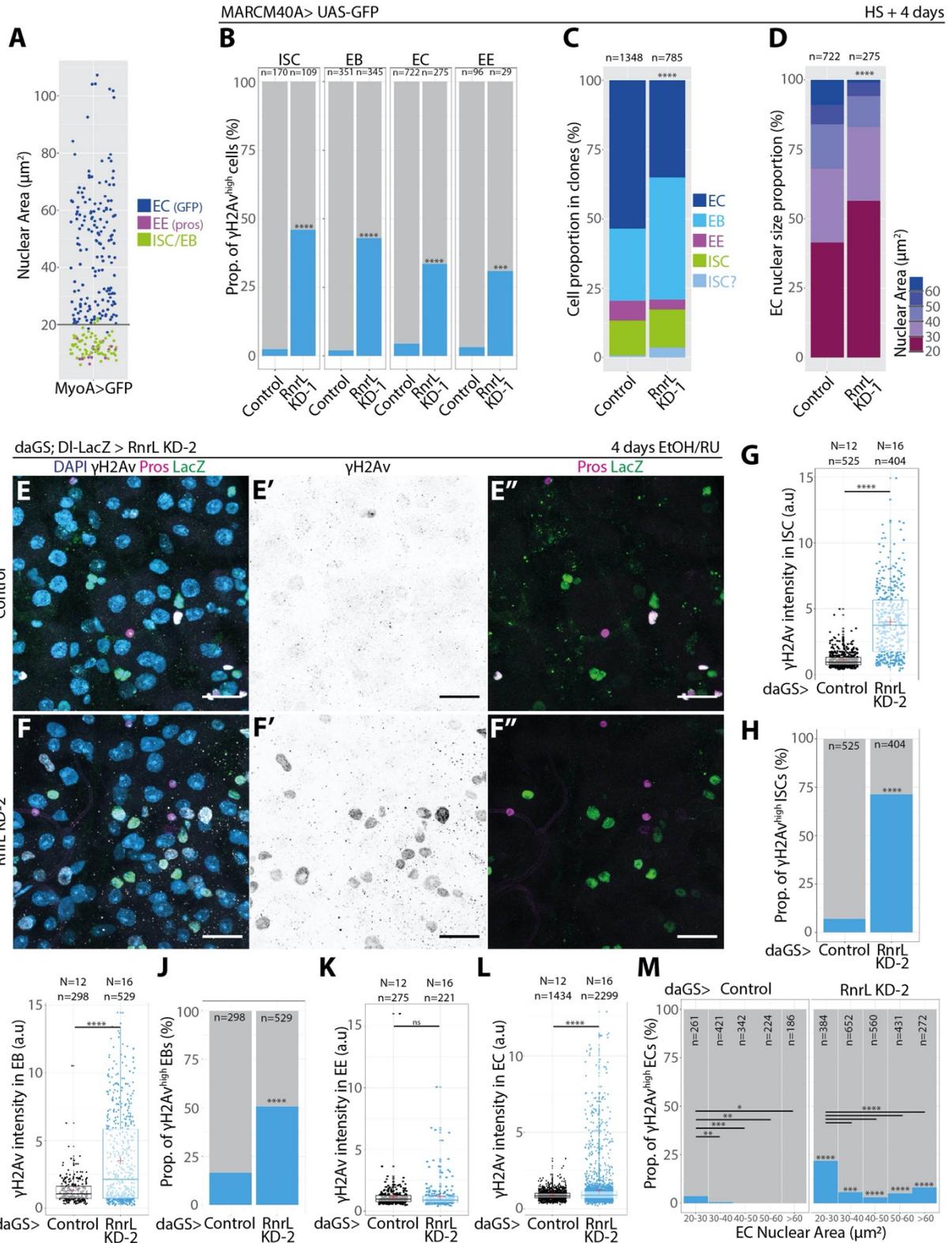


**Figure 1 Supplementary 1. Stem cell-specific induction of replication stress by *RnrL* knockdown.** **A.** PH3<sup>+</sup> cells per gut in female flies fed with control or HU food for 2 or 4 days. Welch's test. **B.** Proportion of  $\gamma$ H2Av<sup>high</sup> cells per gut in HU or *RnrL* RNAi condition, the response to HU is much more variable. **C.**  $\gamma$ H2Av mean intensity in EC and EEs after HU feeding for 2 days. Welch's test. **D-E'''.** *Delta<sup>ts</sup>>RFP*, *RpA70-GFP* female midguts without or with *RnrL* RNAi, stained for DAPI, RFP,  $\gamma$ H2Av, GFP. Scale bar: 20 $\mu$ m. **F.**  $\gamma$ H2Av mean and RpA70 max intensity in each ISC (RFP<sup>+</sup>) nuclei in Control or *RnrL* RNAi condition. Normalized to the mean value in the control. Welch's test. **G.** Proportion of ISC (RFP<sup>+</sup>) with high levels of  $\gamma$ H2Av (>2) or RpA70 max (>3). Fisher's test. **H.**  $\gamma$ H2Av mean intensity in each cell cycle phase after 4 days of *RnrL* RNAi or control in *FUCCI;Delta<sup>ts</sup>*. Normalized to mean of the control. \*: p < 0.05; \*\*: p < 0.01; \*\*\*: p < 0.001; \*\*\*\*: p < 0.0001. N = guts, n = cells. Red cross = mean



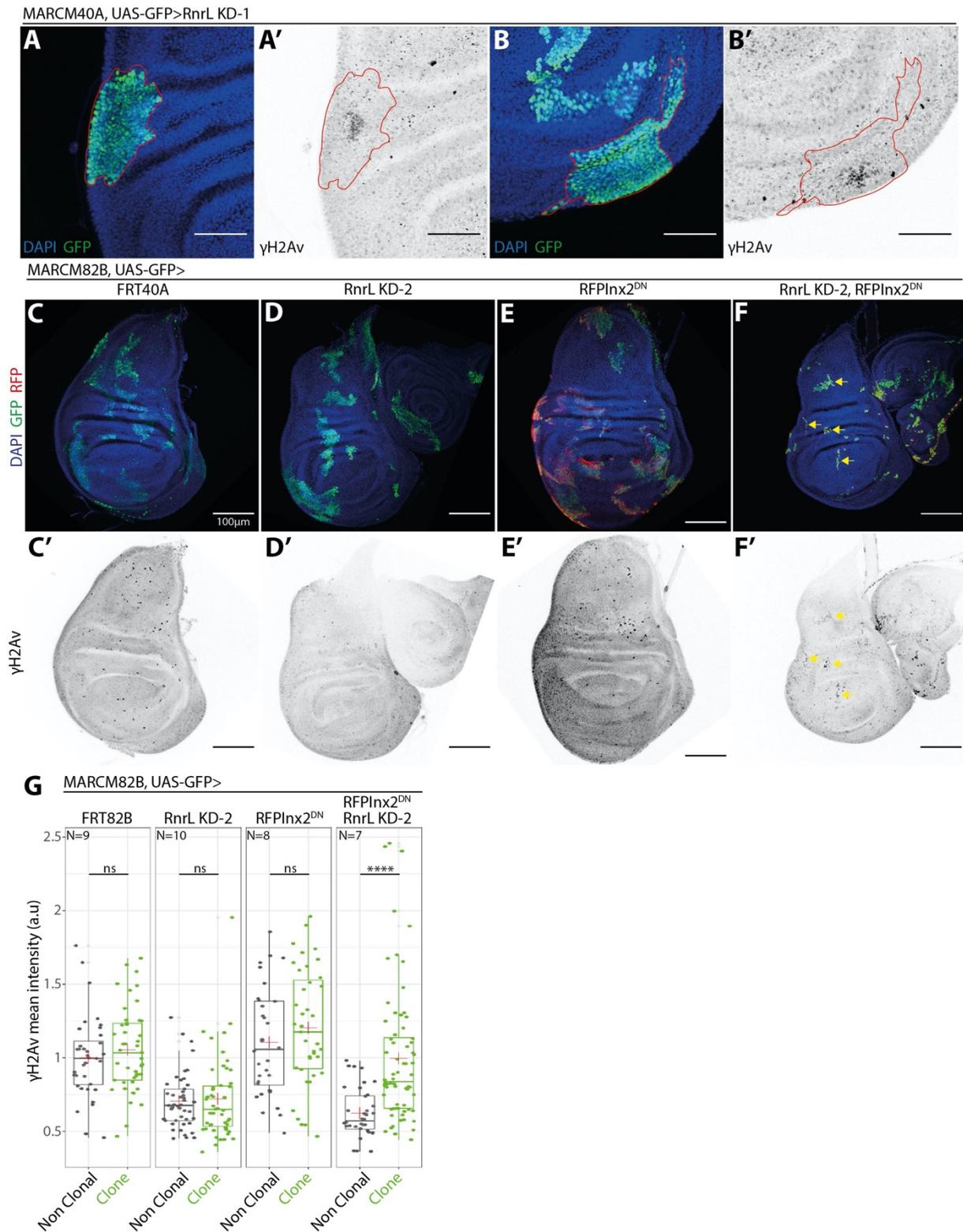
**Figure 1 Supplementary 2. HU is not stem cell specific.** A-C' RpA70-GFP female midguts fed with control (A-A'') or HU-containing food (B-C'') for 2 days and stained for DAPI,  $\gamma$ H2Av, GFP and Pros (not shown). D.  $\gamma$ H2Av mean intensity in EC and EEs after HU feeding for 2 days. Welch's test. EC were determined based on nuclear volume and EE with Pros staining.





**Figure 2 Supplementary 2. *RnrL* knockdown affects stem cell and differentiating progenies.** **A.** *MyoA<sup>ts</sup>>GFP* flies were used to estimate the difference in nuclear area between EB and EC. *MyoAGal4* drives expression in ECs. Most cells with a nucleus  $> 20\mu\text{m}^2$  were GFP+ while cells with nuclei  $< 20\mu\text{m}^2$  were progenitors or EEs marked with Pros. Thus, we used this limit  $20\mu\text{m}^2$  to distinguish EB and EC. **B.** Proportion of cells with high levels of  $\gamma\text{H2Av}$  ( $>2$ ) in MARCM clones with or without *RnrL* RNAi. Fisher's test. **C.** Cell composition of MARCM clones with or without *RnrL* RNAi. Chisquare test. The cells for which the DI staining was not clearly positive or negative were classified as "ISC?". **D.** Proportion of EC with the indicated nuclear size in MARCM clones with or without *RnrL* RNAi. Chisquare test. **E-F''.** *Delta-LacZ; daGS>RnrL RNAi* without (E-E'') or with (F-F'') *RnrL* knockdown induction. The guts were stained for  $\gamma\text{H2Av}$ , the stem cells are marked by the LacZ, and EEs by Pros. Scale bar:  $20\mu\text{m}$ . **G.** Mean  $\gamma\text{H2Av}$  intensity in ISCs, normalized to the Control. Welch's test. **H.** Proportion of ISC with high levels of  $\gamma\text{H2Av}$  ( $>2$ ). Fisher's test. **I.** Mean  $\gamma\text{H2Av}$  intensity in EBs, normalized to the Control. Welch's test. **J.** Proportion of EB with high levels of  $\gamma\text{H2Av}$  ( $>2$ ). Fisher's test. **K.** Mean  $\gamma\text{H2Av}$  intensity in EEs, normalized to the Control. Welch's test. **L.** Mean  $\gamma\text{H2Av}$  intensity in ECs, normalized to the Control. Welch's test. **M.** Proportion of EC with high levels of  $\gamma\text{H2Av}$  ( $>2$ ), depending on their nuclear size. Fisher's test.

\*:  $p < 0.05$ ; \*\*:  $p < 0.01$ ; \*\*\*:  $p < 0.001$ ; \*\*\*\*:  $p < 0.0001$ . N = guts, n = cells. Red cross = mean value



**Figure 3 Supplementary 1. Non-autonomous rescue of DNA damage by gap junctions in the wing disc. A-B'.** *RnrL* RNAi MARCM clones displaying higher levels of DNA damage at the center of the clone. Scale bar: 50 $\mu$ m. **C-F'.** MARCM GFP<sup>+</sup> clones 4 days after induction in the wing disc and stained for  $\gamma$ H2Av, with control clones (C-C'), *RnrL* RNAi (D-D'), *RFP-Inx2<sup>DN</sup>* (E-E') and *RnrLRNAi/RFP-Inx2<sup>DN</sup>* RNAi (I-I'). Scale bar: 100 $\mu$ m. **G.** Quantification of  $\gamma$ H2Av levels inside and outside the clones, normalized to control non-clonal tissue mean intensity. N=Discs. Welch's test, red cross = mean.

\*:  $p < 0.05$ ; \*\*:  $p < 0.01$ ; \*\*\*:  $p < 0.001$ ; \*\*\*\*:  $p < 0.0001$ .

## VI. REFERENCES

1. Blanpain, C., Mohrin, M., Sotiropoulou, P. A. & Passegué, E. DNA–Damage Response in Tissue–Specific and Cancer Stem Cells. *Cell Stem Cell* **8**, 16–29 (2011).
2. Al zouabi, L. & Bardin, A. J. Stem Cell DNA Damage and Genome Mutation in the Context of Aging and Cancer Initiation. *Cold Spring Harb. Perspect. Biol.* a036210 (2020) doi:10.1101/cshperspect.a036210.
3. Weeden, C. E. & Asselin–Labat, M.–L. Mechanisms of DNA damage repair in adult stem cells and implications for cancer formation. *Biochim. Biophys. Acta BBA – Mol. Basis Dis.* **1864**, 89–101 (2018).
4. Hills, S. A. & Diffley, J. F. X. DNA Replication and Oncogene–Induced Replicative Stress. *Curr. Biol.* **24**, R435–R444 (2014).
5. Macheret, M. & Halazonetis, T. D. DNA Replication Stress as a Hallmark of Cancer. *Annu. Rev. Pathol. Mech. Dis.* **10**, 425–448 (2015).
6. Bester, A. C. *et al.* Nucleotide Deficiency Promotes Genomic Instability in Early Stages of Cancer Development. *Cell* **145**, 435–446 (2011).
7. Mannava, S. *et al.* Depletion of Deoxyribonucleotide Pools Is an Endogenous Source of DNA Damage in Cells Undergoing Oncogene–Induced Senescence. *Am. J. Pathol.* **182**, 142–151 (2013).
8. Kotsantis, P., Petermann, E. & Boulton, S. J. Mechanisms of Oncogene–Induced Replication Stress: Jigsaw Falling into Place. *Cancer Discov.* **8**, 537–555 (2018).
9. Primo, L. M. F. & Teixeira, L. K. DNA replication stress: oncogenes in the spotlight. *Genet. Mol. Biol.* (2019) doi:10.1590/1678–4685gmb–2019–0138.
10. Macheret, M. & Halazonetis, T. D. Intragenic origins due to short G1 phases underlie oncogene–induced DNA replication stress. *Nature* **555**, 112–116 (2018).
11. Kotsantis, P. *et al.* Increased global transcription activity as a mechanism of replication stress in cancer. *Nat. Commun.* **7**, (2016).
12. Flach, J. *et al.* Replication stress is a potent driver of functional decline in ageing haematopoietic stem cells. *Nature* **512**, 198–202 (2014).
13. Alvarez, S. *et al.* Replication stress caused by low MCM expression limits fetal erythropoiesis and hematopoietic stem cell functionality. *Nat. Commun.* **6**, (2015).
14. Micchelli, C. A. & Perrimon, N. Evidence that stem cells reside in the adult *Drosophila* midgut epithelium. *Nature* **439**, 475–479 (2006).
15. Ohlstein, B. & Spradling, A. The adult *Drosophila* posterior midgut is maintained by pluripotent stem cells. *Nature* **439**, 470–474 (2006).
16. Miguel–Aliaga, I., Jasper, H. & Lemaitre, B. Anatomy and Physiology of the Digestive Tract of *Drosophila melanogaster*. *Genetics* **210**, 357–396 (2018).
17. Rodríguez–Fernandez, I. A., Tauc, H. M. & Jasper, H. Hallmarks of aging in *Drosophila* intestinal stem cells. *Mech. Ageing Dev.* **190**, 111285 (2020).
18. Siudeja, K. *et al.* Frequent Somatic Mutation in Adult Intestinal Stem Cells Drives Neoplasia and Genetic Mosaicism during Aging. *Cell Stem Cell* **17**, 663–

674 (2015).

19. Siudeja, K., Wurmser, A., Stefanutti, M., Lameiras, S. & Bardin, A. J. Unraveling the features of somatic transposition in the *Drosophila* intestine. *EMBO J.* **19** (2021).
20. Riddiford, N., Siudeja, K., van den Beek, M., Boumard, B. & Bardin, A. J. Evolution and genomic signatures of spontaneous somatic mutation in *Drosophila* intestinal stem cells. *BioRxiv* (2020) doi:10.1101/2020.07.20.188979.
21. Rogakou, E. P., Pilch, D. R., Orr, A. H., Ivanova, V. S. & Bonner, W. M. DNA Double-stranded Breaks Induce Histone H2AX Phosphorylation on Serine 139 \*. *J. Biol. Chem.* **273**, 5858–5868 (1998).
22. Madigan, J. P., Chotkowski, H. L. & Glaser, R. L. DNA double-strand break-induced phosphorylation of *Drosophila* histone variant H2Av helps prevent radiation-induced apoptosis. *Nucleic Acids Res.* **30**, 3698–3705 (2002).
23. Blythe, S. A. & Wieschaus, E. F. Zygotic Genome Activation Triggers the DNA Replication Checkpoint at the Midblastula Transition. *Cell* **160**, 1169–1181 (2015).
24. Zielke, N. *et al.* Fly-FUCCI: A Versatile Tool for Studying Cell Proliferation in Complex Tissues. *Cell Rep.* **7**, 588–598 (2014).
25. Sakaue-Sawano, A. *et al.* Visualizing Spatiotemporal Dynamics of Multicellular Cell-Cycle Progression. *Cell* **132**, 487–498 (2008).
26. Yan, Z. A., Li, X. Z. & Zhou, X. T. The effect of hydroxyurea on the expression of the common fragile site at 3p14. *J. Med. Genet.* **24**, 593–596 (1987).
27. Park, J.-S. *et al.* Requirement of ATR for maintenance of intestinal stem cells in aging *Drosophila*. *Aging* **7**, 307 (2015).
28. Buchon, N., Broderick, N. A., Poidevin, M., Pradervand, S. & Lemaitre, B. *Drosophila* Intestinal Response to Bacterial Infection: Activation of Host Defense and Stem Cell Proliferation. *Cell Host Microbe* **5**, 200–211 (2009).
29. Biteau, B., Hochmuth, C. E. & Jasper, H. JNK Activity in Somatic Stem Cells Causes Loss of Tissue Homeostasis in the Aging *Drosophila* Gut. *Cell Stem Cell* **3**, 442–455 (2008).
30. Feger, G. *et al.* *dpa*, a member of the MCM family, is required for mitotic DNA replication but not endoreplication in *Drosophila*. *EMBO J.* **14**, 5387–5398 (1995).
31. Crevel, G. *et al.* Differential Requirements for MCM Proteins in DNA Replication in *Drosophila* S2 Cells. *PLoS ONE* **2**, e833 (2007).
32. Labib, K., Tercero, J. A. & Diffley, J. F. X. Uninterrupted MCM2–7 Function Required for DNA Replication Fork Progression. *Science* **288**, 1643–1647 (2000).
33. Ekholm-Reed, S. *et al.* Deregulation of cyclin E in human cells interferes with prereplication complex assembly. *J. Cell Biol.* **165**, 789–800 (2004).
34. Beyer, E. C. & Berthoud, V. M. Gap junction gene and protein families: Connexins, innexins, and pannexins. *Biochim. Biophys. Acta BBA – Biomembr.*

1860, 5–8 (2018).

35. Güiza, J., Barría, I., Sáez, J. C. & Vega, J. L. Innexins: Expression, Regulation, and Functions. *Front. Physiol.* **9**, (2018).
36. Ayukawa, T. *et al.* Rescue of Notch signaling in cells incapable of GDP–I–fucose synthesis by gap junction transfer of GDP–I–fucose in *Drosophila*. *Proc. Natl. Acad. Sci.* **109**, 15318–15323 (2012).
37. Spéder, P. & Brand, A. H. Gap Junction Proteins in the Blood–Brain Barrier Control Nutrient–Dependent Reactivation of *Drosophila* Neural Stem Cells. *Dev. Cell* **30**, 309–321 (2014).
38. Dutta, D. *et al.* Regional Cell–Specific Transcriptome Mapping Reveals Regulatory Complexity in the Adult *Drosophila* Midgut. *Cell Rep.* **12**, 346–358 (2015).
39. Xu, C. *et al.* The Septate Junction Protein Tsp2A Restricts Intestinal Stem Cell Activity via Endocytic Regulation of aPKC and Hippo Signaling. *Cell Rep.* **26**, 670–688.e6 (2019).
40. Hung, R.–J. *et al.* A cell atlas of the adult *Drosophila* midgut. *Proc. Natl. Acad. Sci.* **117**, 1514–1523 (2020).
41. Chen, H.–J., Li, Q., Nirala, N. K. & Ip, Y. T. The Snakeskin–Mesh Complex of Smooth Septate Junction Restricts Yorkie to Regulate Intestinal Homeostasis in *Drosophila*. *Stem Cell Rep.* (2020)  
doi:10.1016/j.stemcr.2020.03.021.
42. Halazonetis, T. D., Gorgoulis, V. G. & Bartek, J. An Oncogene–Induced DNA Damage Model for Cancer Development. *Science* **319**, 1352–1355 (2008).
43. Loewenstein, W. R. & Kanno, Y. Intercellular Communication and the Control of Tissue Growth: Lack of Communication between Cancer Cells. *Nature* **209**, 1248–1249 (1966).
44. Aasen, T. *et al.* Connexins in cancer: bridging the gap to the clinic. *Oncogene* **38**, 4429–4451 (2019).
45. Aasen, T., Mesnil, M., Naus, C. C., Lampe, P. D. & Laird, D. W. Gap junctions and cancer: communicating for 50 years. *Nat. Rev. Cancer* **16**, 775–788 (2016).
46. Bardin, A. J., Perdigoto, C. N., Southall, T. D., Brand, A. H. & Schweisguth, F. Transcriptional control of stem cell maintenance in the *Drosophila* intestine. *Development* **137**, 705–714 (2010).
47. Ostrowski, K., Bauer, R. & Hoch, M. The *Drosophila* Innexin7 Gap Junction Protein Is Required for Development of the Embryonic Nervous System. *Cell Commun. Adhes.* **15**, 155–167 (2008).
48. Smendziuk, C. M., Messenberg, A., Vogl, A. W. & Tanentzapf, G. Bi–directional gap junction–mediated soma–germline communication is essential for spermatogenesis. *Development* **142**, 2598–2609 (2015).

# CHAPTER 4 : DISCUSSION AND PERSPECTIVES

In this part, I wanted to discuss freely some of the work I did during my PhD, and some of the limitations I encountered. Interestingly, this work also contributed to raise several additional questions that are currently being further studied in other projects in the lab. In addition, I wanted to discuss some of the questions I am quite interested in, although highly hypothetical at this point, and for which the experimental set up would sometimes rely on the development of techniques not yet available in our model system. However, I believe they represent great perspectives for future work in the next 5 years.

## I. On replication stress and genome instability in ISCs

This PhD work started from the observation that genome complex rearrangements were spontaneously arising in adult intestinal stem cells. We investigated the consequences of replication stress on tissue homeostasis, as mutational signature of replication stress could be found.

### *Evidence for endogenous replication stress in ISCs*

I demonstrated above, that we could artificially induce replication stress in the stem cells using *RnrL* RNAi or hydroxyurea. But, is there replication stress in the stem cells *in vivo*? Since we observed somatic structural variants with a specific signature, we suspected that there was some level of replication stress occurring in ISC. However, can we find evidence of replication stress and is replication stress affecting stem cells *in vivo*?

We reasoned, that we would likely detect replication stress in conditions of forced proliferation of the stem cells. This work is still in progress, but I wanted to list some evidence that could suggest an effect of endogenous replication stress on *Drosophila* intestinal stem cells:

- Detection of somatic mutations with hallmarks of replication stress<sup>1,2</sup> in our previous studies.

- It has been shown by Park et al. that the DNA damage response kinase ATR, specialized in replication fork protection and replication stress signaling, is required for stem cell maintenance in the gut<sup>3</sup>.
- Upon *Ecc15* infection, proliferation increased in the gut and we observed increased levels of DNA damage ( $\gamma$ H2Av) in the stem cells. (L. Alzouabi personal communication)
- Repetitive regenerative responses after *Ecc15* infection drove depletion of the stem pool in Haller et al.<sup>4</sup>.
- *Ecc15* infection promoted a transient increase of DNA repair gene expression<sup>5</sup>, including *WRNexo*. Interestingly, *WRNexo Drosophila* mutants displayed higher somatic genome instability<sup>6,7</sup>, including increased tumor incidence<sup>8</sup>. *WRNexo* was specifically shown to be important for the repair of replication stress-induced DNA damage in *Drosophila*, while dispensable for other types of DNA damage<sup>9</sup>. Thus, *WRNexo* expression in the stem cells might be essential to limit replication stress upon *Ecc15* infection. Consistent with this role, *WRNexo* RNAi limited stem cell proliferation in response to bacterial infection in the midgut<sup>10</sup>

Altogether, it supports the idea that *WRNexo* has a particular role in dealing with endogenous levels of replication stress in the stem cells, more likely upon forced proliferation following tissue injury. To investigate the influence of endogenous replication stress, I started to examine the consequences of *WRNexo* RNAi and *ATR* RNAi on the progenitor's cell-cycle properties and DNA damage upon *Ecc15* infection. This work is still in progress.

### *Cell-specific nucleotide depletion by RnrL RNAi*

We demonstrated that knock-down of *RnrL* leads to high levels of DNA damage in the stem cells, likely responsible for stem cell loss. I think that the levels of replication stress induced in our experiments are very high. It likely does

not correspond to endogenous levels that would be encountered physiologically by the cells in the midgut, but could mimic pathological conditions such as cancer. However, the *RnrL* RNAi system could be fine-tuned to induce lower levels of replication stress, for example by performing the temperature sensitive experiment with the *Delta<sup>ts</sup>* driver at 25°C instead of 29°C. It could give us more insight on how stem cells respond to low levels of replication stress.

Nevertheless, the main advantage of the *RnrL* RNAi is that we can target specifically cell types in different tissues using the UAS/Gal4 system, and we can investigate tissue specificity in the response to replication stress. More interestingly, we can also decipher cell autonomous and non-cell autonomous consequences of nucleotide depletion. This is new, as previous systems investigating the cellular consequences of replication stress were using hydroxyurea or aphidicolin systemically or onto the whole cell culture. Because of the cell-specific targeting of *RnrL* RNAi we could identify the role of gap junctions in buffering nucleotides levels. This could have further implication in our understanding of cell-specific susceptibility to DNA damage in stem cells and cancer.

### *dNTP synthesis pathways*

Rnr produces dNTPs as part of the *de novo* synthesis pathway with the reduction of NTPs to dNTPs. Another pathway, the salvage pathway, can also produce dNTP in the cell by recycling and phosphorylating deoxyribonucleoside (dN), which are the products of DNA degradation. In *Drosophila*, this pathway requires the Deoxyribonucleoside kinase (Dnk)<sup>11</sup>. Knock-down of *dnk* in the ISC or in the wing disc did not induce DNA damage (B. Boumard, M. El Hajj, data not shown). Thus, it suggests that the *de novo* pathway is the main dNTP production pathway. It seems that the loss of Rnr activity cannot be rescued by dNTP production by the salvage pathway. We also tested some other enzymes, like Adenylosuccinate Synthetase (AdSS), Dihydroorotate dehydrogenase (DHOD) or the GMP synthetase burgundy (*bur/GMPS*), implicated in purine or pyrimidine synthesis and so far, we could not see any increase of DNA damage in the wing

disc after knock-down, likely suggesting redundancy at this level of the metabolism.

### *RnrL RNAi and genome stability*

We wanted to determine the effect of induced replication stress on mutation accumulation and characterize the type of mutations it produced. The idea was to first evaluate the effect of *RnrL* RNAi on neoplasia formation in male guts and potentially sequence the neoplasia to identify mutational events induced by replication stress. We used the RU-inducible 5961GeneSwitch driver line to express *RnrL* RNAi in the progenitor cells, aged them and quantify neoplasia frequency in old flies. *RnrL* RNAi expression abolished neoplasia formation in aged male midgut (not shown). Since we observed a strong increase of DNA damage and the depletion of stem cells in our previous experiments, we suspected that the absence of neoplasia was likely due to the loss of stem cells in the tissue. In addition, we could detect significant levels of DNA damage in the stem cells even without induction by RU, suggesting that this driver is leaky and that some expression is induced without RU (not shown).

We could still assess the mutagenicity of *RnrL* RNAi by using a non-leaky driver (daGS), and potentially inducing replication stress only through short pulses. The other possibility would be to reduce the levels of *RnrL* RNAi as suggested before using the *Delta<sup>ts</sup>* driver at 25°C. This could be used to first assess the consequences on neoplasia development in the aging midgut.

## II. Genome fragile sites and chromosomes rearrangements

### *Notch, a chromosome fragile-site?*

As mentioned many times now, previous work in the lab had identified spontaneous inactivation of *Notch* leading to neoplasia formation. The sequencing of more than 50 neoplastic samples from different genetic backgrounds (in part

presented here) confirmed that *Notch* was mutated through deletion and complex rearrangement mostly. However, is *Notch* a mutational hot-spot? Is it more prone to be mutated than most of the other genes? This question is difficult to answer for us for one simple reason, our mutation analysis is biased toward *Notch* phenotype detection. *Notch* mutation drives stem cell and enteroendocrine cell accumulation, which is a very specific and highly recognizable phenotype. In addition, *Notch* is on the X chromosome, a single-hit is sufficient to provoke the gene inactivation. But are there any other frequently mutated genes? From the neoplasia sequencing of different samples, a few genes were found to be mutated in more than one sample, for example *Egfr*, *pointed* and *puckered*. However, those were rare. In addition we have a limited amount of sequencing and lack statistical support for mutational hot-spots.

*Notch* mutation by structural variant resemble the types of events found at chromosome fragile sites of the genome<sup>12</sup>. Could *Notch* be considered a fragile site of the genomes? Fragile sites are genome sequence sensitive to replication stress (see Introduction) and often found in large genes and sequence enriched in AT repeats. *Notch* is quite a large gene in *Drosophila* (~28kb) and it is surrounded by two genes much larger *kirre* (~295kb) and *dunce* (~177kb) which are among the largest genes of the *Drosophila* genome. How this affects the genome organization and the replication program in the intestinal stem cells is not known.

In addition, N. Riddiford in the lab, found that >60% of the breakpoints of structural variants affecting *Notch* fell close to poly(dA:dT) sequences<sup>2</sup>, already known to be associated to replication fork collapse<sup>13</sup>.

Although it seems that *Notch* is very frequently mutated, I think we would need unbiased approaches to truly determine whether *Notch* is a Fragile Site in the ISC genome. It would require a genome-wide profiling of replication stress sites and an unbiased whole-genome sequencing to determine the mutational landscape of ISC independently of *Notch*.

### Genome-wide profiling of replication timing and sites of replication stress

Genome fragile sites are either found at late replicating region, thus far from replication origins<sup>12</sup>, or when replication origins are found in genic region and activated early in S-phase, creating a conflict between replication and transcription<sup>14,15</sup>. The determination of the replication program of ISCs could be interesting to estimate the regions potentially susceptible to replication stress. Replication timing has been investigated in different cell types in *Drosophila*<sup>16-18</sup>. It is known that replication program are highly cell-type specific, for example different mammalian cell lineages showed that 50% of the genome had cell specific replication timing<sup>19,20</sup>. Between wing disc cell and follicular cells in the ovary, the difference was 30%<sup>21</sup>. We could compare *Drosophila* replication timing datasets with the mutations that we identified in the genomes, to extract potential correlations between replication timing and mutagenesis. The determination of the ISC-specific replication program would be possible by whole-genome sequencing of S-phase cells<sup>21,22</sup>, however the required cell number is consequent.

We also investigated genome-wide the potential sites of replication stress. We know that RpA70 binds to ssDNA at sites of replication stress. We designed a RpA70-Dam construct to perform DamID-seq<sup>23</sup> profiling of RpA70 binding sites to the genome (K. Siudeja, B. Boumard, unpublished data). Briefly, Dam is a methylase that methylates adenosine at GATC sites in the genome, thus a fused version of the Dam methylates the genome at site of binding of the protein of interest, here RpA70. Although I believe that this technique can give a good profile of transcription factor or chromatin regulator binding regions<sup>24</sup>, I think that the interpretation for RpA70 binding is difficult, the correlation between the replicates was quite low (~0.6), and we only found 500 sites genome-wide with significant binding in 3 replicates, including one site close to *Notch* (not shown). However, importantly, we are lacking good control in which we artificially induce replication stress, to conclude on this experiment.

Nevertheless, I think that the development of new techniques like and Cut&Tag<sup>25</sup> could be useful to establish the genome-wide binding profile of

replication stress and DNA repair proteins in different conditions. I think it will definitely give us more insight on the mechanisms driving DNA damage and mutation accumulation *in vivo*. However, these techniques are not yet validated in our model system. Importantly, these techniques seem adaptable to a low amount of cellular material.

I believe that in the next few years we could be able to determine more precisely the replication timing in ISC and whether it changes with forced proliferation and aging. We could also be able to map genome-wide the sites of replication stress and DNA damage in the ISC.

### Site-specific induction of replication fork collapse

With the discovery of structural variants in the genome with signature of replication stress, we wondered if we could artificially induce the same type of genome rearrangements to gain insights into how collapsed replication fork are transformed into structural variants responsible for *Notch* inactivation. After reading the work of N. Willis in R. Scully's lab on the Tus/Ter replication blockage system (please refer to the introduction<sup>26</sup>), I hypothesized it would be possible to adapt the system in *Drosophila* to investigate structural variant formation upon fork collapse.

With the help of M. Stefanutti in the lab, we generated 4 fly lines carrying 6xTer sequences inserts at the *Notch* locus. These lines are homozygotes viable. We also made UAS-Tus lines to allow the expression of the fork blocking protein under the control of the Gal4/UAS system. We have 3 different versions of

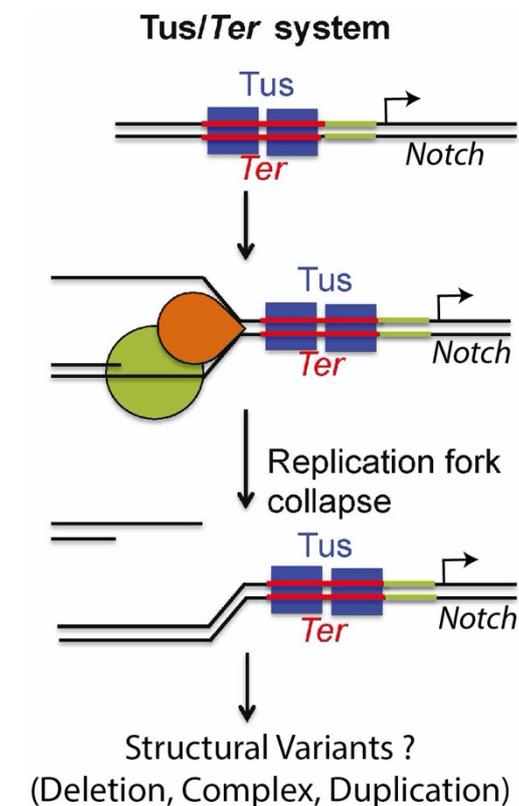


Figure 12. Model of the Tus/Ter system with Ter inserted before Notch

Tus lines, with inactive, active or hyperactive Tus proteins, known to have different efficiency in blocking incoming replication fork.

We were planning to use these lines to induce specific replication fork collapse near *Notch*, by expressing Tus with an intestinal stem cell specific Gal4 driver. The fork collapse would likely result in *Notch* inactivation and neoplastic growth. Thus, we could sequence the neoplasia and identify the type of structural variants generated by replication fork collapse. Unfortunately, we did not have time yet to fully validate the system and get any neoplasia generated by this system. However, when expressed during larval and pupal development, I could observe wing phenotypes potentially reminiscent of clonal *Notch* inactivation, such as notched wing, and extra veins. This is very preliminary, but this phenotype seemed also more penetrant with the hyperactive form of Tus.

I think this system is very interesting to decipher the type of rearrangements induced after replication fork collapse. In the way it was adapted, we have a cell-type specific and controllable timing of induction. The mutations generated can give rise to a detectable neoplasia for sequencing. In addition, if this system is proven to be functional, it could be combined with different DNA repair mutants to assess the contribution of different pathways in structural variant formations. It might help for example to establish the contribution of NHEJ, Pol $\theta$  or Pol32-dependent MMBIR in the complex rearrangements and deletion formation. All these pathways have been implicated in structural variant formation and genome rearrangements in cancer genomes, however the contribution of each in different cancer genome alteration is unclear. The contribution of these pathways could also be established in the same way I presented for Pol $\theta$  before since mutants for Pol32 and NHEJ are homozygotes viables.

### *A complete catalogue of mutations?*

I described in this thesis work the identification of several types of mutations, many of which had already been described by other members of the lab in other experiments: deletions, translocation and complex rearrangements<sup>1,2</sup>;

transposable elements insertions<sup>27</sup>; Loss of Heterozygosity<sup>1,28</sup> (and Alzouabi et al. in prep). I am very proud I could contribute to the extension of the repertoire of structural variants mutations in *Drosophila* intestinal stem cell. Interestingly, data from L. Alzouabi suggest a potential role for replication stress in inducing spontaneous mitotic recombination at transcription–replication conflict sites. All these types of mutations were also largely described in cancer genomes. The sequencing data presented in this thesis also provide the first evidence for two other mutational events, that could be linked to replication stress as well:

- Point mutation and indel accumulation at a specific site, likely due to Break–Induced Replication, and possibly promoted in a Pol $\theta$  mutant context<sup>29</sup>. This will be investigated by L. Alzouabi.
- Loss of the X–chromosome.

A. Suisse in the lab, is now working on understanding the frequency of X–chromosome loss and the mechanism leading to spontaneous aneuploidy. I briefly presented in the introduction mechanisms by which replication stress can lead to aneuploidy, for example, when unreplicated DNA creates chromosomes bridges during cell division. Concerning the X chromosome in *Drosophila*, two interesting facts could link replication stress to the loss of the chromosome: it has only one chromosome arm as its centromere is located at one extremity, and it contains the rDNA locus close to the centromere. The rDNA locus contains a series of repeated sequences coding for the ribosomal RNA, it is highly transcribed and highly susceptible to replication stress because of conflict between transcription and replication<sup>30,31</sup>. Interestingly, previous data from the lab found high frequency of mitotic recombination involving a large segment of the X–chromosome, suggesting frequent DNA DSB close to the centromere<sup>28</sup>. Thus, aneuploidy could be linked to replication stress. There are a lot of assumptions being made here, but I think the coincidences are interesting to be investigated. Of course, other mechanisms of aneuploidy have been described, for example failure in mitotic spindle assembly driving erroneous segregation<sup>32</sup>.

### III. On the functions of gap junctions

We have demonstrated the role of gap junction in buffering replication stress, likely through nucleotide levels, and protects from replication stress in the wing disc. I also characterized gap junction localization in the midgut. *Inx2* and *Inx7* localize at the baso-lateral membrane of enterocytes. However, the function of gap junctions in the intestine is unknown, here we speculate on gap junction functions in the enterocytes and why their absence in intestinal stem cells could be essential to isolate them from mitogenic and apoptotic signals.

#### Gap junction functions

In *Drosophila*, 8 innexin genes have been characterized, with different expression patterns<sup>33</sup>. For example, *Inx7* and *Inx2* were found in the embryonic intestinal epithelium. Different gap junction proteins have been implicated in epithelial embryonic organization<sup>34</sup>, neurogenesis and synapse function, coordinated muscle contraction<sup>35</sup>, stem cell activation in the brain<sup>36</sup> and reproduction<sup>37</sup>. They were not described in the *Drosophila* midgut, although previous reports had identified gap junction structure in the baso-lateral membrane of absorptive cells in the worm *Sagitta setosa*<sup>38</sup>.

Gap junctions have been implicated in intercellular communication via transfer of small metabolites (<1kDa) such as ions (Ca<sup>2+</sup>), reactive oxygen species, amino-acids or various nucleotides like cAMP and GDP-fucose<sup>35,39</sup>. However, we have not characterized gap junction function in the midgut.

#### Gap junctions in the midgut

We showed before that only *Inx2* and *Inx7* are expressed in the adult midgut, and that they are restricted to ECs. I performed preliminary analysis to investigate the role of gap junction in ECs and tissue homeostasis. This is not shown here as it will require future work. I expressed *Inx7* and *Inx2* RNAi in the EC using the *MyoAGal4* driver and did not detect obvious changes in the epithelium structure

although gap junction staining was lost. In addition, EC-specific KD of *Inx7* and *Inx2* did not induce stem cell proliferation, which suggests that loss of gap junction proteins does not perturb epithelium structure and homeostasis. I also noticed a region-specific localization of both *Inx7* and *Inx2* that could be indicative of particular function. Gap junctions were found in the region R2 and R4 of the gut but not in the most anterior (R1) and posterior (R5) regions. The staining was weaker in the acidic copper cell region (R3) of the gut, however a clear demarcation was made between R2–R3 and R3–R4, with a few layers of cells without detectable staining.

In many systems, gap junctions are important for  $\text{Ca}^{2+}$  signaling. Reactive oxygen species also have an important function in response to stress<sup>40</sup>. We think that gap junctions could be important to coordinate EC response to stress and cell death<sup>35</sup> through exchange of ROS,  $\text{Ca}^{2+}$  and potentially other metabolites. Further investigations are required to decipher gap junction function in the midgut.

We also showed before that intestinal stem cells are isolated from the epithelium because they lack gap junctions. This could explain their susceptibility to replication stress induced DNA damage. However, stem cell isolation could also be beneficial to the stem cells. Several mitogenic signals rely on  $\text{Ca}^{2+}$  or ROS signaling in the stem cells<sup>41</sup>. Keeping those pathways separated in the stem cells from the EC could be essential to limit stem cell proliferation or to ensure stem cell maintenance. We are very interested in testing these hypotheses, and we are currently working on trying to add gap junction between stem cells and EC via UAS-*Inx7* and UAS-*Inx2* expression using the progenitor specific *EscargotGal4* driver. We will examine how this expression affect stem cell proliferation and differentiation in the tissue. Importantly, we would like to test whether having gap junctions would protect ISCs from replication stress induced by nucleotide depletion.

### *Gap junction and tumor growth*

We demonstrated the role of gap junctions in limiting DNA damage in the developing wing disc. We now wonder how gap junction could affect tumor growth in this tissue. Several tumor models induced by oncogene expression are used in the wing disc, and mechanisms regulating tumor growth and cell competition with non-tumoral tissue have been investigated there. Increased levels of DNA damage have been described in Ras<sup>V12</sup> expressing cells<sup>42</sup>. Because of increased proliferation, we assume that tumor cells would require higher levels of dNTPs to complete replication. Is the neighboring tissue contributing to tumor growth by providing nucleotides required for DNA synthesis? How would gap junction loss in the tumor impact its growth and levels of DNA damage? We are setting up a new project to address these questions.

### *Gap junctions, tissue specific mutations and cancer*

Finally, we think that gap junction expression could be an indicator of stem cell or cancer cell susceptibility to replication stress and genome instability. In mammals, gap junctions are formed by hexameric channels composed of Connexins (Cx). In humans, 20 connexin isoforms have been identified. Different effects of gap junction on carcinogenesis and tumor progression have been described<sup>43-45</sup>. Some data support the idea that gap junctions act as tumor suppressors in the early steps of tumorigenesis, loss of Cx26 or Cx43 increases breast tumor incidence<sup>44</sup>. However, other data indicate that gap junction could be required for later stages of tumor progression, Cx26 is often highly expressed in lymph node metastasis of breast cancer and lung metastasis of colorectal cancer<sup>46,47</sup>. Cx43 expression is increased in brain metastasis<sup>48</sup>. Several evidences suggest that gap junction promote epithelial to mesenchymal transition and metastasis growth<sup>43</sup>.

Thus, the contribution of gap junction to tumorigenesis is still quite unclear, and how different levels of gap junction expression relates to difference in genome instability has not been investigated. Similarly, whether gap junctions play a role in different stem cell population is understudied.

#### IV. Final statement

This PhD work participated in the description of somatic mutations in intestinal stem cells and in the study of the contribution of replication stress and *Polθ* in mutation formation in our system. We believe that it contributes to the general understanding of the mechanisms driving genome instability in stem cells and cancer cells. This work also provided the basis for future studies on mechanisms of chromosome alterations and has proposed a model for possible tissue-specific susceptibility to replication stress.

#### V. REFERENCES

1. Siudeja, K. *et al.* Frequent Somatic Mutation in Adult Intestinal Stem Cells Drives Neoplasia and Genetic Mosaicism during Aging. *Cell Stem Cell* **17**, 663–674 (2015).
2. Riddiford, N., Siudeja, K., van den Beek, M., Boumard, B. & Bardin, A. J. Evolution and genomic signatures of spontaneous somatic mutation in *Drosophila* intestinal stem cells. *BioRxiv* (2020) doi:10.1101/2020.07.20.188979.
3. Park, J.-S. *et al.* Requirement of ATR for maintenance of intestinal stem cells in aging *Drosophila*. *Aging Cell* **7**, 307 (2015).
4. Haller, S. *et al.* mTORC1 Activation during Repeated Regeneration Impairs Somatic Stem Cell Maintenance. *Cell Stem Cell* **21**, 806–818.e5 (2017).
5. Sousa-Victor, P. *et al.* Piwi Is Required to Limit Exhaustion of Aging Somatic Stem Cells. *Cell Rep.* **20**, 2527–2537 (2017).
6. Saunders, R. D. C., Boubriak, I., Clancy, D. J. & Cox, L. S. Identification and characterization of a *Drosophila* ortholog of WRN exonuclease that is required to maintain genome integrity. *Aging Cell* **7**, 418–425 (2008).
7. Cox, L. S., Clancy, D. J., Boubriak, I. & Saunders, R. D. C. Modeling Werner Syndrome in *Drosophila melanogaster*: Hyper-recombination in Flies Lacking WRN-like Exonuclease. *Ann. N. Y. Acad. Sci.* **1119**, 274–288 (2007).
8. Cassidy, D. *et al.* Evidence for premature aging in a *Drosophila* model of Werner syndrome. *Exp. Gerontol.* **127**, 110733 (2019).
9. Bolterstein, E., Rivero, R., Marquez, M. & McVey, M. The *Drosophila* Werner Exonuclease Participates in an Exonuclease-Independent Response to

- Replication Stress. *Genetics* **197**, 643–652 (2014).
10. Fang, E. F. *et al.* NAD<sup>+</sup> augmentation restores mitophagy and limits accelerated aging in Werner syndrome. *Nat. Commun.* **10**, (2019).
  11. Legent, K. *et al.* In Vivo Analysis of Drosophila Deoxyribonucleoside Kinase Function in Cell Cycle, Cell Survival and Anti–Cancer Drugs Resistance. *Cell Cycle* **5**, 740–749 (2006).
  12. Debatisse, M., Le Tallec, B., Letessier, A., Dutrillaux, B. & Brison, O. Common fragile sites: mechanisms of instability revisited. *Trends Genet.* **28**, 22–32 (2012).
  13. Tubbs, A. *et al.* Dual Roles of Poly(dA:dT) Tracts in Replication Initiation and Fork Collapse. *Cell* **174**, 1127–1142.e19 (2018).
  14. Barlow, J. H. *et al.* Identification of Early Replicating Fragile Sites that Contribute to Genome Instability. *Cell* **152**, 620–632 (2013).
  15. Macheret, M. & Halazonetis, T. D. Intragenic origins due to short G1 phases underlie oncogene–induced DNA replication stress. *Nature* **555**, 112–116 (2018).
  16. Comoglio, F. *et al.* High–Resolution Profiling of Drosophila Replication Start Sites Reveals a DNA Shape and Chromatin Signature of Metazoan Origins. *Cell Rep.* **11**, 821–834 (2015).
  17. Hua, B. L. & Orr–Weaver, T. L. DNA Replication Control During *Drosophila* Development: Insights into the Onset of S Phase, Replication Initiation, and Fork Progression. *Genetics* **207**, 29–47 (2017).
  18. Munden, A. *et al.* Rif1 inhibits replication fork progression and controls DNA copy number in Drosophila. **28** (2018).
  19. Hiratani, I. *et al.* Global Reorganization of Replication Domains During Embryonic Stem Cell Differentiation. *PLOS Biol.* **6**, e245 (2008).
  20. Hiratani, I. *et al.* Genome–wide dynamics of replication timing revealed by in vitro models of mouse embryogenesis. *Genome Res.* **20**, 155–169 (2010).
  21. Armstrong, R. L., Das, S., Hill, C. A., Duronio, R. J. & Nordman, J. T. Rif1 Functions in a Tissue–Specific Manner To Control Replication Timing Through Its PP1–Binding Motif. *Genetics* **215**, 75–87 (2020).
  22. Armstrong, R. L. *et al.* Chromatin conformation and transcriptional activity are permissive regulators of DNA replication initiation in Drosophila. *Genome Res.* **28**, 1688–1700 (2018).
  23. Marshall, O. J., Southall, T. D., Cheetham, S. W. & Brand, A. H. Cell–type–specific profiling of protein–DNA interactions without cell isolation using targeted DamID with next–generation sequencing. *Nat. Protoc.* **11**, 1586–1598 (2016).
  24. Gervais, L. *et al.* Stem Cell Proliferation Is Kept in Check by the Chromatin Regulators Kismet/CHD7/CHD8 and Trr/MLL3/4. *Dev. Cell* **49**, 556–573.e6 (2019).
  25. Kaya–Okur, H. S. *et al.* CUT&Tag for efficient epigenomic profiling of small samples and single cells. *Nat. Commun.* **10**, 1930 (2019).
  26. Willis, N. A. *et al.* BRCA1 controls homologous recombination at Tus/Ter–

- stalled mammalian replication forks. *Nature* **510**, 556–559 (2014).
27. Siudeja, K., Wurmser, A., Stefanutti, M., Lameiras, S. & Bardin, A. J. Unraveling the features of somatic transposition in the *Drosophila* intestine. *EMBO J.* **19** (2021).
  28. Siudeja, K. & Bardin, A. J. Somatic recombination in adult tissues: What is there to learn? *Fly (Austin)* 1–8 (2016) doi:10.1080/19336934.2016.1249073.
  29. Carvajal-García, J., Crown, K. N., Ramsden, D. A. & Sekelsky, J. DNA polymerase theta suppresses mitotic crossing over. *PLOS Genet.* **17**, e1009267 (2021).
  30. Flach, J. *et al.* Replication stress is a potent driver of functional decline in ageing haematopoietic stem cells. *Nature* **512**, 198–202 (2014).
  31. Glover, T. W., Wilson, T. E. & Arlt, M. F. Fragile sites in cancer: more than meets the eye. *Nat. Rev. Cancer* **17**, 489–501 (2017).
  32. Gogondeau, D. *et al.* Aneuploidy causes premature differentiation of neural and intestinal stem cells. *Nat. Commun.* **6**, 8894 (2015).
  33. Curtin, K. D., Zhang, Z. & Wyman, R. J. *Drosophila* has several genes for gap junction proteins. *Gene* **232**, 191–201 (1999).
  34. Bauer, R., Lehmann, C., Martini, J., Eckardt, F. & Hoch, M. Gap Junction Channel Protein Innexin 2 Is Essential for Epithelial Morphogenesis in the *Drosophila* Embryo. *Mol. Biol. Cell* **15**, 2992–3004 (2004).
  35. Güiza, J., Barría, I., Sáez, J. C. & Vega, J. L. Innexins: Expression, Regulation, and Functions. *Front. Physiol.* **9**, (2018).
  36. Spéder, P. & Brand, A. H. Gap Junction Proteins in the Blood–Brain Barrier Control Nutrient–Dependent Reactivation of *Drosophila* Neural Stem Cells. *Dev. Cell* **30**, 309–321 (2014).
  37. Smendziuk, C. M., Messenberg, A., Vogl, A. W. & Tanentzapf, G. Bi-directional gap junction-mediated soma–germline communication is essential for spermatogenesis. *Development* **142**, 2598–2609 (2015).
  38. Duvert, M., Gros, D. & Salat, C. The junctional complex in the intestine of *Sagitta setosa* (Chaetognatha): the paired septate junction. *J. Cell Sci.* **42**, 227–246 (1980).
  39. Sáez, J. C., Berthoud, V. M., Brañes, M. C., Martínez, A. D. & Beyer, E. C. Plasma Membrane Channels Formed by Connexins: Their Regulation and Functions. *Physiol. Rev.* **83**, 1359–1400 (2003).
  40. Patel, P. H. *et al.* Damage sensing by a Nox–Ask1–MKK3–p38 signaling pathway mediates regeneration in the adult *Drosophila* midgut. *Nat. Commun.* **10**, (2019).
  41. Miguel-Aliaga, I., Jasper, H. & Lemaitre, B. Anatomy and Physiology of the Digestive Tract of *Drosophila melanogaster*. *Genetics* **210**, 357–396 (2018).
  42. Murcia, L., Clemente-Ruiz, M., Pierre-Elies, P., Royou, A. & Milán, M. Selective Killing of RAS–Malignant Tissues by Exploiting Oncogene–Induced DNA Damage. *Cell Rep.* **28**, 119–131.e4 (2019).
  43. Nalewajska, M., Marchelek-Myśliwiec, M., Opara-Bajerowicz, M., Dziedziejko, V. & Pawlik, A. Connexins–Therapeutic Targets in Cancers. *Int. J.*

*Mol. Sci.* **21**, (2020).

44. Aasen, T., Mesnil, M., Naus, C. C., Lampe, P. D. & Laird, D. W. Gap junctions and cancer: communicating for 50 years. *Nat. Rev. Cancer* **16**, 775–788 (2016).

45. Aasen, T. *et al.* Connexins in cancer: bridging the gap to the clinic. *Oncogene* **38**, 4429–4451 (2019).

46. Kanczuga-Koda, L. *et al.* Increased expression of connexins 26 and 43 in lymph node metastases of breast cancer. *J. Clin. Pathol.* **59**, 429–433 (2006).

47. Ezumi, K. *et al.* Aberrant expression of connexin 26 is associated with lung metastasis of colorectal cancer. *Clin. Cancer Res. Off. J. Am. Assoc. Cancer Res.* **14**, 677–684 (2008).

48. Chen, Q. *et al.* Carcinoma–astrocyte gap junctions promote brain metastasis by cGAMP transfer. *Nature* **533**, 493–498 (2016).

## INTRODUCTION-REFERENCES

1. Carlson, E. A. Mutation: the History of an Idea from Darwin to Genomics. 1 (2011).
2. Weiner, J. *Time, Love, Memory*. (Faber & Faber, 2000).
3. Villegas, S. N. One hundred years of Drosophila cancer research: no longer in solitude. *Dis. Model. Mech.* **12**, (2019).
4. Vries, H. de. *Die mutationstheorie. Versuche und beobachtungen über die entstehung von arten im pflanzenreich*. (Leipzig, Veit & comp., 1901).
5. Morgan, T. H. Sex Limited Inheritance in Drosophila. *Science* **32**, 120–122 (1910).
6. Nilsson-Ehle, H. Einige Ergebnisse von Kreuzungen bei Hafer und Weizen. – *Nilsson-Ehle Herman Einige Ergeb. Von Kreuzungen Bei Hafer Weizen* – – – (1908).
7. Nilsson-Ehle, H. Kreuzungsuntersuchungen an Hafer und Weizen. *Z. Für Indukt. Abstamm.- Vererbungslehre* **3**, 290–291 (1910).
8. Nilsson-Ehle, H. Mendélisme et acclimatation. in *IVe conférence internationale génétique Paris 1911. Comptes rendus et rapports* (ed. Vilmorin, P. de) 136–157 (Vrin, 1911).
9. Little, C. C. Experimental studies of the inheritance of color in mice. (1913).
10. Morgan, T. H., Sturtevant, A. H., Bridges, C. B. & Muller, H. J. *The Mechanism of Mendelian Heredity*. (H. Holt, 1915).
11. Muller, H. J. Genetic Variability, Twin Hybrids and Constant Hybrids, in a Case of Balanced Lethal Factors. *Genetics* **3**, 422–499 (1918).
12. Muller, H. J. & Altenburg, E. The rate of change of hereditary factors in Drosophila. *Proc. Soc. Exp. Biol. Med.* **17**, 10–14 (1919).
13. Morgan, T. H. The method of inheritance of two sex-limited characters in the same animal. *Proc. Soc. Exp. Biol. Med.* **8**, 17–19 (1910).
14. Morgan, T. H. The Application of the Conception of Pure Lines to Sex-Limited Inheritance and to Sexual Dimorphism. *Am. Nat.* **45**, 65–78 (1911).
15. Morgan, T. H. Random Segregation Versus Coupling in Mendelian Inheritance. *Science* **34**, 384–384 (1911).
16. Sturtevant, A. H. The linear arrangement of six sex-linked factors in Drosophila, as shown by their mode of association. *J. Exp. Zool.* **14**, 43–59 (1913).
17. McClintock, B. A Cytological and Genetical Study of Triploid Maize. *Genetics* **14**, 180–222 (1929).
18. Creighton, H. B. & McClintock, B. The Correlation of Cytological and Genetical Crossing-Over in Zea Mays. A Corroboration. *Proc. Natl. Acad. Sci.* **21**, 148–150 (1935).
19. Castle, W. E. & Nachtsheim, H. Linkage Interrelations of Three Genes for Rex (Short) Coat in the Rabbit. *Proc. Natl. Acad. Sci. U. S. A.* **19**, 1006–1011 (1933).
20. Creighton, H. B. & McClintock, B. A Correlation of Cytological and

- Genetical Crossing-Over in Zea Mays. *Proc. Natl. Acad. Sci.* **17**, 492-497 (1931).
21. Brink, R. A. & Cooper, D. C. A Proof That Crossing Over Involves an Exchange of Segments Between Homologous Chromosomes. *Genetics* **20**, 22-35 (1935).
  22. Stern, C. & Sekiguti, K. Analyse eines Mosaikindividuums bei *Drosophila melanogaster*. *Biol Zentralbl* **51:194-99**, (1931).
  23. Muller, H. J. Variation Due to Change in the Individual Gene. *Am. Nat.* **56**, 32-50 (1922).
  24. Stadler, L. J. The Gene. *Science* **120**, 811-819 (1954).
  25. Bridges, C. B. Non-Disjunction as Proof of the Chromosome Theory of Heredity. *Genetics* **1**, 1-52 (1916).
  26. Dexter, J. S. The Analysis of a Case of Continuous Variation in *Drosophila* by a Study of Its Linkage Relations. *Am. Nat.* **48**, 712-758 (1914).
  27. Muller, H. J. Artificial Transmutation of the Gene. *Science* **66**, 84-87 (1927).
  28. Muller, H. J. THE MEASUREMENT OF GENE MUTATION RATE IN DROSOPHILA, ITS HIGH VARIABILITY, AND ITS DEPENDENCE UPON TEMPERATURE. *Genetics* **13**, 279-357 (1928).
  29. Muller, H. J. The Production of Mutations by X-Rays. *Proc. Natl. Acad. Sci. U. S. A.* **14**, 714-726 (1928).
  30. Mavor, J. W. An Effect of X Rays on the Linkage of Mendelian Characters in the First Chromosome of *Drosophila*. *Genetics* **8**, 355-366 (1923).
  31. Mavor, J. W. On the Elimination of the X-Chromosome from the Egg of *Drosophila Melanogaster* by X-Rays. *Science* **54**, 277-279 (1921).
  32. Mavor, J. W. & Svenson, H. K. An Effect of X Rays on the Linkage of Mendelian Characters in the Second Chromosome of *Drosophila Melanogaster*. *Genetics* **9**, 70-89 (1924).
  33. Mavor, J. W. A comparison of the susceptibility to X-rays of *drosophila melanogaster* at various stages of its life-cycle. *J. Exp. Zool.* **47**, 63-83 (1927).
  34. Muller, H. J. The Regionally Differential Effect of X Rays on Crossing Over in Autosomes of *Drosophila*. *Genetics* **10**, 470-507 (1925).
  35. Stadler, L. J. Mutations in Barley Induced by X-Rays and Radium. *Science* **68**, 186-187 (1928).
  36. Snell, G. D. The Induction by X-Rays of Hereditary Changes in Mice. *Genetics* **20**, 545-567 (1935).
  37. Muller, H. J. & Altenburg, E. The Frequency of Translocations Produced by X-Rays in *Drosophila*. *Genetics* **15**, 283-311 (1930).
  38. Stadler, L. J. On the Genetic Nature of Induced Mutations in Plants. <https://www.abebooks.co.uk/Genetic-Nature-Induced-Mutations-Plants-Stadler/30671330055/bd> (1932).
  39. Weinstein, A. The Production of Mutations and Rearrangements of Genes by X-Rays. *Science* **67**, 376-377 (1928).
  40. Stadler, L. J. The Experimental Modification of Heredity in Crop Plants\*†:

- I. Induced Chromosomal Irregularities‡. *Sci. Agric.* (1931).
41. Bauer, H., Demerec, M. & Kaufmann, B. P. X-Ray Induced Chromosomal Alterations in *Drosophila Melanogaster*. *Genetics* **23**, 610–630 (1938).
  42. Sax, K. Chromosome Aberrations Induced by X-Rays. *Genetics* **23**, 494–516 (1938).
  43. Mohr, O. L. Character Changes Caused by Mutation of an Entire Region of a Chromosome in *Drosophila*. *Genetics* **4**, 275–282 (1919).
  44. Bridges, C. B. Deficiency. *Genetics* **2**, 445–465 (1917).
  45. Sturtevant, A. H. A Case of Rearrangement of Genes in *Drosophila*. *Proc. Natl. Acad. Sci. U. S. A.* **7**, 235–237 (1921).
  46. Muller, H. J. & Mott-Smith, L. M. Evidence That Natural Radioactivity Is Inadequate to Explain the Frequency of “Natural” Mutations. *Proc. Natl. Acad. Sci.* **16**, 277–285 (1930).
  47. Patterson, J. T. & Muller, H. J. Are ‘Progressive’ Mutations Produced by X-Rays? *Genetics* **15**, 495–577 (1930).
  48. Kaplan, W. D. Formaldehyde as a Mutagen in *Drosophila*. *Science* **108**, 43–43 (1948).
  49. Auerbach, C. & Robson, J. M. Action of Mustard Gas on the Bone Marrow. *Nature* **158**, 878–879 (1946).
  50. Auerbach, C. & Robson, J. M. Chemical Production of Mutations. *Nature* **157**, 302–302 (1946).
  51. Loveless, A. Increased Rate of Plaque-type and Host-range Mutation following Treatment of Bacteriophage in vitro with Ethyl Methane Sulphonate. *Nature* **181**, 1212–1213 (1958).
  52. Ward, F. D. The Production of Mutations in *Drosophila Melanogaster* by Irradiation with Alpha-Rays. *Genetics* **20**, 230–249 (1935).
  53. Nagai, M. A. & Locher, G. L. The Production of Mutations in *Drosophila* with Neutron Radiation. *Genetics* **23**, 179–189 (1938).
  54. Auerbach, C., Robson, J. M. & Carr, J. G. The Chemical Production of Mutations. *Science* **105**, 243–247 (1947).
  55. Auerbach, C. Some Recent Results with Chemical Mutagens. *Hereditas* **37**, 1–16 (1951).
  56. Darlington, C. D. & Koller, P. C. The chemical breakage of chromosomes. *Heredity* **1**, 187–221 (1947).
  57. St Johnston, D. The art and design of genetic screens: *Drosophila melanogaster*. *Nat. Rev. Genet.* **3**, 176–188 (2002).
  58. Emerson, R. A. Genetical Studies of Variegated Pericarp in Maize. *Genetics* **2**, 1–35 (1917).
  59. Emerson, R. A. The Frequency of Somatic Mutation in Variegated Pericarp of Maize. *Genetics* **14**, 633–633 (1929).
  60. Hollander, W. F. & Cole, L. J. Somatic Mosaics in the Domestic Pigeon. *Genetics* **25**, 16–40 (1940).
  61. Patterson, J. T. The production of mutations in somatic cells of *Drosophila melanogaster* by means of x-rays. *J. Exp. Zool.* **53**, 327–372

(1929).

62. Patterson, J. T. The Mechanism of Mosaic Formation in *Drosophila*. *Genetics* **18**, 32–52 (1933).
63. Casteel, D. B. Histology of the eyes of x-rayed *Drosophila*. *J. Exp. Zool.* **53**, 373–385 (1929).
64. Auerbach, C. Chemically induced mosaicism in *Drosophila melanogaster*. *Proc. R. Soc. Edinb. Sect. B Biol.* **62**, 211–222 (1946).
65. Stern, C. Somatic Crossing Over and Segregation in *Drosophila Melanogaster*. *Genetics* **21**, 625–730 (1936).
66. McClINTOCK, B. The origin and behavior of mutable loci in maize. *Proc. Natl. Acad. Sci. U. S. A.* **36**, 344–355 (1950).
67. Pianese, G. *Beitrag Zur Histologie Und Aetiologie Des Carcinoms, 1896: Histologische Und Experimentelle Untersuchungen; I/II. Histologische Untersuchungen.*
68. Hansemann, D. Ueber pathologische Mitosen. *Arch. Für Pathol. Anat. Physiol. Für Klin. Med.* **123**, 356–370 (1891).
69. Boveri, T. *Zur Frage der Entstehung maligner Tumoren.* (G. Fischer, 1914).
70. Stark, M. B. An Hereditary Tumor in the Fruit Fly, *Drosophila*. *J. Cancer Res.* **3**, 279–301 (1918).
71. Stark, M. B. & Bridges, C. B. The Linkage Relations of a Benign Tumor in *Drosophila*. *Genetics* **11**, 249–266 (1926).
72. Wilson, I. T. Two New Hereditary Tumors in *Drosophila*. *Genetics* **9**, 343–362 (1924).
73. Stark, M. B. An hereditary tumor. *J. Exp. Zool.* **27**, 509–529 (1919).
74. Scharrer, B. & Lochhead, M. S. Tumors in the Invertebrates: A Review. *Cancer Res.* **10**, 403–419 (1950).
75. Matz, G. Morphogenese tumorale chez les insectes. *J. Insect Physiol.* **11**, 637–639 (1965).
76. Harshbarger, J. C. & Taylor, R. L. Neoplasms of Insects. *Annu. Rev. Entomol.* **13**, 159–190 (1968).
77. Russell, E. S. A comparison of benign and “malignant” tumors in *Drosophila melanogaster*. *J. Exp. Zool.* **84**, 363–385 (1940).
78. Russell, E. S. THE INHERITANCE OF TUMORS IN DROSOPHILA MELANOGASTER, WITH ESPECIAL REFERENCE TO AN ISOGENIC STRAIN OF ST SR TUMOR 36a. *Genetics* **27**, 612–618 (1942).
79. Oftedal, P. The histogenesis of a new tumor in *Drosophila melanogaster*, and a comparison with tumors of five other stocks. *Z. Für Indukt. Abstamm.-Vererbungslehre* **85**, 408–422 (1953).
80. Murray, W. S. & Little, C. C. The Genetics of Mammary Tumor Incidence in Mice. *Genetics* **20**, 466–496 (1935).
81. Little, C. C. The Genetics of Cancer in Mice. *Biol. Rev.* **22**, 315–343 (1947).
82. Gateff, E. Malignant neoplasms of genetic origin in *Drosophila*

- melanogaster. *Science* **200**, 1448–1459 (1978).
83. Gerlach, S. U. & Herranz, H. Genomic instability and cancer: lessons from *Drosophila*. *Open Biol.* **10**, 200060 (2020).
  84. Burdette, W. J. The Significance of Mutation in Relation to the Origin of Tumors: A Review. *Cancer Res.* **15**, 201–226 (1955).
  85. Jeggo, P. A., Pearl, L. H. & Carr, A. M. DNA repair, genome stability and cancer: a historical perspective. *Nat. Rev. Cancer* **16**, 35–42 (2016).
  86. Potter, M. Brief Historical Sketch of Chromosomal Translocations and Tumors. *JNCI Monogr.* **2008**, 2–7 (2008).
  87. Hanahan, D. & Weinberg, R. A. Hallmarks of Cancer: The Next Generation. *Cell* **144**, 646–674 (2011).
  88. Avery, O. T., MacLeod, C. M. & McCarty, M. STUDIES ON THE CHEMICAL NATURE OF THE SUBSTANCE INDUCING TRANSFORMATION OF PNEUMOCOCCAL TYPES. *J. Exp. Med.* **79**, 137–158 (1944).
  89. Hershey, A. D. & Chase, M. INDEPENDENT FUNCTIONS OF VIRAL PROTEIN AND NUCLEIC ACID IN GROWTH OF BACTERIOPHAGE. *J. Gen. Physiol.* **36**, 39–56 (1952).
  90. Chargaff, E. Chemical specificity of nucleic acids and mechanism of their enzymatic degradation. *Experientia* **6**, 201–209 (1950).
  91. Zamenhof, S., Brawerman, G. & Chargaff, E. On the desoxypentose nucleic acids from several microorganisms. *Biochim. Biophys. Acta* **9**, 402–405 (1952).
  92. Wyatt, G. R. THE NUCLEIC ACIDS OF SOME INSECT VIRUSES. *J. Gen. Physiol.* **36**, 201–205 (1952).
  93. Watson, J. D. & Crick, F. H. C. Molecular Structure of Nucleic Acids: A Structure for Deoxyribose Nucleic Acid. *Nature* **171**, 737–738 (1953).
  94. Bautz, E. & Freese, E. On the Mutagenic Effect of Alkylating Agents. *Proc. Natl. Acad. Sci.* **46**, 1585–1594 (1960).
  95. Brookes, P. & Lawley, P. D. The reaction of mustard gas with nucleic acids in vitro and in vivo. *Biochem. J.* **77**, 478–484 (1960).
  96. Brookes, P. & Lawley, P. The reaction of mono- and di-functional alkylating agents with nucleic acids. *Biochem. J.* **80**, 496–503 (1961).
  97. Freese, E. B. TRANSITIONS AND TRANSVERSIONS INDUCED BY DEPURINATING AGENTS\*. *Proc. Natl. Acad. Sci. U. S. A.* **47**, 540–545 (1961).
  98. Strauss, B. S. Specificity of the Mutagenic Action of the Alkylating Agents. *Nature* **191**, 730–731 (1961).
  99. Krieg, D. R. ETHYL METHANESULFONATE-INDUCED REVERSION OF BACTERIOPHAGE T4rII MUTANTS. *Genetics* **48**, 561–580 (1963).
  100. Brenner, S., Barnett, L., Crick, F. H. C. & Orgel, A. The theory of mutagenesis. 4 (1961).
  101. Crick, F. H. C., Barnett, L., Brenner, S. & Watts-Tobin, R. J. General Nature of the Genetic Code for Proteins. *Nature* **192**, 1227–1232 (1961).
  102. Jacob, F. & Monod, J. Genetic regulatory mechanisms in the synthesis of proteins. *J. Mol. Biol.* **3**, 318–356 (1961).

103. Jou, W. M., Haegeman, G., Ysebaert, M. & Fiers, W. Nucleotide Sequence of the Gene Coding for the Bacteriophage MS2 Coat Protein. *Nature* **237**, 82–88 (1972).
104. Wu, R. Nucleotide Sequence Analysis of DNA. *Nature. New Biol.* **236**, 198–200 (1972).
105. Sanger, F., Nicklen, S. & Coulson, A. R. DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci.* **74**, 5463–5467 (1977).
106. Maxam, A. M. & Gilbert, W. A new method for sequencing DNA. *Proc. Natl. Acad. Sci.* **74**, 560–564 (1977).
107. Weeden, C. E. & Asselin-Labat, M.–L. Mechanisms of DNA damage repair in adult stem cells and implications for cancer formation. *Biochim. Biophys. Acta BBA – Mol. Basis Dis.* **1864**, 89–101 (2018).
108. Vitale, I., Manic, G., De Maria, R., Kroemer, G. & Galluzzi, L. DNA Damage in Stem Cells. *Mol. Cell* **66**, 306–319 (2017).
109. Sekelsky, J. DNA Repair in *Drosophila*: Mutagens, Models, and Missing Genes. *Genetics* **205**, 471–490 (2017).
110. Friedberg, E. C. How nucleotide excision repair protects against cancer. *Nat. Rev. Cancer* **1**, 22–33 (2001).
111. Sale, J. E. Translesion DNA Synthesis and Mutagenesis in Eukaryotes. *Cold Spring Harb. Perspect. Biol.* **5**, a012708–a012708 (2013).
112. Kunkel, T. A. & Erie, D. A. Eukaryotic Mismatch Repair in Relation to DNA Replication. *Annu. Rev. Genet.* **49**, 291–313 (2015).
113. Yang, L. *et al.* An enhanced genetic model of colorectal cancer progression history. *Genome Biol.* **20**, (2019).
114. Yamanaka, K., Chatterjee, N., Hemann, M. T. & Walker, G. C. Inhibition of mutagenic translesion synthesis: A possible strategy for improving chemotherapy? *PLoS Genet.* **13**, e1006842 (2017).
115. Scully, R., Panday, A., Elango, R. & Willis, N. A. DNA double-strand break repair-pathway choice in somatic mammalian cells. *Nat. Rev. Mol. Cell Biol.* **20**, 698–714 (2019).
116. Ly, P. & Cleveland, D. W. Rebuilding Chromosomes After Catastrophe: Emerging Mechanisms of Chromothripsis. *Trends Cell Biol.* **27**, 917–930 (2017).
117. Mehta, A., Beach, A. & Haber, J. E. Homology Requirements and Competition between Gene Conversion and Break-Induced Replication during Double-Strand Break Repair. *Mol. Cell* **65**, 515–526.e3 (2017).
118. Ceccaldi, R., Rondinelli, B. & D'Andrea, A. D. Repair Pathway Choices and Consequences at the Double-Strand Break. *Trends Cell Biol.* **26**, 52–64 (2016).
119. Delabaere, L. *et al.* Aging impairs double-strand break repair by homologous recombination in *Drosophila* germ cells. *Aging Cell* (2016) doi:10.1111/accel.12556.
120. McVey, M. & Lee, S. E. MMEJ repair of double-strand breaks (director's cut): deleted sequences and alternative endings. *Trends Genet.* **24**, 529–538 (2008).

121. Roth, D. B. & Wilson, J. H. Nonhomologous recombination in mammalian cells: role for short sequence homologies in the joining reaction. *Mol. Cell. Biol.* **6**, 4295–4304 (1986).
122. Liang, F., Romanienko, P. J., Weaver, D. T., Jeggo, P. A. & Jasin, M. Chromosomal double-strand break repair in Ku80-deficient cells. *Proc. Natl. Acad. Sci.* **93**, 8929–8933 (1996).
123. Canning, S. & Dryja, T. P. Short, direct repeats at the breakpoints of deletions of the retinoblastoma gene. *Proc. Natl. Acad. Sci. U. S. A.* **86**, 5044–5048 (1989).
124. Luzi, P., Rafi, M. A. & Wenger, D. A. Characterization of the large deletion in the GALC gene found in patients with Krabbe disease. *Hum. Mol. Genet.* **4**, 2335–2338 (1995).
125. Magnani, C. *et al.* Short direct repeats at the breakpoints of a novel large deletion in the CFTR gene suggest a likely slipped mispairing mechanism. *Hum. Genet.* **98**, 102–108 (1996).
126. Welcker, A. J., Montigny, J. de, Potier, S. & Souciet, J.–L. Involvement of Very Short DNA Tandem Repeats and the Influence of the RAD52 Gene on the Occurrence of Deletions in *Saccharomyces cerevisiae*. *Genetics* **156**, 549–557 (2000).
127. Yoshida, H. *et al.* Analysis of the joining sequences of the t(15;17) translocation in human acute promyelocytic leukemia: Sequence non-specific recombination between the pml and rara genes within identical short stretches. *Genes. Chromosomes Cancer* **12**, 37–44 (1995).
128. Zhang, J. G., Goldman, J. M. & Cross, N. C. P. Characterization of genomic BCR-ABL breakpoints in chronic myeloid leukaemia by PCR. *Br. J. Haematol.* **90**, 138–146 (1995).
129. Weinstock, D. M., Brunet, E. & Jasin, M. Formation of NHEJ-derived reciprocal chromosomal translocations does not require Ku70. *Nat. Cell Biol.* **9**, 978–981 (2007).
130. Chen, C., Umez, K. & Kolodner, R. D. Chromosomal Rearrangements Occur in *S. cerevisiae* rfa1 Mutator Mutants Due to Mutagenic Lesions Processed by Double-Strand-Break Repair. *Mol. Cell* **2**, 9–22 (1998).
131. Zhang, Y. & Rowley, J. D. Chromatin structural elements and chromosomal translocations in leukemia. *DNA Repair* **5**, 1282–1297 (2006).
132. Mattarucchi, E. *et al.* Microhomologies and interspersed repeat elements at genomic breakpoints in chronic myeloid leukemia. *Genes. Chromosomes Cancer* **47**, 625–632 (2008).
133. Jaüger, U. *et al.* Follicular lymphomas' BCL-2/IgH junctions contain templated nucleotide insertions: novel insights into the mechanism of t(14;18) translocation. *Blood* **95**, 3520–3529 (2000).
134. Zhu, C. *et al.* Unrepaired DNA Breaks in p53-Deficient Cells Lead to Oncogenic Gene Amplification Subsequent to Translocations. *Cell* **109**, 811–821 (2002).
135. Bentley, J., Diggle, C. P., Harnden, P., Knowles, M. A. & Kiltie, A. E.

- DNA double strand break repair in human bladder cancer is error prone and involves microhomology-associated end-joining. *Nucleic Acids Res.* **32**, 5249–5259 (2004).
136. Lieber, M. R., Ma, Y., Pannicke, U. & Schwarz, K. The mechanism of vertebrate nonhomologous DNA end joining and its role in V(D)J recombination. *DNA Repair* **3**, 817–826 (2004).
137. Pan-Hammarström, Q. *et al.* Impact of DNA ligase IV on nonhomologous end joining pathways during class switch recombination in human cells. *J. Exp. Med.* **201**, 189–194 (2005).
138. Yan, C. T. *et al.* IgH class switching and translocations use a robust non-classical end-joining pathway. *Nature* **449**, 478–482 (2007).
139. Jankovic, M., Nussenzweig, A. & Nussenzweig, M. C. Antigen receptor diversification and chromosome translocations. *Nat. Immunol.* **8**, 801–808 (2007).
140. Boulton, S. J. & Jackson, S. P. Identification of a *Saccharomyces Cerevisiae* Ku80 Homologue: Roles in DNA Double Strand Break Rejoining and in Telomeric Maintenance. *Nucleic Acids Res.* **24**, 4639–4648 (1996).
141. Kabotyanski, E. B., Gomelsky, L., Han, J. O., Stamato, T. D. & Roth, D. B. Double-strand break repair in Ku86- and XRCC4-deficient cells. *Nucleic Acids Res.* **26**, 5333–5342 (1998).
142. Difilippantonio, M. J. *et al.* DNA repair protein Ku80 suppresses chromosomal aberrations and malignant transformation. *Nature* **404**, 510–514 (2000).
143. Ferguson, D. O. *et al.* The nonhomologous end-joining pathway of DNA repair is required for genomic stability and the suppression of translocations. *Proc. Natl. Acad. Sci.* **97**, 6630–6633 (2000).
144. Wang, H. *et al.* Biochemical evidence for Ku-independent backup pathways of NHEJ. *Nucleic Acids Res.* **31**, 5377–5388 (2003).
145. Yu, X. & Gabriel, A. Ku-dependent and Ku-independent end-joining pathways lead to chromosomal rearrangements during double-strand break repair in *Saccharomyces cerevisiae*. *Genetics* **163**, 843–856 (2003).
146. Bennardo, N., Cheng, A., Huang, N. & Stark, J. M. Alternative-NHEJ Is a Mechanistically Distinct Pathway of Mammalian Chromosome Break Repair. *PLOS Genet.* **4**, e1000110 (2008).
147. Meyer, D., Fu, B. X. H. & Heyer, W.-D. DNA polymerases  $\delta$  and  $\lambda$  cooperate in repairing double-strand breaks by microhomology-mediated end-joining in *Saccharomyces cerevisiae*. *Proc. Natl. Acad. Sci.* **112**, E6907–E6916 (2015).
148. Villarreal, D. D. *et al.* Microhomology Directs Diverse DNA Break Repair Pathways and Chromosomal Translocations. *PLOS Genet.* **8**, e1003026 (2012).
149. Lee, K. & Lee, S. E. *Saccharomyces cerevisiae* Sae2- and Tel1-Dependent Single-Strand DNA Formation at DNA Break Promotes Microhomology-Mediated End Joining. *Genetics* **176**, 2003–2014 (2007).
150. Payen, C., Koszul, R., Dujon, B. & Fischer, G. Segmental Duplications

- Arise from Pol32-Dependent Repair of Broken Forks through Two Alternative Replication-Based Mechanisms. *PLoS Genet.* **4**, e1000175 (2008).
151. Ma, J.-L., Kim, E. M., Haber, J. E. & Lee, S. E. Yeast Mre11 and Rad1 Proteins Define a Ku-Independent Mechanism To Repair Double-Strand Breaks Lacking Overlapping End Sequences. *Mol. Cell. Biol.* **23**, 8820-8828 (2003).
152. Decottignies, A. Microhomology-Mediated End Joining in Fission Yeast Is Repressed by Pku70 and Relies on Genes Involved in Homologous Recombination. *Genetics* **176**, 1403-1415 (2007).
153. Daley, J. M. & Wilson, T. E. Rejoining of DNA Double-Strand Breaks as a Function of Overhang Length. *Mol. Cell. Biol.* **25**, 896-906 (2005).
154. Deng, S. K., Gibb, B., de Almeida, M. J., Greene, E. C. & Symington, L. S. RPA antagonizes microhomology-mediated repair of DNA double-strand breaks. *Nat. Struct. Mol. Biol.* **21**, 405-412 (2014).
155. Deng, S. K., Chen, H. & Symington, L. S. Replication protein A prevents promiscuous annealing between short sequence homologies: Implications for genome integrity: Prospects & Overviews. *BioEssays* **37**, 305-313 (2015).
156. McVey, M. End-Joining Repair of Double-Strand Breaks in *Drosophila melanogaster* Is Largely DNA Ligase IV Independent. *Genetics* **168**, 2067-2076 (2004).
157. Chan, S. H., Yu, A. M. & McVey, M. Dual Roles for DNA Polymerase Theta in Alternative End-Joining Repair of Double-Strand Breaks in *Drosophila*. *PLoS Genet.* **6**, e1001005 (2010).
158. Yu, A. M. & McVey, M. Synthesis-dependent microhomology-mediated end joining accounts for multiple types of repair junctions. *Nucleic Acids Res.* **38**, 5706-5717 (2010).
159. Boyd, J. B., Golino, M. D., Shaw, K. E. S., Osgood, C. J. & Green, M. M. Third-chromosome mutagen-sensitive mutants of *Drosophila melanogaster*. *Genetics* **97**, 607-623 (1981).
160. Boyd, J. B., Sakaguchi, K. & Harris, P. V. mus308 mutants of *Drosophila* exhibit hypersensitivity to DNA cross-linking agents and are defective in a deoxyribonuclease. *Genetics* **125**, 813-819 (1990).
161. Harris, P. V. *et al.* Molecular cloning of *Drosophila* mus308, a gene involved in DNA cross-link repair with homology to prokaryotic DNA polymerase I genes. *Mol. Cell. Biol.* **16**, 5764-5771 (1996).
162. Pang, M., McConnell, M. & Fisher, P. A. The *Drosophila* mus308 gene product, implicated in tolerance of DNA interstrand crosslinks, is a nuclear protein found in both ovaries and embryos. *DNA Repair* **4**, 971-982 (2005).
163. Arana, M. E., Seki, M., Wood, R. D., Rogozin, I. B. & Kunkel, T. A. Low-fidelity DNA synthesis by human DNA polymerase theta. *Nucleic Acids Res.* **36**, 3847-3856 (2008).
164. Shima, N., Munroe, R. & Schimenti, J. The mouse genomic instability mutation chaos1 is an allele of Polq that exhibits genetic interaction with Atm. *Mol. Cell. Biol.* **24**, 10381-10389 (2004).
165. Muzzini, D. M., Plevani, P., Boulton, S. J., Cassata, G. & Marini, F.

- Caenorhabditis elegans POLQ-1 and HEL-308 function in two distinct DNA interstrand cross-link repair pathways. *DNA Repair* **7**, 941–950 (2008).
166. Thyme, S. B. & Schier, A. F. Polq-Mediated End Joining Is Essential for Surviving DNA Double-Strand Breaks during Early Zebrafish Development. *Cell Rep.* **15**, 707–714 (2016).
167. van Kregten, M. *et al.* T-DNA integration in plants results from polymerase- $\theta$ -mediated DNA repair. *Nat. Plants* **2**, 1–6 (2016).
168. Yousefzadeh, M. J. *et al.* Mechanism of Suppression of Chromosomal Instability by DNA Polymerase POLQ. *PLoS Genet.* **10**, e1004654 (2014).
169. Wyatt, D. W. *et al.* Essential Roles for Polymerase  $\theta$ -Mediated End Joining in the Repair of Chromosome Breaks. *Mol. Cell* **63**, 662–673 (2016).
170. Roerink, S. F., van Schendel, R. & Tijsterman, M. Polymerase theta-mediated end joining of replication-associated DNA breaks in *C. elegans*. *Genome Res.* **24**, 954–962 (2014).
171. Lee-Theilen, M., Matthews, A. J., Kelly, D., Zheng, S. & Chaudhuri, J. CtIP promotes microhomology-mediated alternative end joining during class-switch recombination. *Nat. Struct. Mol. Biol.* **18**, 75–79 (2011).
172. Simsek, D. & Jasin, M. Alternative end-joining is suppressed by the canonical NHEJ component Xrcc4-ligase IV during chromosomal translocation formation. *Nat. Struct. Mol. Biol.* **17**, 410–416 (2010).
173. Feldman, T. *et al.* Recurrent deletions in clonal hematopoiesis are driven by microhomology-mediated end joining. *Nat. Commun.* **12**, 2455 (2021).
174. Kostyrko, K. & Mermod, N. Assays for DNA double-strand break repair by microhomology-based end-joining repair mechanisms. *Nucleic Acids Res.* **44**, e56–e56 (2016).
175. Rodgers, K. & McVey, M. Error-Prone Repair of DNA Double-Strand Breaks: ERROR-PRONE REPAIR OF DNA DSBs. *J. Cell. Physiol.* **231**, 15–24 (2016).
176. Truong, L. N. *et al.* Microhomology-mediated End Joining and Homologous Recombination share the initial end resection step to repair DNA double-strand breaks in mammalian cells. *Proc. Natl. Acad. Sci.* **110**, 7720–7725 (2013).
177. Carvajal-Garcia, J. *et al.* Mechanistic basis for microhomology identification and genome scarring by polymerase theta. *Proc. Natl. Acad. Sci.* 201921791 (2020) doi:10.1073/pnas.1921791117.
178. Kent, T. *et al.* DNA polymerase  $\theta$  specializes in incorporating synthetic expanded-size (xDNA) nucleotides. *Nucleic Acids Res.* gkw721 (2016) doi:10.1093/nar/gkw721.
179. Kent, T., Mateos-Gomez, P. A., Sfeir, A. & Pomerantz, R. T. Polymerase  $\theta$  is a robust terminal transferase that oscillates between three different mechanisms during end-joining. *Elife* **5**, e13740 (2016).
180. Kent, T., Chandramouly, G., McDevitt, S. M., Ozdemir, A. Y. & Pomerantz, R. T. Mechanism of microhomology-mediated end-joining promoted by human DNA polymerase  $\theta$ . *Nat. Struct. Mol. Biol.* **22**, 230–237

(2015).

181. Zahn, K. E., Averill, A. M., Aller, P., Wood, R. D. & Doubl  , S. Human DNA polymerase  $\theta$  grasps the primer terminus to mediate DNA repair. *Nat. Struct. Mol. Biol.* **22**, 304–311 (2015).
182. Black, S., Kashkina, E., Kent, T. & Pomerantz, R. DNA Polymerase  $\theta$ : A Unique Multifunctional End–Joining Machine. *Genes* **7**, 67 (2016).
183. Black, S. J. *et al.* Molecular basis of microhomology–mediated end–joining by purified full–length Pol $\theta$ . *Nat. Commun.* **10**, 4423 (2019).
184. Beagan, K. & McVey, M. Linking DNA polymerase theta structure and function in health and disease. *Cell. Mol. Life Sci.* **73**, 603–615 (2016).
185. Lemee, F. *et al.* DNA polymerase up–regulation is associated with poor survival in breast cancer, perturbs DNA replication, and promotes genetic instability. *Proc. Natl. Acad. Sci.* **107**, 13390–13395 (2010).
186. Sakai, W. *et al.* Secondary mutations as a mechanism of cisplatin resistance in BRCA2 –mutated cancers. *Nature* **451**, 1116–1120 (2008).
187. Edwards, S. L. *et al.* Resistance to therapy caused by intragenic deletion in BRCA2. *Nature* **451**, 1111–1115 (2008).
188. Ceccaldi, R. *et al.* Homologous–recombination–deficient tumours are dependent on Pol $\theta$ –mediated repair. *Nature* **518**, 258–262 (2015).
189. Mateos–Gomez, P. A. *et al.* Mammalian polymerase  $\theta$  promotes alternative NHEJ and suppresses recombination. *Nature* **518**, 254–257 (2015).
190. Mateos–Gomez, P. A. *et al.* The helicase domain of Pol $\theta$  counteracts RPA to promote alt–NHEJ. *Nat. Struct. Mol. Biol.* (2017) doi:10.1038/nsmb.3494.
191. Carvajal–Garcia, J., Crown, K. N., Ramsden, D. A. & Sekelsky, J. DNA polymerase theta suppresses mitotic crossing over. *PLOS Genet.* **17**, e1009267 (2021).
192. Davis, L., Khoo, K. J., Zhang, Y. & Maizels, N. POLQ suppresses interhomolog recombination and loss of heterozygosity at targeted DNA breaks. *Proc. Natl. Acad. Sci.* **117**, 22900–22909 (2020).
193. Kais, Z. *et al.* FANCD2 Maintains Fork Stability in BRCA1/2–Deficient Tumors and Promotes Alternative End–Joining DNA Repair. *Cell Rep.* **15**, 2488–2499 (2016).
194. Alexandrov, L. B. *et al.* Signatures of mutational processes in human cancer. *Nature* **500**, 415–421 (2013).
195. Schrempf, A., Slyskova, J. & Loizou, J. I. Targeting the DNA Repair Enzyme Polymerase  $\theta$  in Cancer Therapy. *Trends Cancer* **7**, 98–111 (2021).
196. Higgins, G. S. & Boulton, S. J. Beyond PARP–POL $\theta$  as an anticancer target. *Science* **359**, 1217–1218 (2018).
197. Inagaki, S. *et al.* Arabidopsis TEBICHI, with Helicase and DNA Polymerase Domains, Is Required for Regulated Cell Division and Differentiation in Meristems. *Plant Cell* **18**, 879–892 (2006).
198. Yoon, J.–H. *et al.* Error–Prone Replication through UV Lesions by DNA Polymerase  $\theta$  Protects against Skin Cancers. *Cell* **176**, 1295–1309.e15 (2019).

199. Koole, W. *et al.* A Polymerase Theta-dependent repair pathway suppresses extensive genomic instability at endogenous G4 DNA sites. *Nat. Commun.* **5**, (2014).
200. Yoshimura, M. *et al.* Vertebrate POLQ and POL $\beta$  Cooperate in Base Excision Repair of Oxidative DNA Damage. *Mol. Cell* **24**, 115–125 (2006).
201. Prasad, R. *et al.* Human DNA polymerase  $\theta$  possesses 5'-dRP lyase activity and functions in single-nucleotide base excision repair in vitro. *Nucleic Acids Res.* **37**, 1868–1877 (2009).
202. Asagoshi, K. *et al.* Single-nucleotide base excision repair DNA polymerase activity in *C. elegans* in the absence of DNA polymerase  $\beta$ . *Nucleic Acids Res.* **40**, 670–681 (2012).
203. Fujimori, H. *et al.* Mismatch repair dependence of replication stress-associated DSB recognition and repair. *Heliyon* **5**, (2019).
204. Feng, W. *et al.* Genetic determinants of cellular addiction to DNA polymerase theta. *Nat. Commun.* **10**, (2019).
205. Czech, B. *et al.* An endogenous small interfering RNA pathway in *Drosophila*. *Nature* **453**, 798–802 (2008).
206. Okamura, K. *et al.* The *Drosophila* hairpin RNA pathway generates endogenous short interfering RNAs. *Nature* **453**, 803–806 (2008).
207. Debatisse, M., Le Tallec, B., Letessier, A., Dutrillaux, B. & Brison, O. Common fragile sites: mechanisms of instability revisited. *Trends Genet.* **28**, 22–32 (2012).
208. Técher, H., Koundrioukoff, S., Nicolas, A. & Debatisse, M. The impact of replication stress on replication dynamics and DNA damage in vertebrate cells. *Nat. Rev. Genet.* **18**, 535–550 (2017).
209. Glover, T. W., Wilson, T. E. & Arlt, M. F. Fragile sites in cancer: more than meets the eye. *Nat. Rev. Cancer* **17**, 489–501 (2017).
210. Macheret, M. & Halazonetis, T. D. DNA Replication Stress as a Hallmark of Cancer. *Annu. Rev. Pathol. Mech. Dis.* **10**, 425–448 (2015).
211. Primo, L. M. F. & Teixeira, L. K. DNA replication stress: oncogenes in the spotlight. *Genet. Mol. Biol.* (2019) doi:10.1590/1678-4685gmb-2019-0138.
212. Glover, T. W., Berger, C., Coyle, J. & Echo, B. DNA polymerase  $\alpha$  inhibition by aphidicolin induces gaps and breaks at common fragile sites in human chromosomes. *Hum. Genet.* **67**, 136–142 (1984).
213. Yan, Z. A., Li, X. Z. & Zhou, X. T. The effect of hydroxyurea on the expression of the common fragile site at 3p14. *J. Med. Genet.* **24**, 593–596 (1987).
214. Gr, S. Heritable fragile sites on human chromosomes. VIII. Preliminary population cytogenetic data on the folic-acid-sensitive fragile sites. *Am. J. Hum. Genet.* **34**, 452–458 (1982).
215. Yunis, J. Fragile sites and predisposition to leukemia and lymphoma. *Cancer Genet. Cytogenet.* **12**, 85–88 (1984).
216. De Braekeleer, M., Smith, B. & Lin, C. C. Fragile sites and structural rearrangements in cancer. *Hum. Genet.* **69**, 112–116 (1985).

217. de Braekeleer, M. Fragile sites and chromosomal structural rearrangements in human leukemia and cancer. *Anticancer Res.* **7**, 417–422 (1987).
218. Le Beau, M. M. Chromosomal fragile sites and cancer-specific rearrangements. *Blood* **67**, 849–858 (1986).
219. Tubbs, A. *et al.* Dual Roles of Poly(dA:dT) Tracts in Replication Initiation and Fork Collapse. *Cell* **174**, 1127–1142.e19 (2018).
220. Rondón, A. G. & Aguilera, A. What causes an RNA–DNA hybrid to compromise genome integrity? *DNA Repair* **81**, 102660 (2019).
221. Minchell, N. E., Keszthelyi, A. & Baxter, J. Cohesin Causes Replicative DNA Damage by Trapping DNA Topological Stress. *Mol. Cell* (2020) doi:10.1016/j.molcel.2020.03.013.
222. Morafraila, E. C. *et al.* Checkpoint inhibition of origin firing prevents DNA topological stress. *Genes Dev.* **33**, 1539–1554 (2019).
223. Canela, A. *et al.* Topoisomerase II–Induced Chromosome Breakage and Translocation Is Determined by Chromosome Architecture and Transcriptional Activity. *Mol. Cell* **75**, 252–266.e8 (2019).
224. Canela, A. *et al.* Genome Organization Drives Chromosome Fragility. *Cell* (2017) doi:10.1016/j.cell.2017.06.034.
225. Gemble, S. *et al.* Topoisomerase II $\alpha$  prevents ultrafine anaphase bridges by two mechanisms. *Open Biol.* **10**, 190259.
226. Boteva, L. *et al.* Common Fragile Sites Are Characterized by Faulty Condensin Loading after Replication Stress. *Cell Rep.* **32**, 108177 (2020).
227. Flach, J. *et al.* Replication stress is a potent driver of functional decline in ageing haematopoietic stem cells. *Nature* **512**, 198–202 (2014).
228. Alvarez, S. *et al.* Replication stress caused by low MCM expression limits fetal erythropoiesis and hematopoietic stem cell functionality. *Nat. Commun.* **6**, (2015).
229. Tufi, R. *et al.* Enhancing nucleotide metabolism protects against mitochondrial dysfunction and neurodegeneration in a PINK1 model of Parkinson’s disease. *Nat. Cell Biol.* **16**, 157–166 (2014).
230. Hämäläinen, R. H. *et al.* Defects in mtDNA replication challenge nuclear genome stability through nucleotide depletion and provide a unifying mechanism for mouse progerias. *Nat. Metab.* **1**, 958–965 (2019).
231. Halazonetis, T. D., Gorgoulis, V. G. & Bartek, J. An Oncogene–Induced DNA Damage Model for Cancer Development. *Science* **319**, 1352–1355 (2008).
232. Guerrero Llobet, S. *et al.* Cyclin E expression is associated with high levels of replication stress in triple–negative breast cancer. *Npj Breast Cancer* **6**, 40 (2020).
233. Kok, Y. P. *et al.* Overexpression of Cyclin E1 or Cdc25A leads to replication stress, mitotic aberrancies, and increased sensitivity to replication checkpoint inhibitors. *Oncogenesis* **9**, 88 (2020).
234. Murga, M. *et al.* Exploiting oncogene–induced replicative stress for the selective killing of Myc–driven tumors. *Nat. Struct. Mol. Biol.* **18**, 1331–1335

- (2011).
235. Macheret, M. & Halazonetis, T. D. Intragenic origins due to short G1 phases underlie oncogene-induced DNA replication stress. *Nature* **555**, 112–116 (2018).
236. Mannava, S. *et al.* Depletion of Deoxyribonucleotide Pools Is an Endogenous Source of DNA Damage in Cells Undergoing Oncogene-Induced Senescence. *Am. J. Pathol.* **182**, 142–151 (2013).
237. Kotsantis, P., Petermann, E. & Boulton, S. J. Mechanisms of Oncogene-Induced Replication Stress: Jigsaw Falling into Place. *Cancer Discov.* **8**, 537–555 (2018).
238. Barlow, J. H. *et al.* Identification of Early Replicating Fragile Sites that Contribute to Genome Instability. *Cell* **152**, 620–632 (2013).
239. Bester, A. C. *et al.* Nucleotide Deficiency Promotes Genomic Instability in Early Stages of Cancer Development. *Cell* **145**, 435–446 (2011).
240. Lamm, N. *et al.* Folate levels modulate oncogene-induced replication stress and tumorigenicity. *EMBO Mol. Med.* **7**, 1138–1152 (2015).
241. Berti, M. & Vindigni, A. Replication stress: getting back on track. *Nat. Struct. Mol. Biol.* **23**, 103–109 (2016).
242. Liu, Y., Nielsen, C. F., Yao, Q. & Hickson, I. D. The origins and processing of ultra fine anaphase DNA bridges. *Curr. Opin. Genet. Dev.* **26**, 1–5 (2014).
243. Ait Saada, A. *et al.* Unprotected Replication Forks Are Converted into Mitotic Sister Chromatid Bridges. *Mol. Cell* **66**, 398–410.e4 (2017).
244. Bhowmick, R., Minocherhomji, S. & Hickson, I. D. RAD52 Facilitates Mitotic DNA Synthesis Following Replication Stress. *Mol. Cell* **64**, 1117–1126 (2016).
245. Minocherhomji, S. *et al.* Replication stress activates DNA repair synthesis in mitosis. *Nature* **528**, 286–290 (2015).
246. Sotiriou, S. K. *et al.* Mammalian RAD52 Functions in Break-Induced Replication Repair of Collapsed DNA Replication Forks. *Mol. Cell* **64**, 1127–1134 (2016).
247. Marco, S. D. *et al.* RECQ5 Helicase Cooperates with MUS81 Endonuclease in Processing Stalled Replication Forks at Common Fragile Sites during Mitosis. *Mol. Cell* **66**, 658–671.e8 (2017).
248. Umbreit, N. T. *et al.* *Mechanisms Generating Cancer Genome Complexity From A Single Cell Division Error*. <http://biorxiv.org/lookup/doi/10.1101/835058> (2019) doi:10.1101/835058.
249. Liu, S. *et al.* Nuclear envelope assembly defects link mitotic errors to chromothripsis. *Nature* **561**, 551–555 (2018).
250. Zhang, C.-Z. *et al.* Chromothripsis from DNA damage in micronuclei. *Nature* **522**, 179–184 (2015).
251. Lambert, S. *et al.* Homologous Recombination Restarts Blocked Replication Forks at the Expense of Genome Rearrangements by Template Exchange. *Mol. Cell* **39**, 346–359 (2010).

252. Lambert, S., Watson, A., Sheedy, D. M., Martin, B. & Carr, A. M. Gross Chromosomal Rearrangements and Elevated Recombination at an Inducible Site-Specific Replication Fork Barrier. *Cell* **121**, 689–702 (2005).
253. Lambert, S., Froget, B. & Carr, A. M. Arrested replication fork processing: Interplay between checkpoints and recombination. *DNA Repair* **6**, 1042–1061 (2007).
254. Willis, N. A. *et al.* BRCA1 controls homologous recombination at Tus/Ter-stalled mammalian replication forks. *Nature* **510**, 556–559 (2014).
255. Hartlerode, A. J., Willis, N. A., Rajendran, A., Manis, J. P. & Scully, R. Complex Breakpoints and Template Switching Associated with Non-canonical Termination of Homologous Recombination in Mammalian Cells. *PLoS Genet* **12**, e1006410 (2016).
256. Willis, N. A. *et al.* Mechanism of tandem duplication formation in BRCA1-mutant cells. *Nature* (2017) doi:10.1038/nature24477.
257. Goullet de Rugy, T. *et al.* Excess Pol $\theta$  functions in response to replicative stress in homologous recombination-proficient cancer cells. *Biol. Open* **5**, 1485–1492 (2016).
258. Alexander, J. L., Beagan, K., Orr-Weaver, T. L. & McVey, M. Multiple mechanisms contribute to double-strand break repair at rereplication forks in *Drosophila* follicle cells. *Proc. Natl. Acad. Sci.* 201617110 (2016) doi:10.1073/pnas.1617110113.
259. Lee, J. A., Carvalho, C. M. B. & Lupski, J. R. A DNA Replication Mechanism for Generating Nonrecurrent Rearrangements Associated with Genomic Disorders. *Cell* **131**, 1235–1247 (2007).
260. Liu, P. *et al.* Chromosome Catastrophes Involve Replication Mechanisms Generating Complex Genomic Rearrangements. *Cell* **146**, 889–903 (2011).
261. Hastings, P. J., Ira, G. & Lupski, J. R. A Microhomology-Mediated Break-Induced Replication Model for the Origin of Human Copy Number Variation. *PLoS Genet.* **5**, e1000327 (2009).
262. Sakofsky, C. J. *et al.* Translesion Polymerases Drive Microhomology-Mediated Break-Induced Replication Leading to Complex Chromosomal Rearrangements. *Mol. Cell* **60**, 860–872 (2015).
263. Costantino, L. *et al.* Break-Induced Replication Repair of Damaged Forks Induces Genomic Duplications in Human Cells. *Science* **343**, 88–91 (2014).
264. Cortés-Ciriano, I. *et al.* Comprehensive analysis of chromothripsis in 2,658 human cancers using whole-genome sequencing. *Nat. Genet.* **52**, 15 (2020).
265. Carvalho, C. M. B. & Lupski, J. R. Mechanisms underlying structural variant formation in genomic disorders. *Nat. Rev. Genet.* **17**, 224–238 (2016).
266. Beck, C. R. *et al.* Megabase Length Hypermutation Accompanies Human Structural Variation at 17p11.2. *Cell* (2019) doi:10.1016/j.cell.2019.01.045.
267. Liu, P. *et al.* An Organismal CNV Mutator Phenotype Restricted to Early Human Development. *Cell* **168**, 830–842.e7 (2017).
268. Ubhi, T. & Brown, G. W. Exploiting DNA Replication Stress for Cancer

- Treatment. *Cancer Res.* **79**, 1730–1739 (2019).
269. Hubackova, S. *et al.* Replication and ribosomal stress induced by targeting pyrimidine synthesis and cellular checkpoints suppress p53-deficient tumors. *Cell Death Dis.* **11**, 110 (2020).
270. Ladds, M. J. G. W. *et al.* A DHODH inhibitor increases p53 synthesis and enhances tumor cell killing by p53 degradation blockage. *Nat. Commun.* **9**, 1107 (2018).
271. Laks, D. R. *et al.* Inhibition of Nucleotide Synthesis Targets Brain Tumor Stem Cells in a Subset of Glioblastoma. *Mol. Cancer Ther.* **15**, 1271–1278 (2016).
272. Pfister, S. X. *et al.* Inhibiting WEE1 Selectively Kills Histone H3K36me3-Deficient Cancers by dNTP Starvation. *Cancer Cell* **28**, 557–568 (2015).
273. Chen, X. *et al.* Cyclin E Overexpression Sensitizes Triple-Negative Breast Cancer to Wee1 Kinase Inhibition. *Clin. Cancer Res.* **24**, 6594–6610 (2018).
274. Nishimura, E. K., Granter, S. R. & Fisher, D. E. Mechanisms of Hair Graying: Incomplete Melanocyte Stem Cell Maintenance in the Niche. **307**, 6 (2005).
275. Inomata, K. *et al.* Genotoxic Stress Abrogates Renewal of Melanocyte Stem Cells by Triggering Their Differentiation. *Cell* **137**, 1088–1099 (2009).
276. Zhang, B. *et al.* Hyperactivation of sympathetic nerves drives depletion of melanocyte stem cells. *Nature* (2020) doi:10.1038/s41586-020-1935-3.
277. Haller, S. *et al.* mTORC1 Activation during Repeated Regeneration Impairs Somatic Stem Cell Maintenance. *Cell Stem Cell* **21**, 806–818.e5 (2017).
278. Barbosa, J. S. *et al.* Live imaging of adult neural stem cell behavior in the intact and injured zebrafish brain. *Science* **348**, 789–793 (2015).
279. Molofsky, A. V. *et al.* Increasing p16 INK4a expression decreases forebrain progenitors and neurogenesis during ageing. *Nature* **443**, 448–452 (2006).
280. Harrison, D. E., Astle, C. M. & Delaittre, J. A. Loss of proliferative capacity in immunohemopoietic stem cells caused by serial transplantation rather than aging. *J. Exp. Med.* **147**, 1526–1531 (1978).
281. Baumgartner, C. *et al.* An ERK-Dependent Feedback Mechanism Prevents Hematopoietic Stem Cell Exhaustion. *Cell Stem Cell* **22**, 879–892.e6 (2018).
282. Singh, S. K. *et al.* Id1 Ablation Protects Hematopoietic Stem Cells from Stress-Induced Exhaustion and Aging. *Cell Stem Cell* **23**, 252–265.e8 (2018).
283. Chakkalakal, J. V., Jones, K. M., Basson, M. A. & Brack, A. S. The aged niche disrupts muscle stem cell quiescence. *Nature* **490**, 355–360 (2012).
284. Al zouabi, L. & Bardin, A. J. Stem Cell DNA Damage and Genome Mutation in the Context of Aging and Cancer Initiation. *Cold Spring Harb. Perspect. Biol.* a036210 (2020) doi:10.1101/cshperspect.a036210.
285. Carlson, M. E. *et al.* Relative roles of TGF- $\beta$ 1 and Wnt in the systemic regulation and aging of satellite cell responses. *Aging Cell* **8**, 676–689 (2009).
286. Conboy, I. M., Conboy, M. J., Smythe, G. M. & Rando, T. A. Notch-

- Mediated Restoration of Regenerative Potential to Aged Muscle. *Science* **302**, 1575–1577 (2003).
287. Choi, S. *et al.* Corticosterone inhibits GAS6 to govern hair follicle stem–cell quiescence. *Nature* **592**, 428–432 (2021).
288. Rossi, D. J. *et al.* Cell intrinsic alterations underlie hematopoietic stem cell aging. *Proc. Natl. Acad. Sci.* **102**, 9194–9199 (2005).
289. Sánchez–Danés, A. *et al.* Defining the clonal dynamics leading to mouse skin tumour initiation. *Nature* **536**, 298–303 (2016).
290. Nassar, D. & Blanpain, C. Cancer Stem Cells: Basic Concepts and Therapeutic Implications. *Annu. Rev. Pathol. Mech. Dis.* **11**, 47–76 (2016).
291. Latil, M. *et al.* Cell–Type–Specific Chromatin States Differentially Prime Squamous Cell Carcinoma Tumor–Initiating Cells for Epithelial to Mesenchymal Transition. *Cell Stem Cell* (2016) doi:10.1016/j.stem.2016.10.018.
292. López–Otín, C., Blasco, M. A., Partridge, L., Serrano, M. & Kroemer, G. The Hallmarks of Aging. *Cell* **153**, 1194–1217 (2013).
293. Booth, L. N. & Brunet, A. The Aging Epigenome. *Mol. Cell* **62**, 728–744 (2016).
294. Brunet, A. & Rando, T. A. Interaction between epigenetic and metabolism in aging stem cells. *Curr. Opin. Cell Biol.* **45**, 1–7 (2017).
295. Chandel, N. S., Jasper, H., Ho, T. T. & Passequé, E. Metabolic regulation of stem cell function in tissue homeostasis and organismal ageing. *Nat. Cell Biol.* **18**, 823–832 (2016).
296. Blanpain, C., Mohrin, M., Sotiropoulou, P. A. & Passequé, E. DNA–Damage Response in Tissue–Specific and Cancer Stem Cells. *Cell Stem Cell* **8**, 16–29 (2011).
297. Matsumura, H. *et al.* Hair follicle aging is driven by transepidermal elimination of stem cells via COL17A1 proteolysis. *Science* **351**, (2016).
298. Rossi, D. J. *et al.* Deficiencies in DNA damage repair limit the function of haematopoietic stem cells with age. *Nature* **447**, 725–729 (2007).
299. Rossi, D. J. *et al.* Hematopoietic Stem Cell Quiescence Attenuates DNA Damage Response and Permits DNA Damage Accumulation During Aging. *Cell Cycle* **6**, 2371–2376 (2007).
300. Mohrin, M. *et al.* Hematopoietic Stem Cell Quiescence Promotes Error–Prone DNA Repair and Mutagenesis. *Cell Stem Cell* **7**, 174–185 (2010).
301. Nijnik, A. *et al.* DNA repair is limiting for haematopoietic stem cells during ageing. *Nature* **447**, 686–690 (2007).
302. Blokzijl, F. *et al.* Tissue–specific mutation accumulation in human adult stem cells during life. *Nature* (2016) doi:10.1038/nature19768.
303. Abascal, F. *et al.* Somatic mutation landscapes at single–molecule resolution. *Nature* 1–6 (2021) doi:10.1038/s41586–021–03477–4.
304. Martincorena, I. *et al.* High burden and pervasive positive selection of somatic mutations in normal human skin. *Science* **348**, 880–886 (2015).
305. Martincorena, I. Somatic mutant clones colonize the human esophagus with age. *14* (2018).

306. Yokoyama, A. *et al.* Age-related remodelling of oesophageal epithelia by mutated cancer drivers. *Nature* **565**, 312–317 (2019).
307. Bolhaqueiro, A. C. F. *et al.* Ongoing chromosomal instability and karyotype evolution in human colorectal cancer organoids. *Nat. Genet.* **51**, 824–834 (2019).
308. Lee-Six, H. *et al.* The landscape of somatic mutation in normal colorectal epithelial cells. *Nature* **574**, 532–537 (2019).
309. Yizhak, K. *et al.* RNA sequence analysis reveals macroscopic somatic clonal expansion across normal tissues. *Science* **364**, eaaw0726 (2019).
310. Moore, L. *et al.* The mutational landscape of normal human endometrial epithelium. *Nature* (2020) doi:10.1038/s41586-020-2214-z.
311. Brazhnik, K. *et al.* Single-cell analysis reveals different age-related somatic mutation profiles between stem and differentiated cells in human liver. *Sci. Adv.* **6**, eaax2659 (2020).
312. Lodato, M. A. *et al.* Aging and neurodegeneration are associated with increased mutations in single human neurons. *Science* eaao4426 (2017).
313. Jaiswal, S. & Ebert, B. L. Clonal hematopoiesis in human aging and disease. *Science* **366**, eaan4673 (2019).
314. Kakiuchi, N. & Ogawa, S. Clonal expansion in non-cancer tissues. *Nat. Rev. Cancer* (2021) doi:10.1038/s41568-021-00335-3.
315. García-Nieto, P. E., Morrison, A. J. & Fraser, H. B. The somatic mutation landscape of the human body. *Genome Biol.* **20**, 298 (2019).
316. Micchelli, C. A. & Perrimon, N. Evidence that stem cells reside in the adult *Drosophila* midgut epithelium. *Nature* **439**, 475–479 (2006).
317. Ohlstein, B. & Spradling, A. The adult *Drosophila* posterior midgut is maintained by pluripotent stem cells. *Nature* **439**, 470–474 (2006).
318. He, L., Si, G., Huang, J., Samuel, A. D. T. & Perrimon, N. Mechanical regulation of stem-cell differentiation by the stretch-activated Piezo channel. *Nature* (2018) doi:10.1038/nature25744.
319. Chen, J. *et al.* Transient Scute activation via a self-stimulatory loop directs enteroendocrine cell pair specification from self-renewing intestinal stem cells. *Nat. Cell Biol.* (2018) doi:10.1038/s41556-017-0020-0.
320. Goulas, S., Conder, R. & Knoblich, J. A. The Par Complex and Integrins Direct Asymmetric Cell Division in Adult Intestinal Stem Cells. *Cell Stem Cell* **11**, 529–540 (2012).
321. Hu, D. J.-K. & Jasper, H. Control of Intestinal Cell Fate by Dynamic Mitotic Spindle Repositioning Influences Epithelial Homeostasis and Longevity. *Cell Rep.* **28**, 2807–2823.e5 (2019).
322. Tian, A., Wang, B. & Jiang, J. Injury-stimulated and self-restrained BMP signaling dynamically regulates stem cell pool size during *Drosophila* midgut regeneration. *Proc. Natl. Acad. Sci.* **114**, E2699–E2708 (2017).
323. Perdigoto, C. N., Schweisguth, F. & Bardin, A. J. Distinct levels of Notch activity for commitment and terminal differentiation of stem cells in the adult fly intestine. *Development* **138**, 4585–4595 (2011).

324. Guo, Z. & Ohlstein, B. Bidirectional Notch signaling regulates *Drosophila* intestinal stem cell multipotency. *Science* **350**, aab0988–aab0988 (2015).
325. Ohlstein, B. & Spradling, A. Multipotent *Drosophila* Intestinal Stem Cells Specify Daughter Cell Fates by Differential Notch Signaling | Science. *Science* **315**, 988–992 (2007).
326. Hung, R.-J. *et al.* A cell atlas of the adult *Drosophila* midgut. *Proc. Natl. Acad. Sci.* **117**, 1514–1523 (2020).
327. Guo, X. *et al.* The Cellular Diversity and Transcription Factor Code of *Drosophila* Enteroendocrine Cells. *Cell Rep.* **29**, 4172–4185.e5 (2019).
328. Guo, X., Huang, H., Yang, Z., Cai, T. & Xi, R. Division of Labor: Roles of Groucho and CtBP in Notch–Mediated Lateral Inhibition that Controls Intestinal Stem Cell Differentiation in *Drosophila*. *Stem Cell Rep.* **12**, 1007–1023 (2019).
329. Bardin, A. J., Perdigoto, C. N., Southall, T. D., Brand, A. H. & Schweisguth, F. Transcriptional control of stem cell maintenance in the *Drosophila* intestine. *Development* **137**, 705–714 (2010).
330. Miguel–Aliaga, I., Jasper, H. & Lemaitre, B. Anatomy and Physiology of the Digestive Tract of *Drosophila melanogaster*. *Genetics* **210**, 357–396 (2018).
331. Khaminets, A., Ronnen–Oron, T., Baldauf, M., Meier, E. & Jasper, H. Cohesin controls intestinal stem cell identity by maintaining association of Escargot with target promoters. *eLife* **9**, (2020).
332. Li, Z. *et al.* A Switch in Tissue Stem Cell Identity Causes Neuroendocrine Tumors in *Drosophila* Gut. *Cell Rep.* **30**, 1724–1734.e4 (2020).
333. Buddika, K. *et al.* Intestinal progenitor P–bodies maintain stem cell identity by suppressing pro–differentiation factors. *BioRxiv* (2020) doi:10.1101/2020.06.27.175398.
334. Xu, C. *et al.* The Septate Junction Protein Tsp2A Restricts Intestinal Stem Cell Activity via Endocytic Regulation of aPKC and Hippo Signaling. *Cell Rep.* **26**, 670–688.e6 (2019).
335. Antonello, Z. A., Reiff, T., Ballesta–Illan, E. & Dominguez, M. Robust intestinal homeostasis relies on cellular plasticity in enteroblasts mediated by miR–8–Escargot switch. *EMBO J.* **34**, 2025–2041 (2015).
336. Korzelius, J. *et al.* The WT1–like transcription factor Klumpfuss maintains lineage commitment of enterocyte progenitors in the *Drosophila* intestine. *Nat. Commun.* **10**, 4123 (2019).
337. Doupé, D. P., Marshall, O. J., Dayton, H., Brand, A. H. & Perrimon, N. *Drosophila* intestinal stem and progenitor cells are major sources and regulators of homeostatic niche signals. *Proc. Natl. Acad. Sci.* 201719169 (2018) doi:10.1073/pnas.1719169115.
338. Jin, Z. *et al.* The *Drosophila* Ortholog of Mammalian Transcription Factor Sox9 Regulates Intestinal Homeostasis and Regeneration at an Appropriate Level. *Cell Rep.* **31**, 107683 (2020).
339. Meng, F. W., Rojas Villa, S. E. & Biteau, B. Sox100B Regulates Progenitor–Specific Gene Expression and Cell Differentiation in the Adult *Drosophila* Intestine. *Stem Cell Rep.* **14**, 226–240 (2020).

340. Rojas Villa, S. E., Meng, F. W. & Biteau, B. zfh2 controls progenitor cell activation and differentiation in the adult *Drosophila* intestinal absorptive lineage. *PLOS Genet.* **15**, e1008553 (2019).
341. Du, G. *et al.* Peroxisome Elevation Induces Stem Cell Differentiation and Intestinal Epithelial Repair. *Dev. Cell* (2020) doi:10.1016/j.devcel.2020.03.002.
342. Zhang, F., Pirooznia, M. & Xu, H. Mitochondria regulate intestinal stem cell proliferation and epithelial homeostasis through FOXO. *Mol. Biol. Cell* **31**, 1538–1549 (2020).
343. Willms, R. J., Zeng, J. & Campbell, S. D. Myt1 Kinase Couples Mitotic Cell Cycle Exit with Differentiation in *Drosophila*. *Cell Rep.* **33**, 108400 (2020).
344. Arnaoutov, A. *et al.* IRBIT Directs Differentiation of Intestinal Stem Cell Progeny to Maintain Tissue Homeostasis. *iScience* 100954 (2020) doi:10.1016/j.isci.2020.100954.
345. Reiff, T. *et al.* Notch and EGFR regulate apoptosis in progenitor cells to ensure gut homeostasis in *Drosophila*. *EMBO J.* (2019) doi:10.15252/embj.2018101346.
346. Arthurton, L., Nahotko, D. A., Alonso, J., Wendler, F. & Baena-Lopez, L. A. Non-apoptotic caspase activation preserves *Drosophila* intestinal progenitor cells in quiescence. *EMBO Rep.* **21**, (2020).
347. Lindblad, J. L., Tare, M., Amcheslavsky, A., Shields, A. & Bergmann, A. Non-apoptotic enteroblast-specific role of the initiator caspase Dronc for development and homeostasis of the *Drosophila* intestine. *BioRxiv* (2020) doi:10.1101/2020.10.25.354431.
348. Erez, N. *et al.* A Non-stop identity complex (NIC) supervises enterocyte identity and protects from premature aging. *eLife* **10**, e62312 (2021).
349. Flint Brodsky, N. *et al.* The transcription factor Hey and nuclear lamins specify and maintain cell identity. *eLife* **8**, e44745 (2019).
350. Gervais, L. & Bardin, A. J. Tissue homeostasis and aging: new insight from the fly intestine. *Curr. Opin. Cell Biol.* **48**, 97–105 (2017).
351. Tauc, H. M. *et al.* Age-related changes in polycomb gene regulation disrupt lineage fidelity in intestinal stem cells. *eLife* **10**, e62250 (2021).
352. Beehler-Evans, R. & Micchelli, C. A. Generation of enteroendocrine cell diversity in midgut stem cell lineages. *Development* **142**, 654–664 (2015).
353. Tracy Cai, X. *et al.* AWD regulates timed activation of BMP signaling in intestinal stem cells to maintain tissue homeostasis. *Nat. Commun.* **10**, (2019).
354. Johansson, J. *et al.* RAL GTPases Drive Intestinal Stem Cell Function and Regeneration through Internalization of WNT Signalosomes. *Cell Stem Cell* **24**, 592–607.e7 (2019).
355. Naszai, M. *et al.* RAL GTPases mediate EGFR/MAPK signalling-driven intestinal stem cell proliferation and tumourigenesis upstream of RAS activation. *BioRxiv* (2020) doi:10.1101/2020.10.07.329607.
356. Zhang, P. *et al.* An SH3PX1-Dependent Endocytosis-Autophagy Network Restrains Intestinal Stem Cell Proliferation by Counteracting EGFR-ERK Signaling. *Dev. Cell* **49**, 574–589.e5 (2019).

357. Gervais, L. *et al.* Stem Cell Proliferation Is Kept in Check by the Chromatin Regulators Kismet/CHD7/CHD8 and Trr/MLL3/4. *Dev. Cell* **49**, 556–573.e6 (2019).
358. Hao, X. *et al.* Lola regulates *Drosophila* adult midgut homeostasis via non-canonical hippo signaling. *Elife* **28** (2020).
359. Andriatsilavo, M. *et al.* Spen limits intestinal stem cell self-renewal. *PLOS Genet.* **14**, e1007773 (2018).
360. Xu, C. *et al.* An in vivo RNAi screen uncovers the role of AdoR signaling and adenosine deaminase in controlling intestinal stem cell activity. *Proc. Natl. Acad. Sci.* 201900103 (2019) doi:10.1073/pnas.1900103117.
361. Morris, O. & Jasper, H. Reactive Oxygen Species in intestinal stem cell metabolism, fate and function. *Free Radic. Biol. Med.* **166**, 140–146 (2021).
362. Patel, P. H. *et al.* Damage sensing by a Nox-Ask1-MKK3-p38 signaling pathway mediates regeneration in the adult *Drosophila* midgut. *Nat. Commun.* **10**, (2019).
363. Mundorf, J., Donohoe, C. D., McClure, C. D., Southall, T. D. & Uhlirova, M. Ets21c Governs Tissue Renewal, Stress Tolerance, and Aging in the *Drosophila* Intestine. *Cell Rep.* **27**, 3019–3033.e5 (2019).
364. Dai, Z. *et al.* *Drosophila* Caliban preserves intestinal homeostasis and lifespan through regulating mitochondrial dynamics and redox state in enterocytes. *PLOS Genet.* **16**, e1009140 (2020).
365. Wisidagama, D. R. & Thummel, C. S. Regulation of *Drosophila* Intestinal Stem Cell Proliferation by Enterocyte Mitochondrial Pyruvate Metabolism. *G3 GenesGenomesGenetics* g3.400633.2019 (2019) doi:10.1534/g3.119.400633.
366. Ngo, S., Liang, J., Su, Y.-H. & O'Brien, L. E. Disruption of EGF Feedback by Intestinal Tumors and Neighboring Cells in *Drosophila*. *Curr. Biol.* (2020) doi:10.1016/j.cub.2020.01.082.
367. Perochon, J., Aughey, G. N., Southall, T. & Cordero, J. B. Dynamic adult tracheal plasticity drives stem cell adaptation to changes in intestinal homeostasis. *BioRxiv* (2021) doi:0.1101/2021.01.10.426079.
368. Amcheslavsky, A., Lindblad, J. L. & Bergmann, A. Transiently “Undead” Enterocytes Mediate Homeostatic Tissue Turnover in the Adult *Drosophila* Midgut. *Cell Rep.* **33**, 108408 (2020).
369. Izumi, Y., Furuse, K. & Furuse, M. Septate junctions regulate gut homeostasis through regulation of stem cell proliferation and enterocyte behavior in *Drosophila*. *J. Cell Sci.* **132**, jcs232108 (2019).
370. Chen, H.-J., Li, Q., Nirala, N. K. & Ip, Y. T. The Snakeskin-Mesh Complex of Smooth Septate Junction Restricts Yorkie to Regulate Intestinal Homeostasis in *Drosophila*. *Stem Cell Rep.* (2020) doi:10.1016/j.stemcr.2020.03.021.
371. Izumi, Y., Furuse, K. & Furuse, M. A novel membrane protein Hoka regulates septate junction organization and stem cell homeostasis in the *Drosophila* gut. *BioRxiv* 75 (2020) doi:10.1101/2020.11.10.376210.
372. Resnik-Docampo, M. *et al.* Tricellular junctions regulate intestinal stem

- cell behaviour to maintain homeostasis. *Nat. Cell Biol.* (2016)  
doi:10.1038/ncb3454.
373. Nagai, H., Tataru, H., Tanaka-Furuhashi, K., Kurata, S. & Yano, T. Homeostatic Regulation of ROS-Triggered Hippo-Yki Pathway via Autophagic Clearance of Ref(2)P/p62 in the *Drosophila* Intestine. *Dev. Cell* **56**, 81–94.e10 (2021).
374. Ameku, T. *et al.* Midgut-derived neuropeptide F controls germline stem cell proliferation in a mating-dependent manner. *PLOS Biol.* **16**, e2005004 (2018).
375. White, M. A., Bonfini, A., Wolfner, M. F. & Buchon, N. *Drosophila melanogaster* sex peptide regulates mated female midgut morphology and physiology. *Proc. Natl. Acad. Sci.* **118**, e2018112118 (2021).
376. Reiff, T. *et al.* Endocrine remodelling of the adult intestine sustains reproduction in *Drosophila*. *eLife* **4**, e06930 (2015).
377. Zipper, L., Jassmann, D., Burgmer, S., Görlich, B. & Reiff, T. Ecdysone steroid hormone remote controls intestinal stem cell fate decisions via the PPAR $\gamma$ -homolog Eip75B in *Drosophila*. *eLife* **9**, e55795 (2020).
378. Ahmed, S. M. H. *et al.* Fitness trade-offs incurred by ovary-to-gut steroid signalling in *Drosophila*. *Nature* **584**, 415–419 (2020).
379. Al Hayek, S. *et al.* Steroid-dependent switch of OvoL/Shavenbaby controls self-renewal versus differentiation of intestinal stem cells. *EMBO J.* (2020)  
doi:10.15252/embj.2019104347.
380. Hadjieconomou, D. *et al.* Enteric neurons increase maternal food intake during reproduction. *Nature* **587**, 455–459 (2020).
381. Scopelliti, A. *et al.* A Neuronal Relay Mediates a Nutrient Responsive Gut/Fat Body Axis Regulating Energy Homeostasis in Adult *Drosophila*. *Cell Metab.* **29**, 269–284.e10 (2019).
382. Zhou, X. *et al.* Physiological and Pathological Regulation of Peripheral Metabolism by Gut-Peptide Hormones in *Drosophila*. *Front. Physiol.* **11**, 577717 (2020).
383. Hudry, B. *et al.* Sex Differences in Intestinal Carbohydrate Metabolism Promote Food Intake and Sperm Maturation. *Cell* **178**, 901–918.e16 (2019).
384. Biteau, B., Hochmuth, C. E. & Jasper, H. JNK Activity in Somatic Stem Cells Causes Loss of Tissue Homeostasis in the Aging *Drosophila* Gut. *Cell Stem Cell* **3**, 442–455 (2008).
385. Guo, L., Karpac, J., Tran, S. L. & Jasper, H. PGRP-SC2 Promotes Gut Immune Homeostasis to Limit Commensal Dysbiosis and Extend Lifespan. *Cell* **156**, 109–122 (2014).
386. Ferguson, M. *et al.* Differential effects of commensal bacteria on progenitor cell adhesion, division symmetry and tumorigenesis in the *Drosophila* intestine. *Development* dev.186106 (2021) doi:10.1242/dev.186106.
387. von Frieling, J. *et al.* A high-fat diet induces a microbiota-dependent increase in stem cell activity in the *Drosophila* intestine. *PLOS Genet.* **16**, e1008789 (2020).

388. Zhou, J. & Boutros, M. JNK-dependent intestinal barrier failure disrupts host-microbe homeostasis during tumorigenesis. *Proc. Natl. Acad. Sci.* **117**, 9401–9412 (2020).
389. Grenier, T. & Leulier, F. How commensal microbes shape the physiology of *Drosophila melanogaster*. *Curr. Opin. Insect Sci.* **41**, 92–99 (2020).
390. Lesperance, D. N. & Broderick, N. A. Microbiomes as modulators of *Drosophila melanogaster* homeostasis and disease. *Curr. Opin. Insect Sci.* **39**, 84–90 (2020).
391. Li, H., Qi, Y. & Jasper, H. Preventing Age-Related Decline of Gut Compartmentalization Limits Microbiota Dysbiosis and Extends Lifespan. *Cell Host Microbe* **19**, 240–253 (2016).
392. Salazar, A. M. *et al.* Intestinal Snakeskin Limits Microbial Dysbiosis during Aging and Promotes Longevity. *iScience* **9**, 229–243 (2018).
393. Resnik-Docampo, M. *et al.* Keeping it tight: The relationship between bacterial dysbiosis, septate junctions, and the intestinal barrier in *Drosophila*. *Fly (Austin)* **12**, 34–40 (2018).
394. Rera, M., Clark, R. I. & Walker, D. W. Intestinal barrier dysfunction links metabolic and inflammatory markers of aging to death in *Drosophila*. *Proc. Natl. Acad. Sci.* **109**, 21528–21533 (2012).
395. Rera, M. *et al.* Modulation of Longevity and Tissue Homeostasis by the *Drosophila* PGC-1 Homolog. *Cell Metab.* **14**, 623–634 (2011).
396. Regan, J. C. *et al.* Sex difference in pathology of the ageing gut mediates the greater response of female lifespan to dietary restriction. *eLife* **5**, e10956 (2016).
397. Rodriguez-Fernandez, I. A., Tauc, H. M. & Jasper, H. Hallmarks of aging in *Drosophila* intestinal stem cells. *Mech. Ageing Dev.* **190**, 111285 (2020).
398. Jeon, H.-J. *et al.* Effect of heterochromatin stability on intestinal stem cell aging in *Drosophila*. *Mech. Ageing Dev.* **173**, 50–60 (2018).
399. Rodriguez-Fernandez, I. A., Qi, Y. & Jasper, H. Loss of a proteostatic checkpoint in intestinal stem cells contributes to age-related epithelial dysfunction. *Nat. Commun.* **10**, (2019).
400. Koehler, C. L., Perkins, G. A., Ellisman, M. H. & Jones, D. L. Pink1 and Parkin regulate *Drosophila* intestinal stem cell proliferation during stress and aging. *J. Cell Biol.* **216**, 2315–2327 (2017).
401. Schell, J. C. *et al.* Control of intestinal stem cell function and proliferation by mitochondrial pyruvate metabolism. *Nat. Cell Biol.* (2017) doi:10.1038/ncb3593.
402. Bensard, C. L. *et al.* Regulation of Tumor Initiation by the Mitochondrial Pyruvate Carrier. *Cell Metab.* (2019) doi:10.1016/j.cmet.2019.11.002.
403. Morris, O., Deng, H., Tam, C. & Jasper, H. Warburg-like Metabolic Reprogramming in Aging Intestinal Stem Cells Contributes to Tissue Hyperplasia. *Cell Rep.* **33**, 108423 (2020).
404. Park, J.-S. *et al.* Age- and oxidative stress-induced DNA damage in *Drosophila* intestinal stem cells as marked by Gamma-H2AX. *Exp. Gerontol.* **47**,

- 401–405 (2012).
405. Sousa–Victor, P. *et al.* Piwi Is Required to Limit Exhaustion of Aging Somatic Stem Cells. *Cell Rep.* **20**, 2527–2537 (2017).
406. Park, J.–S. *et al.* Requirement of ATR for maintenance of intestinal stem cells in aging *Drosophila*. *Aging* **7**, 307 (2015).
407. Park, J.–S., Jeon, H.–J., Pyo, J.–H., Kim, Y.–S. & Yoo, M.–A. Deficiency in DNA damage response of enterocytes accelerates intestinal stem cell aging in *Drosophila*. *Aging* **10**, 322–338 (2018).
408. Na, H., Akan, I., Abramowitz, L. K. & Hanover, J. A. Nutrient–Driven O–GlcNAcylation Controls DNA Damage Repair Signaling and Stem/Progenitor Cell Homeostasis. *Cell Rep.* **31**, 107632 (2020).
409. Siudeja, K. *et al.* Frequent Somatic Mutation in Adult Intestinal Stem Cells Drives Neoplasia and Genetic Mosaicism during Aging. *Cell Stem Cell* **17**, 663–674 (2015).
410. Riddiford, N., Siudeja, K., van den Beek, M., Boumard, B. & Bardin, A. J. Evolution and genomic signatures of spontaneous somatic mutation in *Drosophila* intestinal stem cells. *BioRxiv* (2020)  
doi:10.1101/2020.07.20.188979.
411. Siudeja, K., Wurmser, A., Stefanutti, M., Lameiras, S. & Bardin, A. J. Unraveling the features of somatic transposition in the *Drosophila* intestine. *EMBO J.* **19** (2021).
412. Koyama, L. A. J. *et al.* Bellymount enables longitudinal, intravital imaging of abdominal organs and the gut microbiota in adult *Drosophila*. *PLOS Biol.* **18**, e3000567 (2020).
413. Martin, J. L. *et al.* Long–term live imaging of the *Drosophila* adult midgut reveals real–time dynamics of division, differentiation and loss. *eLife* **7**, e36248 (2018).

## RÉSUMÉ

---

L'intégrité du génome dans les cellules souches des tissus à longue durée de vie est essentielle pour maintenir la fonction des tissus et prévenir l'apparition de cancers. Comment les cellules souches font face aux lésions de l'ADN détermine leur taux de mutation, leur susceptibilité au cancer, et leur déclin fonctionnel avec l'âge. Ma thèse visait à comprendre les facteurs causant des dommages à l'ADN, et les mécanismes agissant dans les cellules souches adultes pour prévenir les mutations spontanées. Il est proposé que le stress de réplication provoqué par l'appauvrissement en nucléotides augmente l'instabilité du génome dans les cellules cancéreuses, car il peut conduire à des réarrangements du génome. Cependant, on ne sait pas encore dans quelle mesure le stress de réplication est responsable des altérations génomiques dans les cellules souches adultes. De même, les mécanismes de réparation de l'ADN qui agissent pour réparer les lésions de l'ADN et qui sont liés aux réarrangements du génome ne sont pas complètement compris. L'intestin de la drosophile est un bon modèle pour étudier les cellules souches et l'homéostasie des tissus et répondre à ces questions.

L'équipe de A. Bardin a précédemment montré que différents types de mutations somatiques apparaissent dans les cellules souches intestinales au cours du vieillissement. Notamment, les intestins mâles âgés développent fréquemment des néoplasies spontanées dues à l'inactivation du gène suppresseur de tumeur Notch par de petites ou grandes délétions ou des réarrangements génomiques plus complexes. Au cours de ma thèse, j'ai d'abord examiné comment les cellules souches âgées font face aux dommages à l'ADN. J'ai également contribué à caractériser la formation spontanée de néoplasies dans différents contextes génétiques. La microdissection et l'extraction de l'ADN des néoplasies pour le séquençage ultérieur du génome (avec K. Siudeja) et le développement de pipelines bioinformatiques (par N. Riddiford), ont permis la caractérisation des mutations somatiques des cellules souches intestinales. Comme de nombreux variants structuraux identifiés suggéraient un mécanisme de réparation de l'ADN reposant sur les microhomologies, j'ai étudié le rôle de Pol Theta - une polymérase impliquée dans la réparation de l'ADN médiée par les microhomologies - dans les mutations somatiques et la formation de néoplasies dans les cellules souches intestinales.

Plusieurs éléments du séquençage ont suggéré que le stress de réplication pourrait être une cause importante de mutation somatique dans les cellules souches. J'ai donc examiné les conséquences du stress de réplication induit par la déplétion de nucléotides sur les cellules souches. J'ai développé une approche cellule-spécifique pour induire le stress de réplication par le knockdown de RnrL (une enzyme essentielle pour la production de dNTPs). Dans l'intestin, le knockdown de RnrL a induit une accumulation de dommages à l'ADN dans les cellules souches en phase S et des défauts de prolifération, conduisant probablement à la perte de cellules souches. Cependant, j'ai observé que le knockdown de RnrL dans le disque d'aile en développement induisait rarement des dommages à l'ADN. Ensuite, j'ai démontré le rôle des jonctions GAP dans la limitation du stress de réplication, probablement en ajustant les niveaux de nucléotides entre cellules adjacentes. Enfin, j'ai montré que les jonctions GAP sont uniquement localisées entre les entérocytes dans l'intestin de drosophile mais pas dans les cellules souches. Les différences d'expression des jonctions GAP pourraient expliquer les différences de sensibilité des tissus et des cellules au stress de réplication et aux dommages à l'ADN.

## MOTS CLÉS

---

Stress de réplication, Dommages à l'ADN, Cellules Souches, Drosophile

## ABSTRACT

---

Genome integrity in long-lived tissue stem cells is essential to maintain tissue function and prevent cancer initiation. How stem cells cope with DNA lesions determines their mutation rate, susceptibility to cancer, and likely age-related functional decline. My thesis aimed to understand what are the DNA damage causing factors, and what mechanisms are acting in adult stem cells to prevent spontaneous mutation. Replication stress driven by nucleotide depletion is proposed to increase genome instability in cancer cells, as it may lead to genome rearrangements. However, to what extent replication stress is responsible for genomic alterations in adult stem cells remains unclear. Likewise, the erroneous DNA repair mechanisms acting to repair DNA damage and linked to genome rearrangements are not completely understood. The *Drosophila* intestine is a good model system to study stem cells and tissue homeostasis and address these questions.

The Bardin lab previously showed that different types of somatic mutations are arising in aging intestinal stem cells. Notably, wild type aged male guts frequently develop spontaneous neoplasias due to the inactivation of the tumor suppressor gene Notch by small or large deletions or more complex genomic rearrangements. During my thesis, I first examined how aged stem cells cope with DNA damage. I also helped to characterize the spontaneous formation of neoplasias in different genetic background. Microdissection and DNA extraction of neoplasias for subsequent whole-genome sequencing (work with K. Siudeja) and the development of bioinformatic pipelines (by N. Riddiford), allowed the characterization of the somatic mutations of intestinal stem cells. Since many structural variants identified suggested a DNA repair mechanism relying on microhomologies, I investigated the role of Pol Theta, a polymerase involved in microhomology mediated DNA repair, in somatic mutations and neoplasia formation in intestinal stem cells.

Several evidence from the sequencing suggested that replication stress might be an important cause of somatic mutation in the stem cells. Therefore, I examined the consequences of nucleotide depletion-induced replication stress on stem cells. Importantly, I developed a cell specific approach to induce replication stress by RnrL knockdown (an enzyme essential for the production of dNTPs). In the gut, the knockdown of RnrL induced DNA damage accumulation in S phase stem cells and proliferation defects, likely leading to stem cell loss. However, I observed that knockdown of RnrL in the developing wing disc rarely induced DNA damage cell autonomously. Following this, I demonstrated a role for GAP junctions in non-cell autonomously limiting replication stress, likely by buffering nucleotide levels between adjacent cells. Finally, I showed that GAP junctions are only localized between enterocytes in the midgut but not in stem cells. Differences in GAP junctions expression could explain tissue and cell specific sensitivity to replication stress and DNA damage.

## KEYWORDS

---

Replication Stress, DNA damage, Stem Cells, *Drosophila*