THÈSE

En vue de l'obtention du

DOCTORAT DE L'UNIVERSITÉ DE TOULOUSE

Délivré par l'Université Toulouse 3 - Paul Sabatier

Présentée et soutenue par

Douaa MOUSSALEM

Le 12 octobre 2020

Titre : Functional characterization of alternative splice variants of the *Drosophila* GATA transcription factor Serpent containing either one or two zinc finger domains.

Ecole doctorale : **BSB - Biologie, Santé, Biotechnologies**

Spécialité : **GENETIQUE MOLECULAIRE**

Unité de recherche :

CBD - Centre de Biologie du Développement

Thèse dirigée par

Marc HAENLIN et Dani OSMAN

Jury

M. David CRIBBS, Président

Mme Annarita MICCIO, Rapporteure

Mme Kyra CAMPBELL, Rapporteure

M. Samir MERABET, Rapporteur

M. Marc HAENLIN, Directeur de thèse

M. Dani OSMAN, Co-directeur de thèse

“I dedicate this work to my parents, who have always supported and encouraged me to achieve my goals. This work is also dedicated to my adorable sisters: Amani, Mariam and Aya, who are an inexhaustible source of motivation, love and strength.”

Acknowledgments

First, I would like to thank Drs. Annarita Miccio, Kyra Campbell, Samir Merabet and David Cribbs, for accepting to be part of my thesis jury. I also profoundly thank my supervisors Dani Osman and Marc Haenlin, the CBD (Centre de Biologie du développement), its director Fabienne Pituello, the doctoral school BSB (Biologie-Santé-Biotechnologies) and its director Philippe Valet, for giving me the opportunity to do my PhD.

I would like to express my sincere gratitude to Dr. Dani Osman, who returned to Lebanon from Switzerland five years ago. When Dani arrived at the Lebanese University with his tiny flies, he sparked my curiosity to know everything about what was happening in this wonderful organism. I was impressed by his positivity, his intelligence and his supportive soul. One of the best days of my life was when Dani accepted to be my thesis co-director. He subsequently helped me to get my Lebanese fellowship, and then he introduced me *via* skype to Marc and Lucas. But he didn't stop there, his support and care are endless.

I would now like to thank Dr. Marc Haenlin for accepting me in his group and for helping me get third and fourth year fellowships from the UPS (Université Paul Sabatier). Thank you, Marc, for helping me to get settled in France once I had arrived, and thank you for being around every day. Thank you for your support and guidance throughout my work, the scientific discussions that we had, and also for correcting my thesis manuscript.

I am also grateful to the greatest researcher and the queen of Genetics, Dr. Vanessa Gobert. Thank you, Vanessa, for always being so active and dynamic. Thank you for staying around and for listening to all of the scientific questions from the PhD students. You definitely transmitted your passion for science. And thank you for reading and correcting my thesis manuscript and for helping me every single time I needed help.

I also thank the GATA factor enthusiast Benoit Auge for his technical help in the lab, the awesome colleague Thomas Genais for being so kind and active throughout my entire thesis period, Delhia Gigan for her energy and sense of humour, and the graduated PhD student Aichun Chen for sending us every year greeting cards from China.

I thank our neighbours, the *Zebrafish* people, Patrick's group members, for the joy and the sense of humour that they spread. Special thanks to Patrick Blader and Pascale Dufourcq who have reminded me several times that I should take a break and not only think about work. And thanks to Pascale again, for giving me a ride home each time we finished late in lab.

I am also thankful to Alain Vincent and Michele Crozatier's group members for the useful discussions that we had during all these years, for sharing materials, and also for the group spirit. Special thanks to Caroline Monod for reading and correcting my thesis manuscript and for always being a gentle person.

My deep gratitude to François Payre for his kindness and care. My sincere thanks to Philippe Valenti for all the useful advice that he gave me. I absolutely thank the entire Payre group for being so friendly and active.

Of course, I thank Alice Davy and Cathy Soula's group. Thanks also to the dearest funniest colleague Mohammad Fawal, and the kindest David Ohayon.

I would like to express my deep gratitude to all the members of the CBD, including the groups of Fabienne Pituello, Bertrand Bénazéraf, Muriel Boube, Andreas Merdes and Eric Théveneau.

Special thanks to Gwenaelle Juffroy and Marie Pelletier and everyone working in the secretariat, you are the sweetest and the kindest. Also, thanks to Julien Favier and Amélie Destenabes for fly medium preparation. Thanks again to Julien for always finding solutions to my requests. Thanks to the service informatique, Laurent Sanchou and Franck Cressent,

especially for providing me with a PC. I used it throughout my entire thesis period. And thank you Franck, for always helping me.

Finally, thank you all for being part of my four years of thesis. Seeing you was a constant source of pleasure and encouragement. Your kind words and your passion for science helped generate an uplifting environment every day. I will always remember the CBD outings, which were full of learning and funny times. I wish all of you, years of exciting research, interesting results, health and happiness.

Content

Abstract

GATA transcription factors play crucial roles in various developmental processes throughout bilaterian animals. In mammals, six GATA factors are present and they play essential functions in different tissues such as the blood, the gut, the liver and the gonads. GATA proteins have two highly conserved domains, the N-terminal and the C-terminal zinc fingers. The C-terminal finger recognizes GATA DNA-binding consensus motif, while the N-terminal finger stabilizes fixation to DNA palindromic sequences and allows their interaction with cofactors of the Friend Of GATA (FOG) family. GATA zinc finger mutations are associated to a vast panel of human diseases whose severity depends on the affected GATA gene and on the position of the mutation in the zinc fingers.

Numerous studies have demonstrated the high level of molecular and functional similarities existing between flies and humans. *Drosophila melanogaster* has five GATA factors containing either one or two zinc fingers, whose sequences are almost identical to those of the canonical zinc fingers of vertebrates. Among them, the *Drosophila* GATA factor Serpent (Srp) is required for the formation of blood cells, gut and fat body as well as during oogenesis. In all these tissues, two isoforms of Srp are generated through an alternative splicing event giving rise to proteins containing either both zinc fingers (N- and C-terminal, hence the name of this isoform: SrpNC) or only the C-terminal zinc finger (SrpC). In a previous work, our team has shown that SrpC and SrpNC activate some genes in a similar manner but also they regulate others differently. Moreover, interaction between SrpNC and its cofactor FOG, U-shaped, is responsible for some but not all aspects of the distinct activities of SrpC and SrpNC. The purpose of this study is to provide a deep genetic investigation of possible differential functional roles of Srp isoforms

during *Drosophila* development. Using CRISPR/Cas9 technology, we generated two mutant fly lines deleted either of SrpC or of SrpNC. In addition, we produced a third mutant fly line in which we specifically introduced into the N-terminal zinc finger of Srp a single point mutation that alters its interaction with U-shaped.

Analysis of these mutants revealed that both isoforms regulate redundantly the transcription of a common set of genes during gut development as well as few genes involved during early hematopoiesis. Surprisingly, flies devoid of SrpNC (isoform containing two-zinc fingers as the mammalian GATA factors) are viable, showing that this isoform is dispensable for most of the developmental processes controlled by Srp. Nonetheless, SrpNC appears to be specifically required in the maintenance of blood cell homeostasis and for fly fertility. Furthermore, disrupting the interaction of Srp and its FOG cofactor U-shaped is equivalent to the complete loss of the isoform SrpNC, showing that SrpNC forms a complex with U-shaped to ensure its functions. In contrast, our genetic approach unraveled that SrpC isoform is essential for viability and fat body development, suggesting that this isoform regulate different developmental programs compared to SrpNC. Altogether, our results reveal a greater functional flexibility played by the GATA zinc fingers to fulfil their many roles throughout development. Also, this work illustrates that, like genome duplication in vertebrates, alternative splicing provides an efficient strategy to promote subfunctionalization and generate GATA functional diversity in invertebrates.

Key words

GATA, Friend of GATA, *Drosophila*, Zinc finger, Development, Hematopoiesis.

Titre/Résumé

Caractérisation fonctionnelle des variants d'épissage alternatifs du facteur de transcription GATA de la drosophile Serpent contenant un ou deux domaines de doigt de zinc

Les facteurs de transcription GATA jouent un rôle crucial dans divers processus de développement chez les animaux bilatéraux. Chez les mammifères, six facteurs GATA sont présents et ils jouent des rôles essentiels dans différents tissus tels que le sang, l'intestin, le foie et les gonades. Les protéines GATA possèdent deux domaines hautement conservés, les doigts de zinc N-terminal et C-terminal. Le doigt C-terminal reconnaît le motif consensus de liaison à l'ADN GATA, tandis que le doigt N-terminal stabilise la fixation aux séquences palindromiques d'ADN et permet leur interaction avec les cofacteurs de la famille Friend Of GATA (FOG). Les mutations des doigts de zinc GATA sont associées à un vaste éventail de maladies humaines dont la gravité dépend du gène GATA affecté et de la position de la mutation dans les doigts de zinc.

De nombreuses études ont démontré le haut niveau de similarités moléculaires et fonctionnelles existant entre les mouches et les humains. La drosophile possède cinq facteurs GATA contenant un ou deux doigts de zinc, dont les séquences sont presque identiques à celles des doigts de zinc canoniques des vertébrés. Parmi eux, le facteur GATA de la drosophile Serpent (Srp) est requis pour la formation des cellules sanguines, de l’intestin et du corps gras ainsi que pendant l'ovogenèse. Dans tous ces tissus, deux isoformes de Srp sont générées par un événement d'épissage alternatif donnant naissance à des protéines contenant soit les deux doigts de zinc

(N- et C-terminal, d'où le nom de cette isoforme: SrpNC) ou uniquement le doigt de zinc C-terminal (SrpC). Dans un travail précédent, notre équipe a montré que SrpC et SrpNC activent certains gènes cibles de manière similaire mais aussi elles en régulent d'autres différemment. En plus, l'interaction entre SrpNC et son cofacteur FOG, U-shaped, est responsable de certaines mais pas de toutes les activités distinctes de SrpC et SrpNC. Le but de cette étude est de fournir une investigation génétique approfondie des rôles fonctionnels différentiels possibles des isoformes Srp au cours du développement de la drosophile. En utilisant la technologie CRISPR/Cas9, nous avons généré deux lignées de mouches mutantes invalidées soit pour SrpC ou pour SrpNC. En outre, nous avons produit une troisième lignée de mouche mutante dans laquelle nous avons spécifiquement introduit dans le doigt de zinc N-terminal de Srp une mutation ponctuelle qui modifie son interaction avec U-shaped.

L'analyse de ces mutants a révélé que les deux isoformes régulent d'une manière redondante la transcription d'un ensemble commun de gènes au cours du développement intestinal ainsi que de quelques gènes impliqués dans l'hématopoïèse précoce. Étonnamment, les mouches dépourvues de SrpNC (isoforme contenant deux doigts de zinc comme les facteurs GATA des mammifères) sont viables, montrant que cette isoforme est dispensable pour la plupart des processus de développement contrôlés par Srp. Néanmoins, SrpNC semble être spécifiquement nécessaire pour le maintien de l'homéostasie des cellules sanguines et pour la fertilité des mouches. En outre, la perturbation de l'interaction de Srp et de son cofacteur FOG U-shaped équivaut à la perte complète de l'isoforme SrpNC, montrant que SrpNC forme un complexe avec U-shaped pour assurer ses fonctions. En revanche, notre approche génétique a révélé que l'isoforme SrpC est essentielle pour la viabilité et le développement du corps gras, suggérant que cette isoforme régule différents programmes développementaux par rapport à SrpNC. Dans l'ensemble, nos résultats révèlent une plus grande flexibilité fonctionnelle jouée par les doigts

de zinc GATA pour remplir leurs nombreux rôles tout au long du développement. En outre, ce travail illustre que, comme la duplication du génome chez les vertébrés, l'épissage alternatif fournit une stratégie efficace pour promouvoir la sous-fonctionnalisation et générer la diversité fonctionnelle des facteurs GATA chez les invertébrés.

Mots clés

GATA, Friend of GATA, Drosophile, Doigt de zinc, Développement, Hématopoïèse.

Foreword

Cells are the units that constitute all living organisms. The human body is composed of trillions of cells that are organized into at least 200 different cell types. For every single type, the cells are tensely programmed in order to acquire specific shape and to accomplish particular functions. Cells of one type constitute a tissue, and different tissues coordinate together in order to form multifunctional organs. Although all the cells possess the same genetic material, the diversity of physical and functional properties between the different cell types depends on the activation or the repression of different set of genes in these cells. The determination of the gene expression state in every cell and at every specific time is under the precise control of thousands of transcription factors and cofactors. Mutations affecting these transcriptional regulators have been widely associated to a broad range of human diseases including cancer. Therefore, over the years, a huge number of studies were focusing on deciphering the functions of multitude of transcription factors and their cofactors, in order to understand how these proteins act, in normal and/or pathologic sistuations. Herein, I will focus on the transcription factors belonging to the GATA family and their cofactors of the Friend of GATA (FOG) family.

The fruit fly *Drosophila melanogaster* has been introduced into the genetic research world in 1910 by the American scientist Thomas Hunt Morgan who discovered the *white-eye* mutation in the fly and its linkage to the X-chromosome. Throughout the twentieth century, lots of efforts have been dedicated to accumulating knowledge on *Drosophila* genes and their role in different vital processes such as the development, homeostasis maintenance, adaptive behaviors and response to stress. The major outcome of these works is the unexpected level of similarity between flies and humans at the molecular and physiological levels. Indeed, it has been

estimated that about 77% of genes associated to human diseases have a functional homolog in the fly (Reiter et al., 2001). In addition, flies are easy to raise in laboratory conditions, they have many offspring and they have a short life cycle. All of these advantages make *Drosophila* a powerful model organism to study and to understand the mode of action of different proteins and regulators, notably those associated to human diseases.

During my PhD work, I was interested in studying the roles played in different organs and at different stages of development by two proteins of the GATA and of the FOG families respectively, called Serpent and U-shaped.

Through this thesis, I will start by a short description of the state of the art concerning transcription factors of the GATA family and their FOG cofactors, at the molecular and functional levels, both in mammals and in *Drosophila*. Next, I will present the experimental results I obtained during my PhD internship. Finally, I will briefly present and discuss the conclusions dawned from these results.

List of abbreviations

α-Spec: alpha spectrin

Adh: Alcohol dehydrogenase

AEL: Acute erythroid leukemia

AF: Atrial fibrillation

AGM: Aorta-gonad-mesonephros

ALL/LBL: Acute lymphoblastic leukemia/lymphoma

Amel*: *Apis mellifera

AML: Acute myeloid leukemia

ASD: Atrial septal defects

BAV: Bicuspid aortic valve

BC: blast crisis

BR-C: Broad-Complex

BRG1: Brahma-related gene-1

C.elegans*: *Caenorabditis elegans

C-ZnF: C-terminal zinc finger

Cas9: CRISPR-associated protein 9

CBP: CREB-binding protein

CML: chronic myeloid leukemia

Col: Collier

Crb2: Crumbs Cell Polarity Complex Component 2

CRISPR: Clustered regularly interspaced short palindromic repeats

Crq: Croquemort

CtBP: C-terminal binding protein

CYP11A1: Cytochrome P450 Family 11 Subfamily A Member 1

CYP19: Cytochrome P450 aromatase

DCM: Dilated cardiomyopathy

DCML: Dentritic cell, monocyte, B and NK lymphoid deficiency

***Dmel*:** *Drosophila melanogaster*

dsRNA: double stranded RNA

E2A: E2-alpha

EKLF: Erythroid Krüppel-like factor

EMT: Epithelial-to-mesenchymal transition

EPP: erythropoietic porphyria

ETP: Early T-cell precursor

Ets: E26 transformation-specific

Fas: Fasciclin

FOG: Friend of GATA

FOXA: Forkhead box subfamily A

Glt: Glutactin

Grn: Grain

H3K9: Histone H3 lysine 9

H3K27: Histone H3 lysine 27

HDR: Hypoparathyroidism, sensorineural deafness and renal insufficiency

***Hsap*:** *Homo sapiens*

HSC: Hematopoietic stem cell

Kit: Tyrosine-protein kinase

KLF: Krüppel-like family factor

Lcup*: *Lucilia cuprina

LDB1: LIM domain binding 1

LGL: large granular lymphocytic leukemia

LMO2: LIM-only protein 2

Lz: Lozenge

MADS : MCM1- AGAMOUS-DEFICIENS-SRF

Mdom*: *Musca domestica

MDS: myelodysplastic syndromes

MED1: Mediator Complex Subunit 1

MEF2: Myocyte enhancer factor 2

msn: mishhapen

MTA3: Metastasis-associated protein

N-ZnF: N-terminal zinc finger

NCBI: National Center for Biotechnology and Information

Nkx2-5: NK2 homeobox 5

NuRD: Nucleosome remodeling and histone deacetylase

Orb: Oo18 RNA-binding protein

Pnr: Pannier

PPH3: Phosphohistone H3

PPO: Prophenoloxidase

PSC: Posterior signaling center

PTA: Persistent truncus arteriosus

PU.1: PU box binding-1

Pxn: Peroxidasin

Refseq: Reference Sequence

RNAi: RNA interference

RT-PCR: Reverse-transcription polymerase chain reaction

RUNX1: Runt-related transcription factor 1

sgRNA: single guide RNA

shRNA: Short-hairpin RNA

SP1: Specificity protein 1

SRF: Serum response factor

Srp: Serpent

SrpC: Serpent containing C-terminal zinc finger domain

SrpNC: Serpent containing N and C-terminal zinc finger domains

ssODN: Single-stranded donor oligonucleotides

STAR: Steroidogenic acute regulatory protein

SWI/SNF: Switch/sucrose nonfermenting

TAL1: T-cell acute leukemia 1

Tcas*: *Tribolium castaneum

Tig: Tiggrin

Tin: Tinman

Tj: Traffic-jam

TOF: Tetralogy of Fallot

Ush: U-shaped

VDRC: Vienna Drosophila Resource Center

Vkg: Viking

VSD: Ventricular septal defects

VT: Vienna tile

Wdr77: WD Repeat Domain 77

ZFPM: Zinc Finger Protein, FOG Family

List of figures

Chapter I. Introduction

(A) Introduction to GATA transcription factors

GATA proteins belong to a well characterized and widely studied family of transcription factors, whose founding member was identified in 1988 in chicken erythroid nuclear extract, for its ability to bind two distinct sites within an enhancer region of the *β-globin* gene and thus activate *β-globin* expression. The two DNA sequences bound by this factor contain a common motif WGATAR, in which W corresponds to an A or T nucleotide and R to an A or G (Evans et al., 1988). This is why this factor is named GATA1 (Orkin, 1990). After GATA1's discovery, screening of cDNA libraries using a murine GATA1 cDNA clone as a probe led to the identification of two other members of GATA family: GATA2 and GATA3 (Orkin, 1990; Yamamoto et al., 1990). Later, a similar technique allowed the discovery of GATA4, GATA5 and GATA6 factors (Arceci et al., 1993; Laverriere et al., 1994). The particularity of all these identified GATA proteins is their ability to recognize and bind GATA motif-containing DNA sequences (Morimoto et al., 1999; Romano & Miccio, 2020).

1- Molecular structure of GATA transcription factors

GATA transcription factors are highly conserved proteins that are present in organisms ranging from flies to humans. In mammals, the six identified GATA factors (GATA1 to GATA6) contain two zinc finger domains, referred to as N-terminal and C-terminal zinc fingers (M. H. Lentjes et al., 2016; M. Tremblay et al., 2018). Each of the two zinc finger domains is formed of 4 cysteine residues that coordinate a single zinc ion. These cysteine residues are positioned into a sequence with the characteristic Cys-X2-Cys-X17-Cys-X2-Cys spacing, and the two zinc

finger domains are separated by a linker of 29 amino acids (Figure 1). At the C-terminal side of each zinc finger domain, a basic amino acid-containing region is found, which is necessary for the binding of GATA proteins to DNA (Omichinski et al., 1993; Pedone et al., 1997). The binding of GATA to DNA is mainly established by the C-terminal zinc finger domain (C-ZnF) and its adjacent basic C-terminal region (Martin & Orkin, 1990; Omichinski et al., 1993; Yang & Evans, 1992). Although dispensable for binding to the GATA-containing DNA motif, the N-terminal zinc finger domain (N-ZnF) contributes to stabilization of Protein/DNA interaction, predominantly on palindromic GATA sequences (Martin & Orkin, 1990; Trainor et al., 1996; Yang & Evans, 1992). In addition, it was shown that the N-ZnF of GATA2 and GATA3 proteins is able to bind to DNA sites containing a GATC sequence, in a manner dependent on its adjacent basic C-terminal region (Pedone et al., 1997). Furthermore, some basic residues in this adjacent region can regulate GATA transcriptional activity without affecting the Protein/DNA interaction. For example, it was found that mutating the KRR basic amino acids located between GATA3's zinc finger domains, abolishes GATA3-mediated activation of minimal T cell receptor alpha and beta enhancers in vitro, but has no effect on the protein's ability to associate with DNA (V. M. Smith et al., 1995). Finally, the GATA N-terminal and C-terminal zinc finger domains also play a role during GATA's interaction with other transcriptional regulators (Jason A. Lowry & Mackay, 2006). Mammalian GATA factors contain a nuclear localization signal and transcriptional trans-activating domains located in the C-terminal and the N-terminal domain of the protein respectively (E. E. Morrisey, Ip, Tang, & Parmacek, 1997). However, contrary to the zinc finger domains that are highly conserved among all six mammalian GATA factors, the GATA N-terminal and C-terminal regions display only a low level of amino acid similarity (Figure 2) (M. Tremblay et al., 2018).

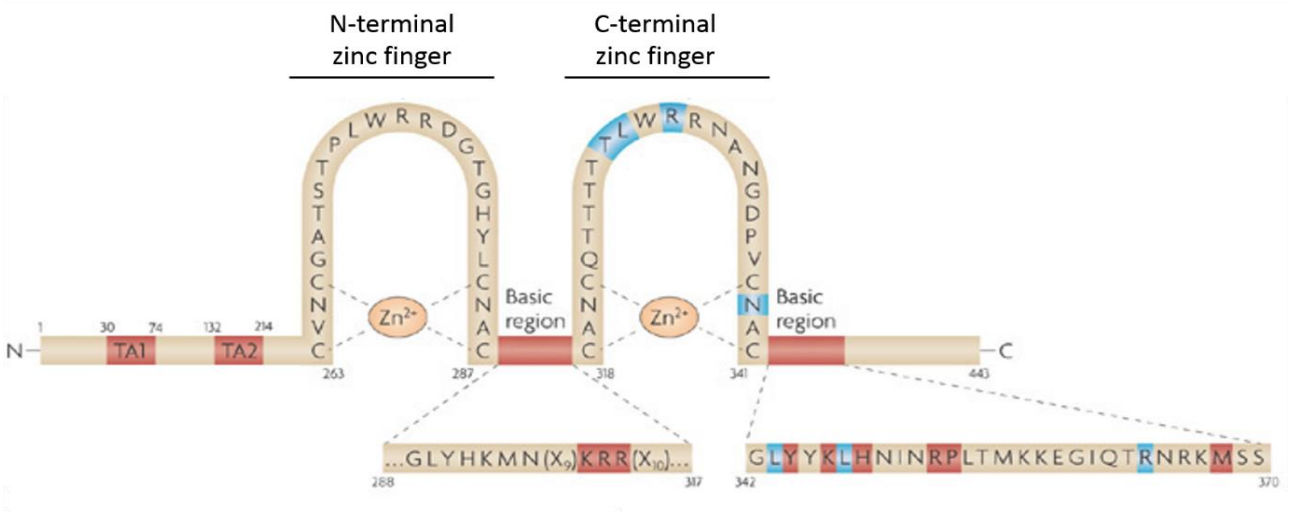


Figure 1. Functional domains and essential amino acids in the mammalian GATA3 transcription factor.

GATA3 protein length is 443 amino acids. GATA3 has two transactivation domains (TA1 and TA2) in the N-terminal region of the protein (N-) and two zinc finger domains (N-terminal zinc finger and C-terminal zinc finger) followed each by a conserved basic region. Each zinc finger domain has the characteristic Cys-X2-Cys-X17-Cys-X2-Cys spacing where the four cysteine residues coordinate a single zinc ion (Zn^{2+}). The two zinc finger domains are separated by a linker of 29 amino acids. The amino acid residues marked in blue were shown in a crystal structure of the C-terminal zinc finger to make direct contact with DNA while residues marked in red have been shown to be involved in normal GATA3 function (adapted from (Ho et al., 2009).

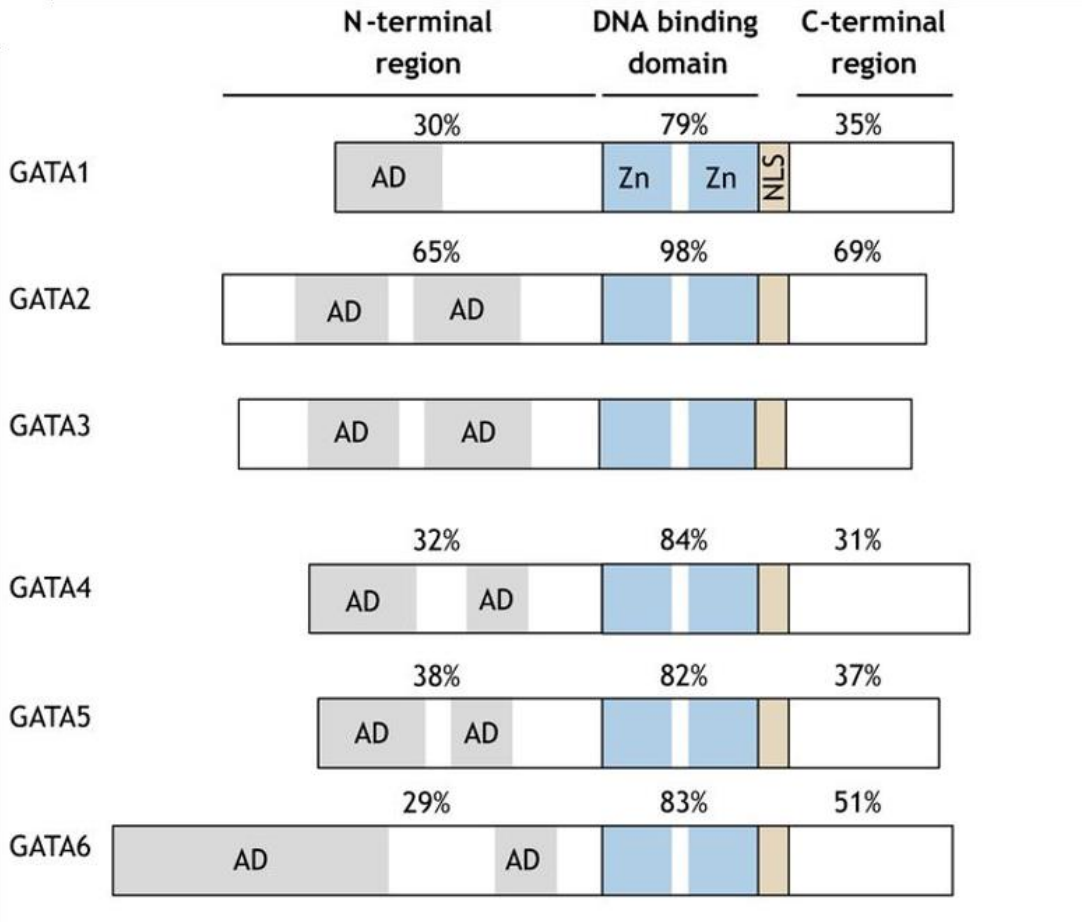


Figure 2. Conservation of the zinc finger domains in the six mammalian GATA.

The six mammalian GATA transcription factors contain two highly conserved zinc finger domains (Zn), a nuclear localization signal (NLS) and two less conserved C-terminal and N-terminal regions. The N-terminal region contains transcriptional activation domains (AD). Each percentage represents the similarity level of the protein sequence of one part of the corresponding protein in comparison to its equivalent part in GATA3 protein (M. Tremblay et al., 2018).

In *Drosophila,* five transcription factors of the GATA family have been identified: GATAa or Pannier (Pnr), GATAb or Serpent (Srp), GATAc or Grain (Grn), GATAd and GATAe (Abel et al., 1993; Lin et al., 1995; Okumura et al., 2005; Ramain et al., 1993; Rehorn et al., 1996; Winick et al., 1993). Interestingly, Pnr, Srp and Grn proteins display a canonic structure, containing both N-ZnF and C-ZnF domains; in addition, alignment of their zing finger domains with those of vertebrate GATA proteins, reveals a high degree of identity between them (Figure 3) (Waltzer et al., 2002). Srp, the fourth member of this *Drosophila* family, is somewhat special, since following alternative splicing of its transcript, the gene yields two distinct protein isoforms: the SrpNC isoform containing the two canonical N-terminal and C-terminal zinc finger domains, or the SrpC isoform containing only the C-ZnF (Rehorn et al., 1996; Waltzer et al., 2002). Of note, GATA transcription factors with a single zinc finger domain are found in nematodes, in fungi and in *Drosophila* GATA factors GATAe and GATAd (W. Q. Gillis et al., 2008; J. A. Lowry & Atchley, 2000).

2- GATA factors interact with other transcriptional regulators

a- GATA factors interact with different partners

In most cases, GATA factors alone are not sufficient to regulate gene expression, and they cooperate with other transcriptional regulatory proteins in order to synergistically or antagonistically modify gene expression levels (Jason A. Lowry & Mackay, 2006). Among these partners, the LIM-only protein 2, LMO2, acts as a molecular bridge, linking murine GATA1 to three other proteins, TAL1, E2A and LDB1 (Osada et al., 1995; Wadman et al., 1997). Together, these proteins constitute a so called pentameric complex that allows

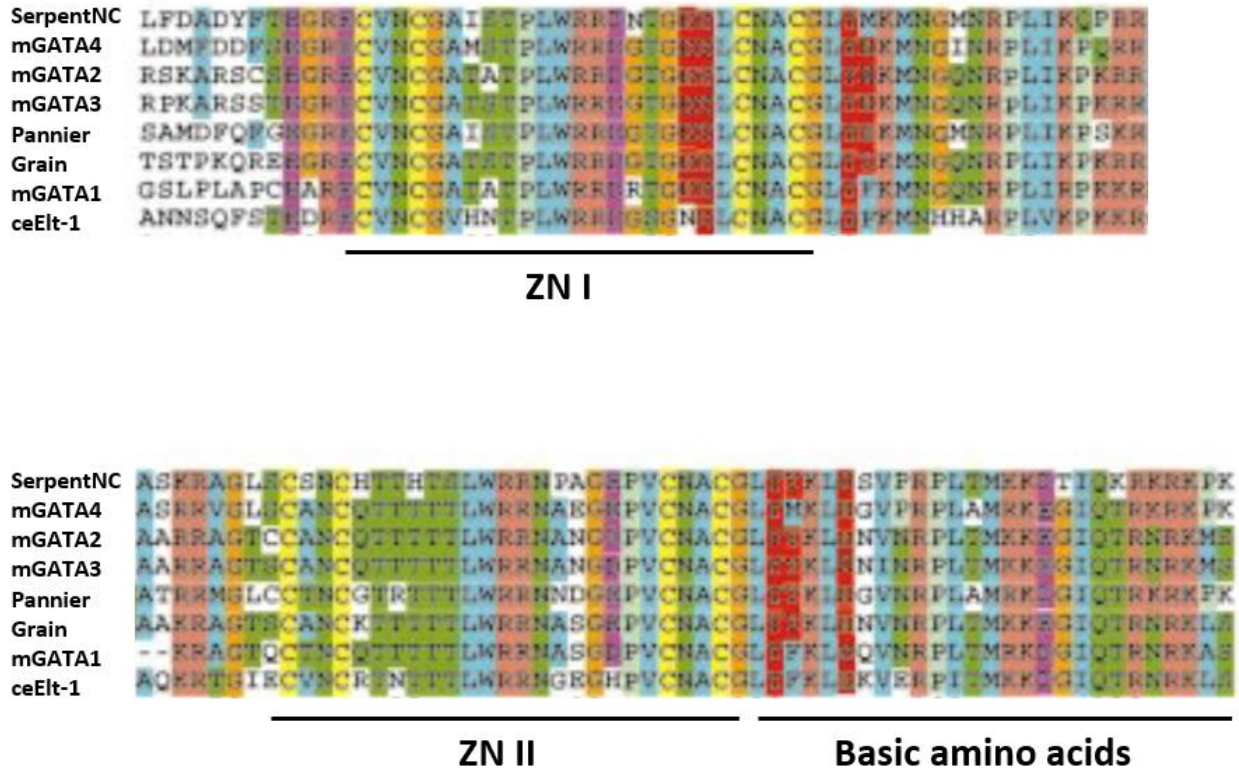


Figure 3. Molecular conservation of the region constituting and surrounding the GATA zinc finger domains between *Drosophila* and mammals.

Alignment of *Drosophila* and mammalian GATA amino acid sequences shows that *Drosophila* GATA factors (SerpentNC, Pannier and Grain) contain N-terminal (ZNI) and C-terminal (ZNII) zinc finger and C-terminal basic tail (Basic amino acids) that are highly conserved with those of mammalian (mGATA4, mGATA2, mGATA3, mGATA1) and *Caenorabditis elegans* (CeElt-1) GATA transcription factors. Conserved amino acids in each column are coloured according to the consensus character assigned to that column. Cysteine residues of the zinc finger domains are highlighted in yellow. ZNI: first zinc finger, ZNII: second zinc finger (adapted from (Waltzer et al., 2002).

simultaneous recognition of the GATA DNA binding site (WGATAR) by GATA1, and of the E-box motif (CANNTG) by TAL1 and E2A. Once bound to DNA, this complex activates reporter gene expression (Osada et al., 1995; Wadman et al., 1997). Murine GATA1 is also able to interact with factors of the Krüppel-like family, Sp1 and EKLF, themselves zinc finger containing proteins. SP1 recognizes both consensus DNA sequences GC and CACC, while EKLF binds a subset of extended CACC sequences. Both of these factors associate with GATA1 in order to synergistically promote erythroid cell-specific gene expression (Lavallée et al., 2006; Merika & Orkin, 1995). In addition, it was demonstrated that human and rat GATA4 proteins interact with Nkx factors, such as Nkx2-5. Nkx factors are homeodomain-containing proteins that recognize a TNAAGTG DNA sequence and that interact with GATA4 to synergistically activate cardiac target genes (Durocher et al., 1997; Liu et al., 2002; Zhu et al., 2000). Similarly, members of the MADS box family of transcription factors, SRF and MEF2, interact with rat GATA4 proteins to promote transcriptional activation of their target genes. The MADS box motif is a DNA-binding and dimerization domain, and proteins of this family recognize A/T-rich DNA regions (Belaguli et al., 2000; S. Morin et al., 2000; Steves Morin et al., 2001). It is noteworthy that depending on the cellular context, interaction with the same partner can yield different outcomes; for instance, in mammals, PU.1 transcription factors of the Ets family interact with GATA1 factors in order to trigger gene expression in eosinophil cells, while they functionally antagonize each other during the differentiation of hematopoietic myeloid progenitors into erythroid *versus* myeloid cells (Du et al., 2002; Nerlov et al., 2000; Rekhtman et al., 1999). Ets family members contain an Ets domain, a winged helix-turn-helix structure recognizing the DNA sequence harboring the core GGAA motif. Finally, factors of the RUNX family contain a Runt domain that binds the TGYGGTY DNA sequence. The RUNX proteins RUNX1 and Lozenge (Lz), interact with the mammalian GATA1 and the *Drosophila* Srp, respectively, during mouse megakaryocyte differentiation and fly crystal cell

(megakaryocyte-like) development (Elagib et al., 2003; Waltzer et al., 2003). In conclusion, proteins of different families such as LIM-only protein 2, Krüppel-like factors, Nkx homeodomain-containing proteins, MADS box containing factors, transcription factors of the Ets family and RUNX family proteins, interact with GATA proteins to regulate their transcriptional activity in a large number of cell types.

b- GATA factors interact with cofactors of the Friend of GATA family

The most widely studied co-factors of GATA factors belong to the Friend of GATA (FOG) family. To date, two members of the FOG family have been identified in mouse, FOG1 and FOG2, also known as ZFPM1 and ZFPM2, respectively (Lu et al., 1999; Svensson et al., 1999; Tevosian et al., 1999; Tsang et al., 1997). FOG1 proteins contain nine zinc finger domains (1 to 9), distributed throughout the protein. Four of these zinc fingers are of the C2H2 type, while the five other zinc fingers are of the C2HC type. FOG2 proteins have eight zinc finger domains that are highly conserved with their eight equivalent zinc finger domains in FOG1 protein (Chlon & Crispino, 2012). Structural comparison between FOG1 and FOG2 shows that the equivalent of the eighth FOG1 zinc finger, a C2H2-type zinc finger, is absent in FOG2, therefore giving rise to a protein with five C2HC type fingers and only three C2H2 type finger domains (Figure 4A). None of the FOG zinc fingers are able to bind DNA, and among the nine FOG1 zinc fingers, only the fingers 1, 5, 6 and 9, which are of the C2HC type, are able to interact with GATA factors. In *Drosophila*, only one FOG cofactor called U-shaped (Ush) has been identified (Cubadda et al., 1997; Haenlin et al., 1997). Ush shares 20% homology with mammalian FOG1 and contains nine zinc finger domains. Similar to FOG1, four of these zinc finger domains are of the C2H2 type, while the five others are of the C2HC type (Figure 4B).

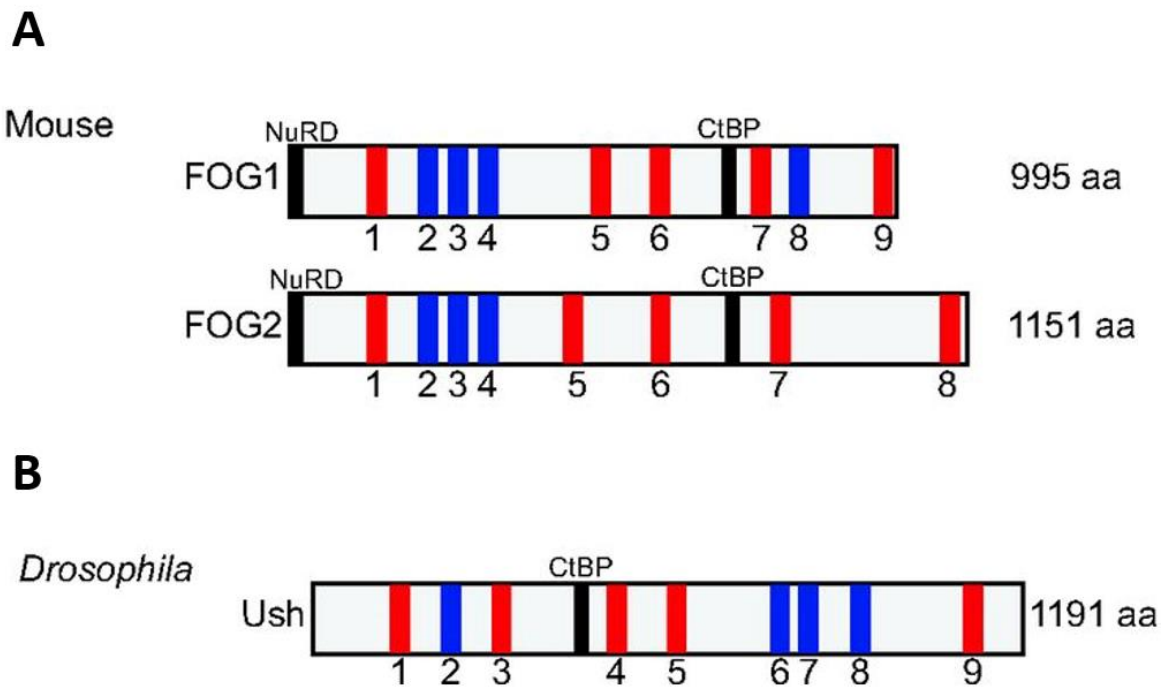


Figure 4. Molecular conservation of mammalian and *Drosophila* GATA cofactors of the Friend of GATA family.

(A) Two members of the Friend of GATA (FOG) family are present in the mouse: FOG1 and FOG2, formed of 995 and 1151 amino acids, respectively. Both cofactors contain zinc finger domains of the C2HC (red vertical bars) and of C2H2 (blue vertical bars) type and also two co-repressor interaction motifs that are colored in black (NuRD or CtBP). (B) In *Drosophila*, U-shaped (Ush) protein of the FOG family is constituted of 1191 amino acid and it contains 8 zinc finger dmains conserved with those of mammalian FOG, and contains also a CtBP interaction motif (adapted from (Chlon & Crispino, 2012)).

Amino acid sequence alignment of FOG1, FOG2 and Ush zinc fingers, which are able to interact with GATA factors, revealed a number of conserved residues. The resulting consensus sequence for FOG binding to GATA factors was predicted to be X3-Phe-X-Cys-X2-Cys-X-Ile-X2-Arg/Ser-X3-Thr/Asn-X3-His-X2-Tyr-Tyr-Cys-X3, where X stands for any amino acid residue (A. H. Fox et al., 1999). Mutation of key residues in zinc finger 1 of FOG1 protein was shown to interfere with the FOG1/GATA1 interaction, demonstrating their importance during FOG cofactor binding to GATA proteins (A. H. Fox et al., 1999).

FOG zinc fingers interact with GATA proteins by specifically binding the core of the GATA N-ZnF (A. Fox et al., 1999; A. H. Fox et al., 1999; Tsang et al., 1997). Interestingly, the ability of FOG to recognize GATA N-ZnF, and not C-ZnF, is due to the presence in the N-ZnF of key residues that are important for the interaction with FOG, and which are not found in the GATA C-ZnF. These residues (referring to the murine GATA1 protein) are Glu203, Val205, Gly208, Ala209, His222 and Tyr223 (Figure 5) (A. Fox et al., 1999). Substitution of the murine GATA1 Val205 residue into glycine (V205G) impairs the interaction with FOG1 and modulates the chromatin occupancy of GATA1 during hematopoietic cell lineages specification (Chlon et al., 2012; J. D. Crispino et al., 1999). Mapping of the key FOG interacting residues onto the solved three-dimensional structure of the GATA1 N-ZnF domain, shows that they form a single contiguous surface, essentially located outside of the DNA binding region (A. H. Fox et al., 1999; Kowalski et al., 1999). Interestingly, these identified residues are highly conserved among the N-ZnF of all six mammalian GATA transcription factors (M. H. Lentjes et al., 2016). In agreement with the fact that FOG proteins recognize the GATA N-terminal, but not C-terminal, zinc finger domain, it was demonstrated that the *Drosophila* FOG cofactor Ush binds GATA factors Pnr and SrpNC, which both contain the canonical N-ZnF domain. However, Ush

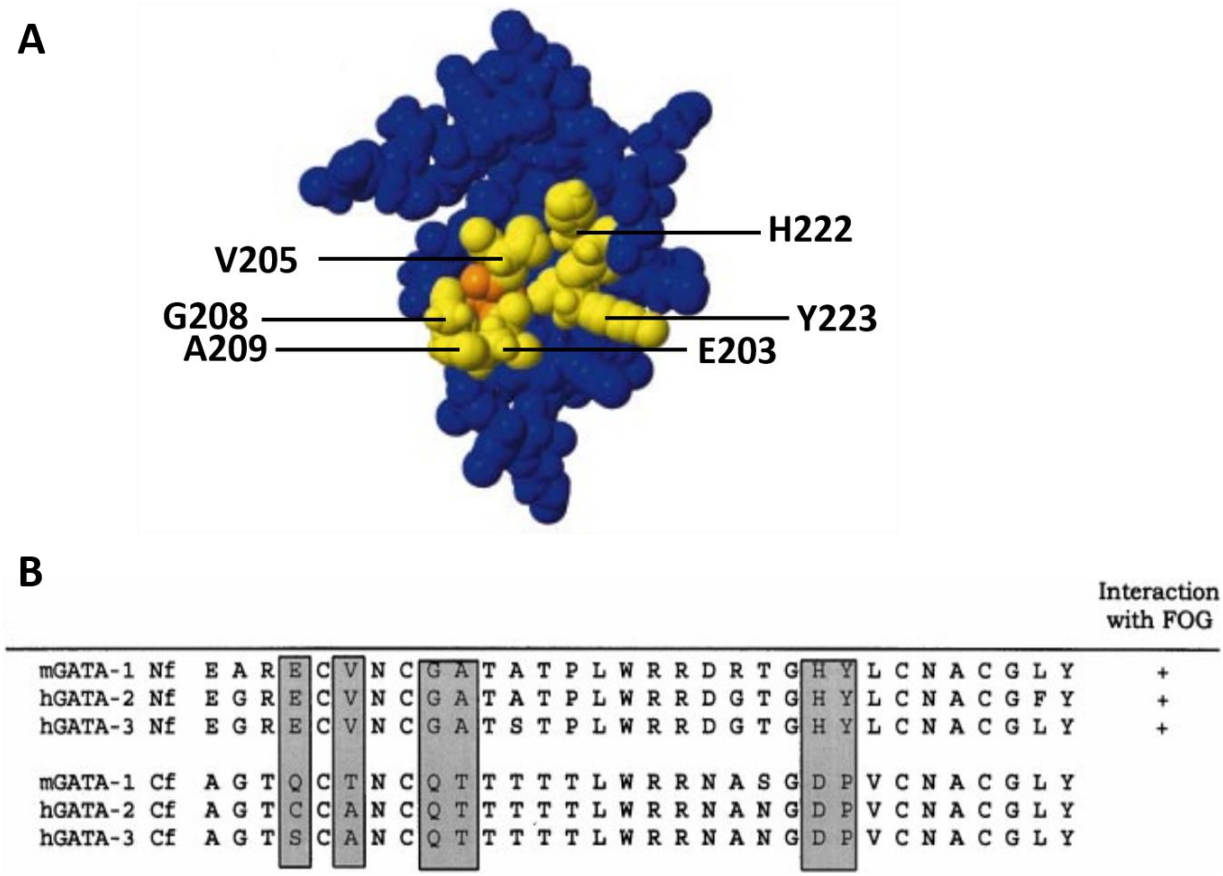


Figure 5. Key amino acid residues identified in GATA1 N-terminal zinc finger domain as required for GATA/FOG interaction.

(A) Representation of 3D structure of GATA1 N-terminal zinc finger domain. The six amino acid residues identified as critical for the interaction with FOG are shown in yellow and their position in the protein is added. The cysteine residues colored in orange attache the zinc ion of the finger. (B) The key amino acid residues (marked in grey) are conserved among the N-terminal zinc finger of mammalian GATA factors (mGATA-1, hGATA-2, hGATA-3) but they are absent from the C-terminal zinc finger domain (adapted from (A. Fox et al., 1999; Kowalski et al., 1999).

is unable to interact with the SrpC isoform that is devoid of the N-ZnF (Haenlin et al., 1997; Waltzer et al., 2002). Finally, the SrpNC^{V421G} protein variant harboring a valine to glycine substitution equivalent to the mammalian GATA1 V205G, displays altered interaction with Ush and behaves as the SrpC isoform rather than as the SrpNC isoform (Nancy Fossett et al., 2003). These results indicate a conserved mode of interaction of GATA and FOG proteins between *Drosophila* and mammals (Nancy Fossett et al., 2003; A. Fox et al., 1999; Waltzer et al., 2002).

Depending on the cellular context and the promoter bound by the GATA/FOG complexes, FOG may acts as a co-activator or as a co-repressor of the GATA-dependent transcription (Chlon & Crispino, 2012). Supporting the idea of their co-repressor function, two co-repressor binding motifs have been identified in both mammalian FOG1 and FOG2 (Figure 4A). The first one, the PIDLS motif, allows the fixation of the C-terminal binding protein (CtBP), while the second is known to recruit the nucleosome remodeling and histone deacetylase (NuRD) complex (A. H. Fox et al., 1999; Hong et al., 2005). Similarly, in *Drosophila*, Ush acts as negative regulator for GATA factors activity (N. Fossett et al., 2000, 2001; Haenlin et al., 1997) and contains a co-repressor binding motif allowing interaction with CtBP. However, unlike FOG1 and FOG2, Ush does not contain any NuRD complex binding site (Chlon & Crispino, 2012).

Altogether, GATA proteins have the capacity to interact with different partners belonging to the Friend of GATA family, as well as with members of other protein families, in order to modulate their own functions. This ability to contrastingly interact with a variety of proteins, opens up the functional range of GATA proteins in mammals and other GATA producing organisms.

3- GATA mutations are associated to human diseases

a- Mutations in GATA factors induce the formation of numerous pathologies

The relation between GATA gene mutations and human diseases has been widely studied. Genome, exome and transcriptome sequencing have led to the identification of a huge number of GATA mutations in patients with different types of biological disorders (M. H. Lentjes et al., 2016). The type of disease depends on the affected GATA gene and its expression pattern. For example, GATA1, GATA2 and GATA3 proteins are expressed in hematopoietic cell lineages, and mutations affecting these factors are related to numerous hematological disorders: acute myeloid leukemia (AML), dendritic cell, monocyte, B and NK lymphoid deficiency (DCML), myelodysplastic syndromes (MDS), large granular lymphocytic leukemia (LGL), chronic myeloid leukemia (CML) with blast crisis (BC), Emberger syndrome, neutropenia, thrombocytopenia, dyserythropoietic anemia, β-thalassemia, erythropoietic porphyria (EPP), acute erythroid leukemia (AEL) and early T-cell precursor (ETP) acute lymphoblastic leukemia/lymphoma (ALL/LBL) (John D. Crispino & Horwitz, 2017; Ping et al., 2017; Spinner et al., 2014; J. Zhang et al., 2012; S.-J. Zhang et al., 2008).

Furthermore, GATA3 is also expressed in developing and differentiated mammary glands (Asselin-Labat et al., 2007; Kouros-Mehr et al., 2006), as well as in embryonic kidney, inner ear and parathyroid glands (Debacker et al., 1999; Labastie et al., 1995; Rivolta & Holley, 1998). Accordingly, GATA3 mutations are found associated to breast cancer, hypoparathyroidism, sensorineural deafness and renal insufficiency (HDR) syndrome (Ellis et al., 2012; Muroya et al., 2001; Okawa et al., 2015; Usary et al., 2004).

Similarly, the factors GATA4, GATA5 and GATA6 that are expressed in the mammalian developing heart (Pikkarainen et al., 2004), are associated to cardiac diseases, such as atrial septal defects (ASD), ventricular septal defects (VSD), bicuspid aortic valve (BAV), dilated cardiomyopathy (DCM), tetralogy of Fallot (TOF), familial atrial fibrillation (AF) and persistent truncus arteriosus (PTA). Additionally, alteration of these factors provokes other diseases such as neonatal diabetes, adult-onset diabetes and congenital diaphragmatic hernia (Allen et al., 2011; De Franco et al., 2013; L. Yu et al., 2013). Finally, loss-of-function mutations in GATA4 are also associated to the 46, XY disorder of sex development (Lourenço et al., 2011; Martinez de LaPiscina et al., 2018).

b- Point mutations in GATA zinc fingers provoke human diseases whose severity depends on the position and the nature of the mutation

Interestingly, most GATA associated human diseases were identified in patients carrying point mutations affecting residues located in and around the two GATA zinc finger domains (Figure 6) (M. H. Lentjes et al., 2016). The severity of the diseases associated to GATA zinc finger point mutations varies depending on the affected zinc finger domain, on the residue altered in the zinc finger domain, but also on the nature of the substitution, as the same residue can be the target of more than one type of substitution, resulting in a different clinical outcome.

In GATA2 protein, many identified point mutations in the N-ZnF domain are associated to AML, while point mutations found in the GATA2 C-terminal zinc finger lead to the development of other hematopoietic disorders, such as MDS, AML and CML with BC formation (M. H. Lentjes et al., 2016).

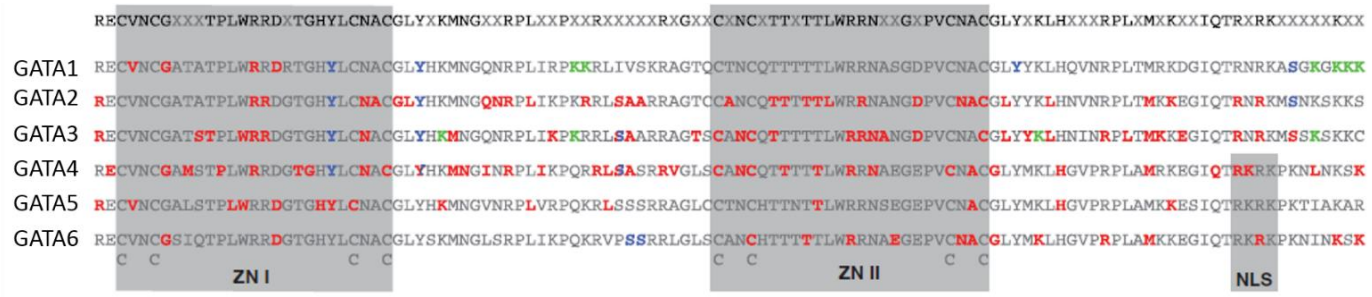


Figure 6. Point mutations inside and surrounding the mammalian GATA zinc finger domains are associated to human diseases.

The mammalian GATA amino acid sequences are aligned and the corresponding GATA type is written beside each sequence. Amino acids marked in blue and in green represent respectively protein phosphorylation and acetylation sites. Amino acids wrote in red are disease-associated alteration sites. Post-transcriptional modification and disease-associated alteration sites are marked on top of the corresponding amino acid. The two GATA zinc finger domains and the nuclear localization sequence (NLS) are highlighted in grey and positions of zinc finger cysteine residues are shown on bottom of the alignment. ZNI: first zinc finger, ZNII: second zinc finger (adapted from (M. H. Lentjes et al., 2016)).

Also, in the same N-ZnF domain of GATA1, mutating amino acids that are required for the GATA1/FOG1 interaction, such as the residues V205, G208 and D218, alters erythrocyte and thrombocyte formation, leading to hematological disorders including anemia and thrombocytopenia (K. Freson et al., 2001; Kathleen Freson et al., 2002; Mehaffey et al., 2001; Nichols et al., 2000). Meanwhile, mutation of the R218 arginine residue, located in the same GATA1 N-ZnF, affects the ability of the factor to bind to DNA palindromic sequences, without affecting the interaction GATA1/FOG1, and leads to thrombocytopenia and β-thalassemia, associated to gray platelet syndrome and porphyria, respectively (Balduini et al., 2004; Phillips et al., 2007; Tubman et al., 2007; C. Yu et al., 2002).

Finally, at the same amino acid position, the nature of the substitution affects the severity of the outcome. For instance, two mutations affect the GATA1 Gly208 residue, G208S and G208R, but the former is associated to macrothrombocytopenia and mild dyserythropoietic anemia, while the latter is associated to severe macrothrombocytopenia and severe dyserythroipoietic anemia (Mehaffey et al., 2001; Vecchio et al., 2005). Analyzing the effect of these two mutations on GATA1's ability to interact with FOG1, showed that G208R generates a stronger disruption of GATA1/FOG1 interaction as compared to G208S, which might explain the more severe phenotypes developed by patients with the G208R substitution as compared to those carrying the G208S mutation (A. E. Campbell et al., 2013). Similar results were observed in the case of the D218G and D218Y mutations. The D218G mutation is associated to thrombocytopenia without anemia, while the D218Y mutation provokes thrombocytopenia and severe dyserythropoietic anemia (K. Freson et al., 2001; Kathleen Freson et al., 2002). The more severe phenotype obtained in patients with the D218Y mutation mirrors the fact that the substitution of the D218 residue into a tyrosine alters more extensively the affinity of GATA1 for FOG1 than its substitution into a glycine (A. E. Campbell et al., 2013; Kathleen Freson et

al., 2002). All these results show that the GATA zinc finger domain amino acids have critical and sophisticated roles during the establishment of GATA functions.

In conclusion, altering the ability of GATA proteins to function properly is associated to several human disorders, where the type and the severity of the disease vary depending on the nature and on the position of the mutation. Given the important and critical roles played by GATA factors in humans, understanding their mode of action is of definite interest.

(B) GATA factors mode of action

1- Mechanistic functions of GATA factors

a- GATA and transcriptional activation

GATA1, the first member of GATA factors was identified by its ability to bind the *β-globin* gene enhancer and to activate its expression in chicken red blood cell precursors (Evans et al., 1988). Two transactivation domains are present in GATA1. They are located in the N-terminal and the C-terminal regions of the protein and they play redundant as well as specific functions during regulation of hematopoiesis in mice (Kaneko et al., 2012). Although poorly conserved, those transactivation domains have been identified in the six mammalian GATA factors (Chlon & Crispino, 2012).

Studies in mice and isolated human cells revealed several target genes that are positively regulated by GATA factors (Cheng et al., 2009; Kurek et al., 2007; Martynova et al., 2019; Welch et al., 2004; Ming Yu et al., 2009). As many other transcriptional factors, GATA proteins recruit co-activators with histone acetyl- and methyl-transferase activities, in order to render the DNA more accessible to transcription and thus regulate gene expression (Figure 7D). Accordingly, both histone acetyltransferase proteins p300 and CBP have been shown to interact with various GATA factors and to act as transcriptional co-activators (Blobel et al., 1998; Dai & Markham, 2001; Kakita et al., 1999; Wada et al., 2000). Moreover, by recruiting H3K79 monomethyltransferase protein, GATA1 allows the methylation of their bound elements within the murine *β-globin* enhancer, prior to the activation of *β-globin* transcription (Steger et al., 2008).

In *Drosophila*, like in mammals, GATA factors act as transcriptional activators. For example, Pnr and Srp interact with components of the mediator transcriptional co-activator complex in order to promote expression of the pro-neural genes *achaete* and *scute*, and of the anti-microbial peptide coding gene *Metchnikowin*, respectively (Garcia-Garcia et al., 1999; Immarigeon et al., 2019; Kuuluvainen et al., 2014).

b- GATA and transcriptional repression

In addition to their functions as transcriptional activators, GATA factors act as transcriptional repressors (Figure 7E). For example, GATA3 directly interacts with transcriptional co-repressors such as the NuRD complex member MTA3, or the H3K9 mono- and dimethyltransferase G9A protein in human breast adenocarcinoma cells.

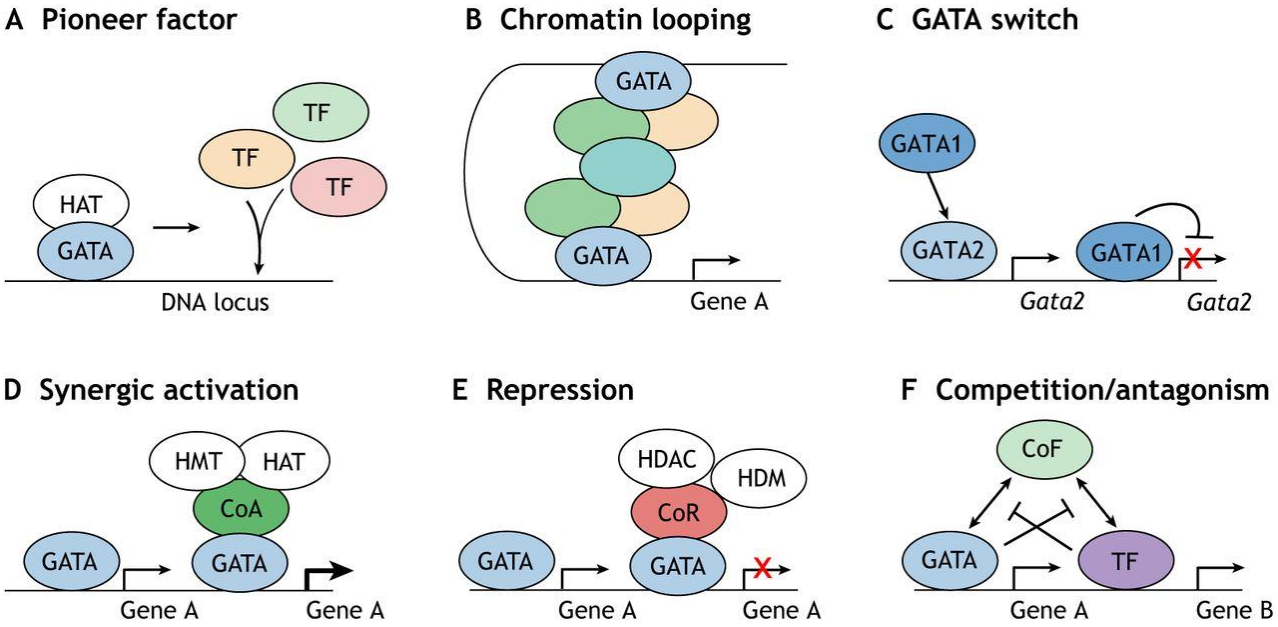


Figure 7. GATA transcription factors mode of action.

(A) GATA proteins act as pioneer factors. They bind heterochromatic DNA and recruit chromatin modifier proteins, such as the histone acetyltransferase (HAT), in order to make the chromatin less compact and more accessible for binding of other transcription factors (TF). (B) GATA can participate to chromatin looping. They bind distant enhancers and change chromatin conformation in order to interact with specific genes promoter. (C) GATA factors can displace each other by the GATA switch process. GATA2 displaces GATA1 on the *GATA2* gene regulatory region in order to inhibit the GATA2 auto-regulation (red cross). (D) GATA factors recruit regulatory coactivators (CoA) and activate gene expression after histone modification by HAT and histone methyltransferase (HMT). (E) GATA factors can recruit regulatory corepressors (CoR) to inhibit gene expression after histone modification by histone deacetylase proteins (HDAC) and histone demethylase (HDM). (F) GATA factors can antagonize the function of other TF by competing for binding to mutual cofactor (CoF) (M. Tremblay et al., 2018).

The GATA3/MTA3/G9A complex represses the expression of proteins implicated in breast cancer metastasis (Si et al., 2015). Furthermore, GATA factors might also indirectly interact with the NuRD complex. Indeed, mammalian GATA1 physically interacts with FOG1, which can in turn directly associate with NuRD. FOG-dependent GATA-mediated transcriptional repression has been demonstrated for various target genes, such as *c-kit* and *GATA2* (J. D. Crispino et al., 1999; Hong et al., 2005).

In *Drosophila*, Ush is responsible for the inhibition of different target gene expression. For example, in the larval hematopoietic progenitors, Ush represses the expression of *hedgehog*, possibly through the formation of a repressive SrpNC/Ush complex (Y. Tokusumi et al., 2010). In addition, Ush negatively regulates the differentiation of *Drosophila* embryonic blood cells into lamellocytes (Avet-Rochex et al., 2010). Lamellocytes are *Drosophila* blood cells that are massively produced in response to infestation by parasitic wasps or under different stress conditions, but are undetected under normal conditions (Lanot et al., 2001; R. Sorrentino et al., 2002; Y. Tokusumi et al., 2018).

c- GATA and chromatin looping

In order to control gene expression, GATA factors participate in 3-dimensional chromatin reorganization. They bind distant regulatory elements and after changing chromatin conformation, they form a physical link between the distant bound elements and the promoter of their transcriptional target genes (Figure 7B). This mode of action has notably been described for the regulation of murine *c-kit* expression (in hematopoietic progenitor cells), *β-globin* (in erythroblasts), and the cytokines *interleukin -4*, *-5* and *-13* (in T helper type 2 cells) (Jing et al., 2008; Spilianakis & Flavell, 2004; Vakoc et al., 2005). However, chromatin

conformation alteration by GATA proteins might also require other gene expression regulators, such as the mediator protein Med1, the chromatin remodeler BRG1, the bridging molecule LDB1 and the cofactor FOG1 (Kim et al., 2009; Song et al., 2007; M. Stumpf et al., 2006; Vakoc et al., 2005).

2- Dynamic functions of GATA factors

a- GATA as pioneer factors

In addition to their classical function of binding GATA motifs in DNA and regulating gene transcription, GATA proteins also contribute to the remodeling of DNA packaging. For instance, GATA1 physically interacts through its zinc finger domains with the mammalian chromatin remodeling complex SWI/SNF, in order to efficiently activate transcription from nucleosome assembled promoters *in vitro* (Kadam et al., 2000). Also, murine GATA4 associates to heterochromatic DNA, in order to decompact chromatin and promote DNA accessibility for other transcription factors (Cirillo et al., 2002). The ability to affect chromatin conformation prior to gene regulation by other factors, is a characteristic feature of ‘‘pioneer factors’’ (Figure 7A) (Zaret & Carroll, 2011). The role of GATA proteins as pioneer factors was discovered in 2002, when GATA4 and FOXA factors were both found to bind the *albumin* gene enhancer, in order to decompact chromatin and promote hepatocyte specification (Bossard & Zaret, 1998; Cirillo et al., 2002).

b- The GATA switch

Different members of the GATA family were shown to bind the same chromatin sites sequentially, in order to yield different transcriptional outputs in a dynamic fashion (Bresnick et al., 2010; Doré et al., 2012; Huang et al., 2016). This process of GATA factor displacement by another member of the family is called the ‘‘GATA switch’’ (Figure 7C). The most studied GATA-switch context is the displacement of murine GATA2 by GATA1 at the *GATA2* locus upon erythroid differentiation. This displacement is responsible for switching off the feed-forward autoregulatory loop of GATA2, by removing the histone acetyltransferase CBP and altering chromatin looping conformation (Grass et al., 2003, 2006; Martowicz et al., 2005). Many other genes are known to be the target of GATA switch events. For example, expression of *Wdr77* and *kit* genes is restricted after GATA switches, in order to inhibit proliferation of the murine developing blood cells (Rylski et al., 2003; Min Yu et al., 2016).

In *Drosophila*, regulation of the same gene by two different GATA proteins has been illustrated in the developing embryonic gut, where Srp and GATAe act sequentially in the endoderm to control expression of the ectodermal protein coding gene *brachyenteron*. Srp inhibits *brachyenteron* expression during early stages of embryogenesis, and GATAe acts during later stages of embryogenesis, when *srp* expression has ceased (R. Murakami et al., 1999; Okumura et al., 2005; Reuter, 1994).

c- GATA antagonism with other transcriptional regulators

GATA and other transcription factors have been shown to antagonize the function of each other in some contexts (Figure 7F). For instance, mouse GATA1 promotes common myeloid progenitor differentiation towards an erythrocytic fate, while PU.1 directs their differentiation into myeloid cells (Hoppe et al., 2016). PU.1 in developing erythrocytes inhibits GATA1-mediated erythropoiesis, and reciprocally GATA1 expression inhibits PU.1-mediated transcription of myeloid specific genes (Nerlov et al., 2000; Rekhtman et al., 1999; P. Zhang et al., 1999). However, although GATA1 and PU.1 ensure antagonistic functions during determination of erythroid *versus* myeloid cell fates, they act synergistically during differentiation of myeloid cell progenitors into mast cells (Du et al., 2002). However, during this process, FOG1 has to be down-regulated. In the case of FOG1 ectopic expression in mast cells progenitors, FOG competes with PU.1 for interaction with GATA1 and hence affects GATA1/PU.1-dependent mast cell differentiation (Cantor et al., 2008; Sugiyama et al., 2008).

Moreover, competition for binding to GATA proteins has been identified in *Drosophila*, as Srp interacts with the transcription factor of the RUNX family Lozenge in order to trigger differentiation of crystal cells, which are reminiscent of mammalian megakaryocytes. However, Ush competes with Lz for the association with Srp, and overexpression of *ush* in crystal cell progenitors alters Srp/Lz-mediated crystal cell differentiation (N. Fossett et al., 2001; Nancy Fossett et al., 2003; Muratoglu et al., 2007).

Taken together, these results show that GATA factors have many different modes of action. They cooperate with or antagonize other transcriptional regulators to control gene expression, and they play additional roles in facilitating the activity of other transcriptional factors. Depending on the context (cell type and developmental stage), GATA proteins display an

important versatility in their modes of action, which allows them to ensure proper regulation during development (M. Tremblay et al., 2018).

(C) Role of GATA factors during mammalian and *Drosophila* development

1- GATA transcription factors are expressed in numerous mammalian and *Drosophila* tissues

GATA transcription factors are expressed in different mammalian tissues, and they have distinct yet partially overlapping expression patterns (Figure 8). Members of the GATA1/2/3 sub-family are expressed in the mesodermal-derived hematopoietic system and kidneys, and in the ectodermal-derived nervous system. Members of the GATA4/5/6 sub-family are expressed in the nervous system and in several other tissues, such as the endodermal gastrointestinal tract and liver, and the mesodermal cardiovascular system and gonads (M. H. Lentjes et al., 2016; Patient & McGhee, 2002). In *Drosophila*, GATA factors are also expressed in different cell types. *grn* is expressed like the GATA1/2/3 sub-family members in the ectodermal-derived neural tissues, while *Pnr* is expressed in the cardiac tissue like GATA4/5/6 sub-family members (Brown & Castelli-Gair Hombría, 2000; Gajewski et al., 1999).

Interestingly, analysis of *srp* expression patterns showed that it is related to both GATA1/2/3 and GATA4/5/6 sub-families, since it is expressed in the mesodermal hematopoietic cells and also in the endodermal developing gut (Rehorn et al., 1996; Reuter, 1994). Serpent is found in

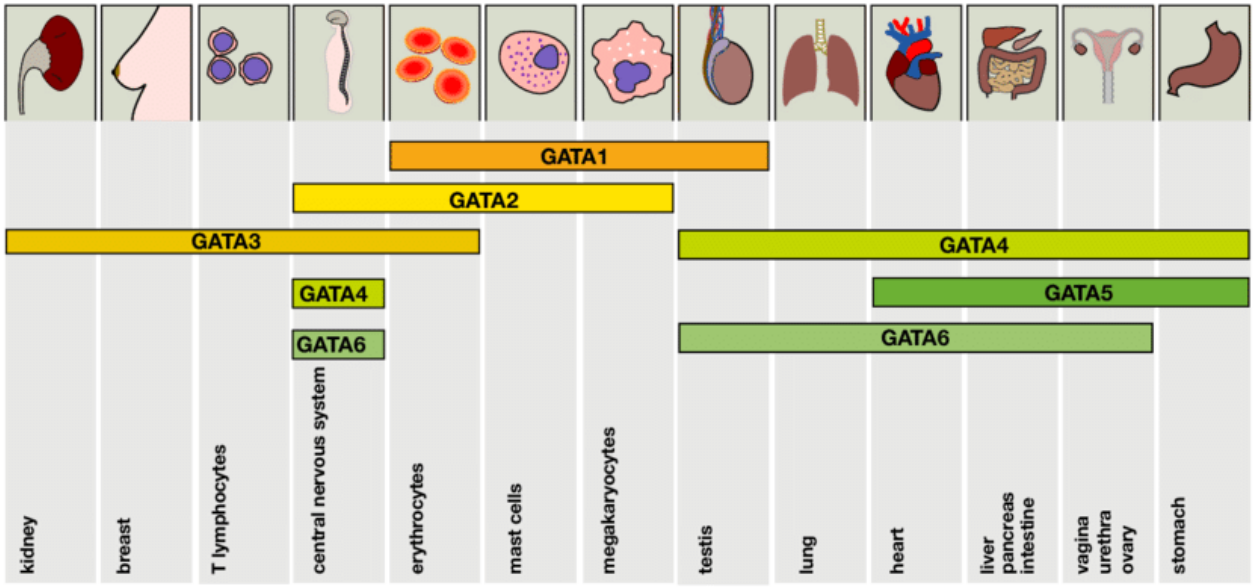


Figure 8. GATA factors expression in various organs during mammalian development.

GATA factors are expressed in numerous mammalian organs during the individual development. GATA1, GATA2 and GATA3 are specifically expressed in hemocytes, kidney and breast while GATA4, GATA5 and GATA6 are specifically found in lungs, heart, digestif and genital system (M. H. Lentjes et al., 2016).

other tissues, namely the fat body (the functional homologue of mammalian liver and adipose tissue) and the ovaries (Lepesant et al., 2020; Rehorn et al., 1996). *GATAe* is expressed in the endodermal derived gut and Malpighian tubules (the equivalent of mammalian kidney) (Martínez-Corrales et al., 2019; Okumura et al., 2005). However, no expression pattern is found for *GATAd* (Okumura et al., 2005). Importantly, the GATA factors cited above are not just expressed in the mentioned tissues, they also ensure essential roles in those tissues.

2- GATA transcription factors have essential functions in mammalian and *Drosophila* organs

a- The hematopoietic system

1- Mammalian and Drosophila hematopoiesis

In mammals, distinct waves of hematopoiesis take place (Figure 9B). The first wave, called primitive hematopoiesis, occurs in the extra-embryonic yolk sac, and gives rise to mainly erythrocytes. The second wave, known as definitive hematopoiesis, happens in the embryo proper in the aorta-gonad-mesonephros (AGM) region, as well as in the extra-embryonic placenta. During definitive hematopoiesis appear self-renewing hematopoietic stem cells (HSCs) that later give rise to all blood cell types. The HSCs colonize and expand in other hematopoietic organs: fetal liver, spleen and thymus and subsequently in the adult bone marrow, the major site of postnatal hematopoiesis (Ivanovs et al., 2017).

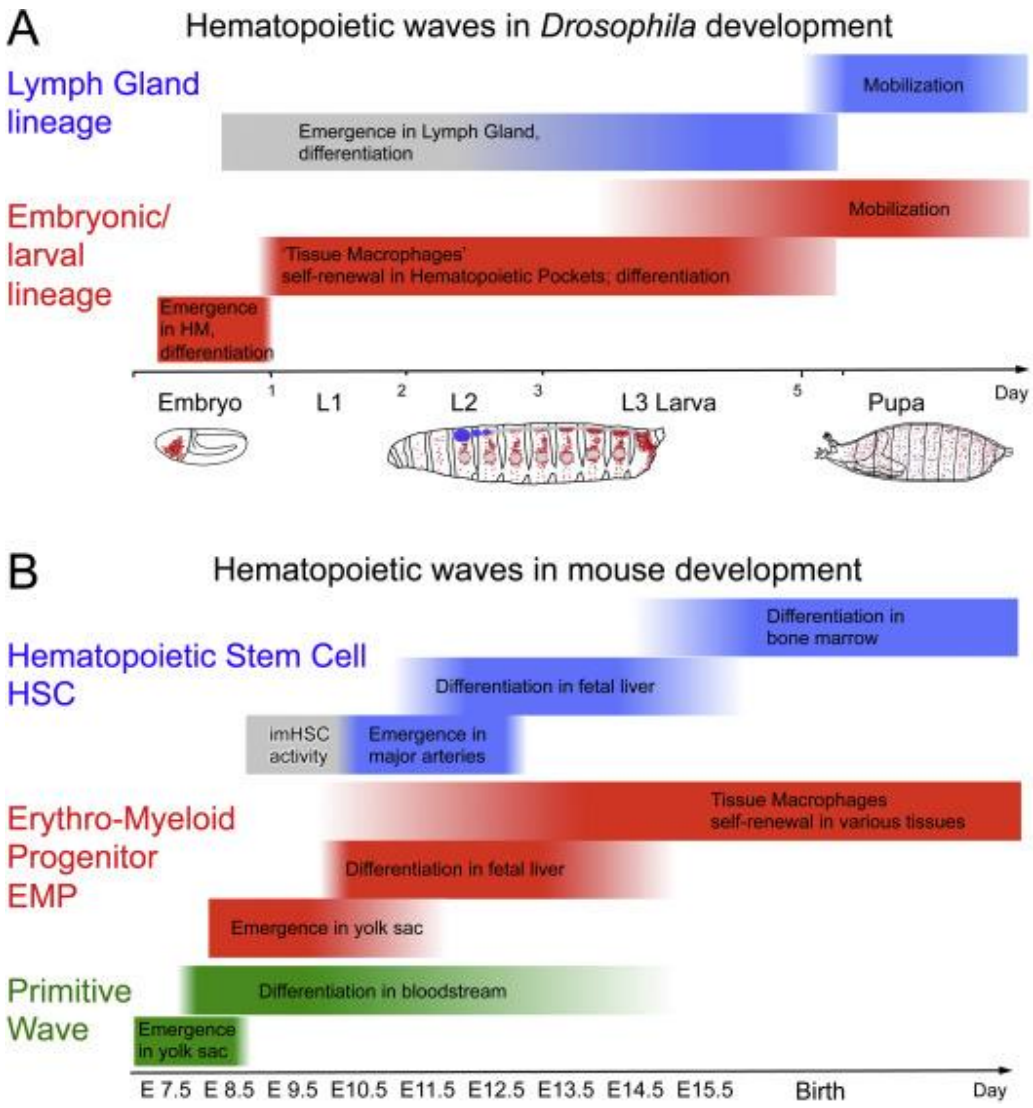


Figure 9. Mammalian and *Drosophila* hematopoietic waves.

(A) Two hematopoietic waves take place during *Drosophila* development (Embryonic/larval and lymph gland depending waves). (B) Three waves of hematopoiesis take place in the developing mouse (Primitive wave, Erythro-Myeloid Progenitor (EMP), and hematopoietic stem cell (HSC) depending waves). L1, L2 and L3: First, second and third instar larval stages, E: embryonic day (Gold & Brückner, 2015).

As in mammals, two distinct hematopoietic waves are present in *Drosophila*, one occurs in the embryo, while the second takes place in a larval hematopoietic organ called the lymph gland (Figure 9A) (Crozatier & Meister, 2007). The embryonic wave begins when the hematopoietic primordium is determined in the embryonic head mesoderm, during early stages of embryogenesis (Figure 10A). Later, these cells differentiate into prohemocytes (blood cell progenitors) that proliferate and differentiate into two mature hemocyte types: crystal cells and plasmatocytes (see below). After their differentiation, plasmatocytes scatter throughout the embryo until the end of embryonic development (Figure 10 B, C) (Tepass et al., 1994). The cells formed during this wave will persist in circulation until adulthood (Holz et al., 2003). The larval hematopoietic wave occurs in a multi-lobed bilateral organ, the lymph gland, which is derived from the embryonic dorsal mesoderm. At embryonic and early larval stages, the lymph gland lobes are formed only of prohemocytes. Larval hematopoiesis takes place mainly during the third instar larval stage. At this stage, the lymph gland is formed of a pair of anterior lobes and several smaller posterior lobes (Figure 11A). The posterior lobes are formed mostly of prohemocytes while the anterior lobes are composed of three compartments: (i) an inner region apposed to the cardiac tube and composed of highly proliferating prohemocytes, called the medullary zone, (ii) a peripheral region formed of differentiated hemocytes, called the cortical zone (Jung et al., 2005), and (iii) a cluster of cells in the posterior region of each anterior lobe called the Posterior Signaling Center (PSC) (Figure 11A, B) (Tim Lebestky et al., 2003), which acts as a microenvironment that participates in the control of lymph gland homeostasis in normal conditions or in response to immune stress (Benmimoun et al., 2015; Oyallon et al., 2016).

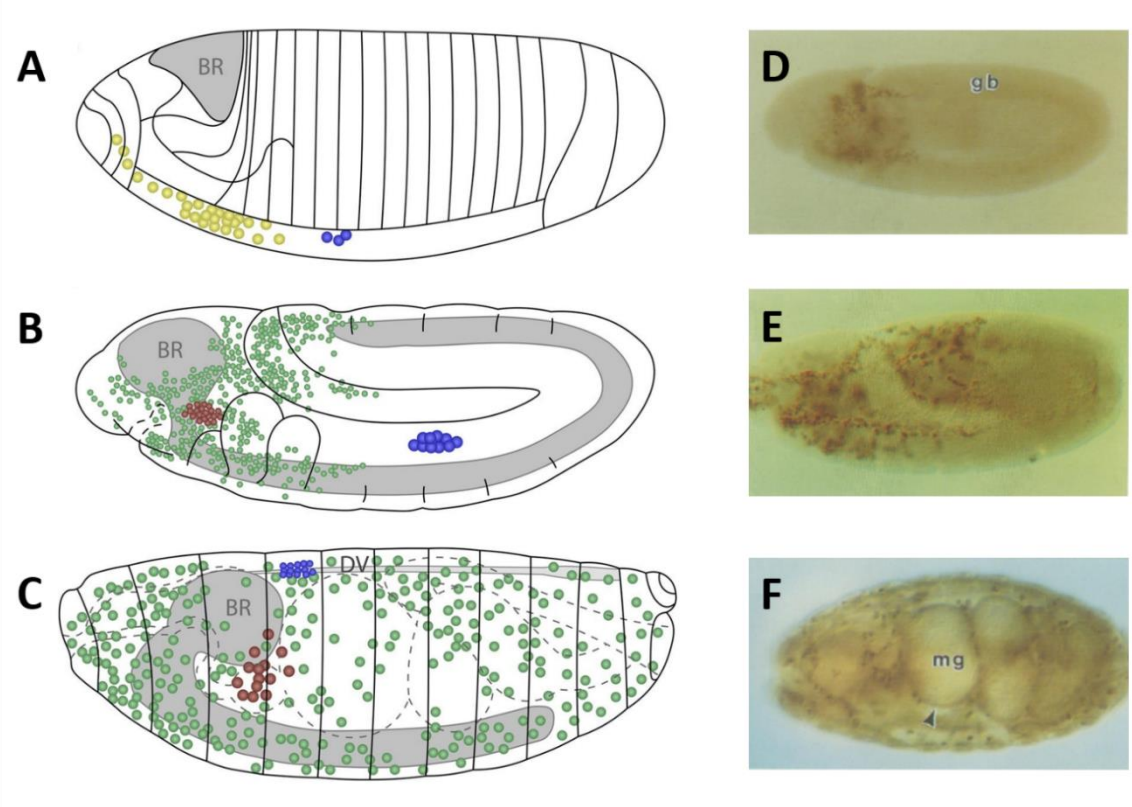


Figure 10. *Drosophila* embryonic hematopoiesis.

Schematic representation of hematopoietic populations during embryonic development (A-C). Anterior to the left, dorsal is up. (A) Precursors of the embryonic hemocytes (yellow) are formed in the head mesoderm while lymph gland precursors (blue) are specified in the thoracic region of the dorsal mesoderm at stage 5 of embryogenesis. (B) At satge 11, the differentiated hemocytes: plasmatocytes (green) and crystal cells (red), migrate in the embryo head and the lymph gland anlage proliferate and reside in the trunk. (C) At stage 17, plasmatocytes are dispersed throughout the embryo while crystal cells accumulate in the anterior part of the embryo and the lymph gland become positioned on either side of the dorsal vessel (DV). BR: Brain (adapted from (Banerjee et al., 2019). (D-F) Immunostaining of whole-mount embryos against Peroxidasin protein shows the plasmatocytes of stage 10 embryo (D), stage 12 (E) and final stages of embryogenesis (F) (adapted from (Nelson et al., 1994).

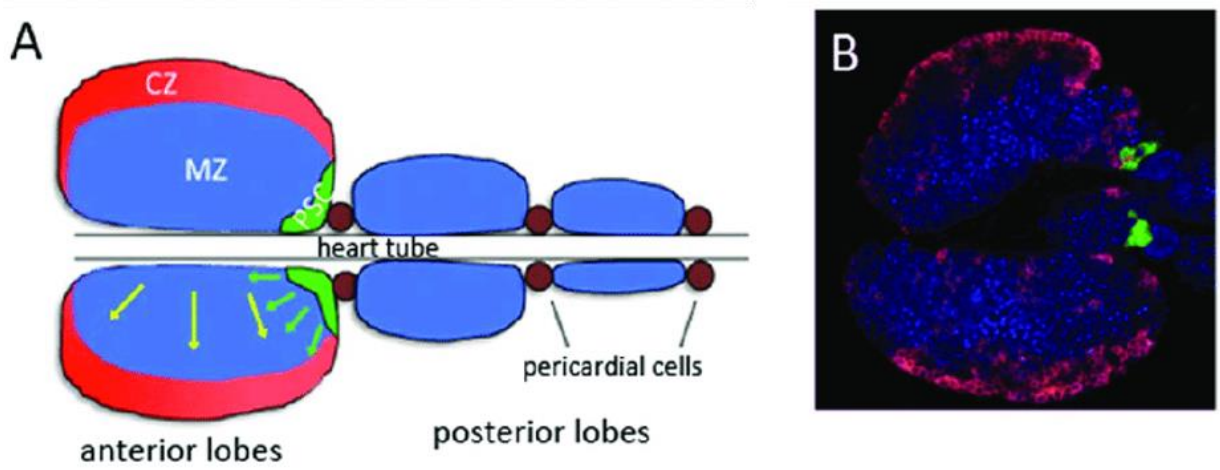


Figure 11. Structure of the *Drosophila* larval lymph gland.

(A) Schematic diagram of the third instar larval lymph gland. The lymph gland flanks the heart tube and it is formed of two large lobes called anterior lobes and several (2 to 4 pairs) smaller lobes called posterior lobes formed of progenitor cells (blue). At each side of the heart tube the lobes are separated by pericardial cells. Each anterior lobe contains a medullar zone (MZ) formed of progenitor cells (blue), of cortical zone (CZ) harboring the differentiated hemocytes (red) and the posterior signaling center (PSC) (green). The green arrows indicate the presumptive PSC-dependent prohemocyte maintenance signals and the yellow arrows represent the main axes of differentiation. (B) Confocal section of third instar lymph gland anterior lobes stained for the PSC (green), blood cell progenitors (blue) and differentiated hemocytes (red) (adapted from (Benmimoun et al., 2015).

Differentiation of the mammalian hematopoietic progenitor cells gives rise to multiple cell types (Figure 12). The erythroid lineage consists of erythrocytes, or red blood cells responsible for transport of oxygen to tissues and for recovery of carbon dioxide produced as waste. The megakaryocytic lineage produces the platelets implicated in blood clotting. The myeloid cell lineages are composed of granulocytes (eosinophils, mast cells and neutrophils), macrophages and dendritic cells, all of which are responsible for innate immune defense. Finally, lymphoid cell lineages regroup the lymphocytes (B-cells, T-cells and natural killer cells) that are the major players of adaptive immunity (Hartenstein, 2006; C. Smith, 2003).

Contrary to mammals, *Drosophila* has neither red blood cells nor lymphoid cells. Only three types of blood cells are produced in the fly (Figure 12). Plasmatocytes are the equivalent of mammalian macrophages and ensure phagocytic functions (apoptotic body removal and micro-organism clearance) (Gold & Brückner, 2015). Crystal cells are the equivalent of granulocytes and owe their name to the presence of crystalline inclusions in their cytoplasm. These inclusions contain melanin precursors that participate in *Drosophila* melanization reactions that allow sequestration and destruction of invading pathogens, as well as wound healing. *Drosophila* crystal cells are also regarded as the equivalent of mammalian megakaryocytes, due to their role in the wound healing (Rizki & Rizki, 1959; Tang, 2009). Finally, under stress conditions (such as after wasp infestation, mechanical stress challenge or genetic deregulation) the fly can produce a third blood cell type: these large flat cells with extended filaments are called lamellocytes (Lanot et al., 2001; R. Sorrentino et al., 2002; Y. Tokusumi et al., 2018). They are found only at larval stages and they surround and destroy particles that are too big to be phagocytosed by plasmatocytes (Wood & Jacinto, 2007).

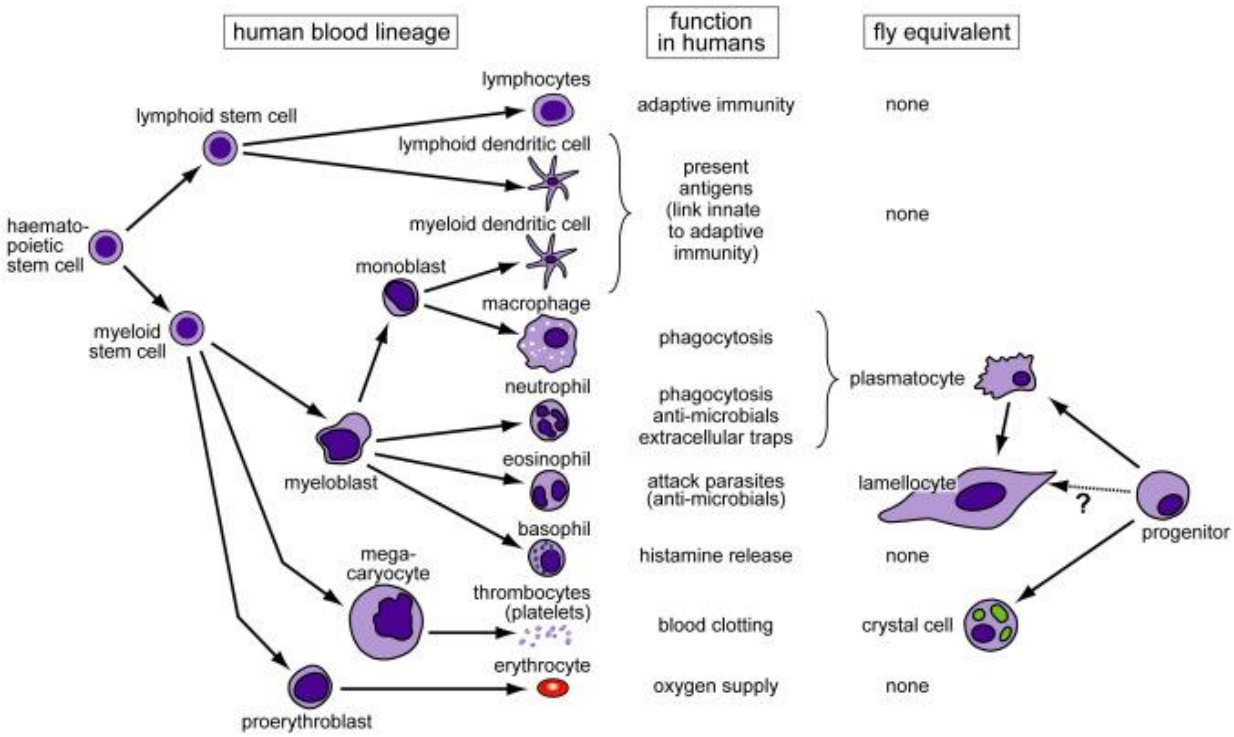


Figure 12. Mammalian and *Drosophila* blood cells.

Different blood cell lineages are produced; they all emerge from hematopoietic stem cells that differentiate later into myeloid or lymphoid stem cells. The lymphoid stem cells give rise to lymphocytes and lymphoid dendritic cell that play a role during the adaptive immunity and antigen presentation, respectively. The myeloid stem cells give rise to more restraint hematopoietic precursors (myeloblast, monoblast, pro-erythroblast) that will be at the origin of several types of mature blood cells: myeloid dendritic cell, macrophage, neutriophil, eosinophil, basophil and erythrocyte. Megakaryocytes are also derived of myeloid stem cells and are the origin of thrombocytes or platelets. The function of each cell lineage in human is added beside the corresponding cell type. In *Drosophila*, three blood cell types are produced: plasmatocyte represents macrophage-like phagocyte, crystal-cell is megacaryocytes-like cell that participate in blood clotting and lamellocyte constitutes an encapsulating hemocyte that participate in the parasites' attack (« Organs », 2015).

2- Role of GATA factors during mammalian and Drosophila hematopoiesis

GATA1 and GATA2 play essential functions during mammalian primitive hematopoiesis. Although not required for erythroid precursor cell formation in the murine embryonic yolk sac, GATA1 is crucial for the differentiation of these precursors into mature erythrocytes (McDevitt et al., 1997; Pevny et al., 1991, 1995; Takahashi et al., 1997). GATA1 loss-of-function causes a blockage of the yolk sac erythropoiesis at the pro-erythroblast stage, followed by apoptosis of the pro-erythroblasts (Y. Fujiwara et al., 1996; Pevny et al., 1995). Contrary to GATA1 null mice that display altered mature erythrocyte formation, mice with GATA2 loss-of-function have somewhat normal maturation of red blood cells, but the number of mature erythrocytes is reduced two- to seven-fold (F.-Y. Tsai et al., 1994). Thus, GATA2 and GATA1 proteins regulate proliferation and viability of embryonic developing blood cells.

Contrary to the primitive hematopoiesis, the mammalian definitive hematopoiesis relies on the presence of pluripotent HSCs that give rise not only to erythrocytes but also to megakaryocytes, myeloid cells and lymphoid cells. Interestingly, GATA1, GATA2 and GATA3 play essential functions throughout the development of the hematopoietic lineages of this post-embryonic hematopoiesis (T. Fujiwara, 2017; Ho et al., 2009; Katsumura & Bresnick, 2017). First, GATA3 participates in the maintenance of the number of long-term repopulating HSCs, their entry into the cell cycle, and their subsequent proliferation (Frelin et al., 2013; Ku et al., 2012). Second, GATA1 and GATA2 are essential for differentiation of HSCs into myeloid cells. In the granulocyte-monocyte progenitors, GATA2 determines the mast cells fate, while in the megakaryocyte-erythrocyte progenitors, GATA2 promotes megakaryocyte differentiation at the expense of erythrocytes (Ikonomi et al., 2000; F. Y. Tsai & Orkin, 1997). In parallel,

GATA1 proteins activate differentiation of common myeloid progenitors into erythrocytes, megakaryocytes, eosinophils, mast cells and dendritic cells (Gutiérrez et al., 2007; Harigae et al., 1998; Hirasawa et al., 2002; Shivdasani et al., 1997; Vyas et al., 1999; Weiss & Orkin, 1995).

Drosophila embryonic hematopoiesis is similar to the initial waves of mammalian hematopoiesis that gives rise to early and intermediate progenitors. Like in mammals, the fly GATA factors play essential functions during the first wave of hematopoiesis. From early stages of embryogenesis on, *srp* is expressed in embryonic prohemocytes, where it is required for the proliferation and survival of these developing hematopoietic cells. The loss-of-function of *srp* is associated to a loss of all embryonic hemocytes: plasmatocytes and crystal cells (T. Lebestky et al., 2000; Rehorn et al., 1996; Sam et al., 1996).

In the fly, the lymph gland is the main site of post-embryonic hematopoiesis, and it is characterized by the specification and the maintenance of progenitor cells that constitute the source of all mature hemocytes. The GATA factor Srp is expressed in the lymph gland hematopoietic progenitor cells, where it plays an essential function in the specification of the hematopoietic fate. Indeed, in the absence of Serpent activity, while lymph gland progenitor cells are formed in their expected position around the cardiac tube, they show inappropriate expression of the pericardial cell marker, *pericardin* (Frandsen et al., 2008; Han & Olson, 2005; Jung et al., 2005; Mandal et al., 2004). In addition, Pnr (another *Drosophila* GATA factor) is required in a cell-autonomous manner during differentiation of lymph gland progenitors into mature plasmatocytes (Minakhina et al., 2011).

Members of the FOG family have also essential roles during hematopoiesis, as they participate in the determination of hematopoietic cell fate. Murine FOG1 is required for the GATA1-mediated activation of mammalian eryhthropoiesis and megakaryopoiesis. Its interaction with GATA1 limits mast cell and eosinophil formation (Cantor et al., 2008; Chang et al., 2002; J. D. Crispino et al., 1999; Querfurth et al., 2000). Similarly, in *Drosophila*, Ush has two opposing roles during the differentiation of mature blood cells. Ush promotes the differentiation of prohemocytes into mature crystal cells and plasmatocytes in the lymph gland; however, it inhibits their differentiation into lamellocytes (Avet-Rochex et al., 2010; Dragojlovic-Munther & Martinez-Agosto, 2013; H. Gao et al., 2009).

Finally, at the molecular level, it has been shown that human GATA1 promotes expression of megakaryocyte specific genes, due to its interaction with the transcription factor of the RUNX family, RUNX1 (Elagib et al., 2003). Interestingly, this physical and functional interaction has been conserved throughout evolution, since Srp was shown to collaborate with Lozenge (the *Drosophila* homolog of RUNX1) to promote differentiation of *Drosophila* megakaryocyte-like crystal cells (Waltzer et al., 2003).

b- The cardiac system

1- Mammalian and Drosophila cardiogenesis

The mammalian heart derives from the mesoderm. During early stages of mammalian development, primitive cardiac cells develop in the splanchnic mesoderm as two tubes formed

from different cardiac precursors, called primary and secondary heart fields. The primitive heart tubes then merge into one single tube called the linear heart tube (Brade et al., 2013). This linear heart tube swells and forms various anatomic features of the heart. Later, the cardiac tube twists and turns on itself by a process called heart looping, to form a heart with four chambers: right ventricle, left ventricle, right atrium and left atrium (Figure 13A) (Bruneau, 2013; Schleich et al., 2013).

The mammalian heart is composed of three tissue layers. The first layer (the endocardium) consists of endothelial cells underlying connective tissue, and it represents the innermost layer of the heart. The second layer (the myocardium) is composed of cardiomyocytes that are responsible for heart contractions. The third layer (the epicardium) comprises epithelial cells and constitutes the outermost layer of the heart. The epicardium receives the nervous and blood vessels that supply the heart, and it surrounds a pericardial cavity containing a serous fluid that prevents friction during heart contractions (Gavaghan, 1998).

In order to establish a directional blood flow, the heart chambers are separated by septa and cardiac valves (Figure 13A). The atrial septum separates the right and left atria, while the ventricular septum separates the right and left ventricles. The atrium and ventricle of each side of the heart are separated by atrioventricular valves (Krishnan et al., 2014; Lamers Wouter H. & Moorman Antoon F.M., 2002).

Compared to the vertebrate heart, the *Drosophila* heart is quite simple: it is basically a simple tube of bilateral cardiomyocytes (Figure 13B). Fly cardiac cell precursors form 11 clusters of cells located along the most dorsal part of the embryonic mesoderm (Gajewski et al., 1999). By

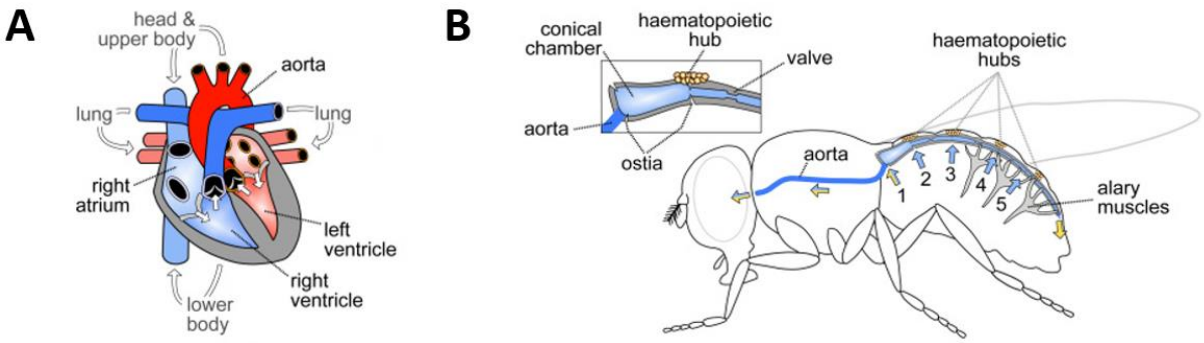


Figure 13. Mammalian and *Drosophila* heart.

(A) The mammalian heart is formed of four chambers: right and left ventricle plus right and left atrium. Chambers are separated by septa and cardiac valves (white arrows) that control blood circulation direction into the heart. The transport of the blood from the heart to the other body organs and *vice versa* is occurred through heart vessels. The aorta is the vessel that transport oxygenated blood from the heart towards the upper and lower body organs. (B) The fly heart is formed of simple tube of bilateral cardiomyocytes located at the dorsal midline. The regulation of the *Drosophila* hemolymph flow direction in the heart depends on ostia and valves. The anterior portion of the heart is named conical chamber from which is attached the fly aorta. The alary mucles are associated with the heart likely to give it lateral stability and support (adapted from (« Organs », 2015)).

the end of embryogenesis, the bilateral heart progenitor cells have migrated to the dorsal midline, where they form a linear heart tube (Figure 14A-C).

The fly cardiac tube consists of only two cell types: cardioblasts that differentiate into contractile cardiomyocytes, and pericardial cells that flank the cardioblasts and act as nephrocytes (Vogler & Bodmer, 2015). In *Drosophila*, the function of the cardiac tube is to pump the hemolymph containing essential components such as nutrients, signaling molecules and hemocytes, from the posterior part of the body towards the anterior part. Like in mammals, regulation of the flow direction in the fly heart is achieved by ostia and valves present in the heart tube (Figure 13B). At the adult stage, the ostia form 5 pairs of openings distributed along the heart which allow the hemolymph to enter but not to exit the cardiac tube, while the valves are formed of intracardiac specialized cardiomyocytes that can close off the heart tube at five different locations (Rotstein & Paululat, 2016).

2- Role of GATA factors in mammalian and Drosophila cardiogenesis

In mammals, the transcription factors of the GATA family including GATA4 and GATA6 play essential functions during heart formation (Clowes et al., 2014; Peterkin et al., 2005). Although the loss-of-function of either GATA4 or GATA6 alone has no effect on differentiation of mammalian cardiomyocytes, mutant mice with simultaneous loss-of-function of GATA4 and GATA6 are devoid of a heart, which results from defects in cardiomyocyte differentiation (Kuo et al., 1997; Molkentin et al., 1997; Zhao et al., 2008). Therefore, GATA4 and GATA6 factors

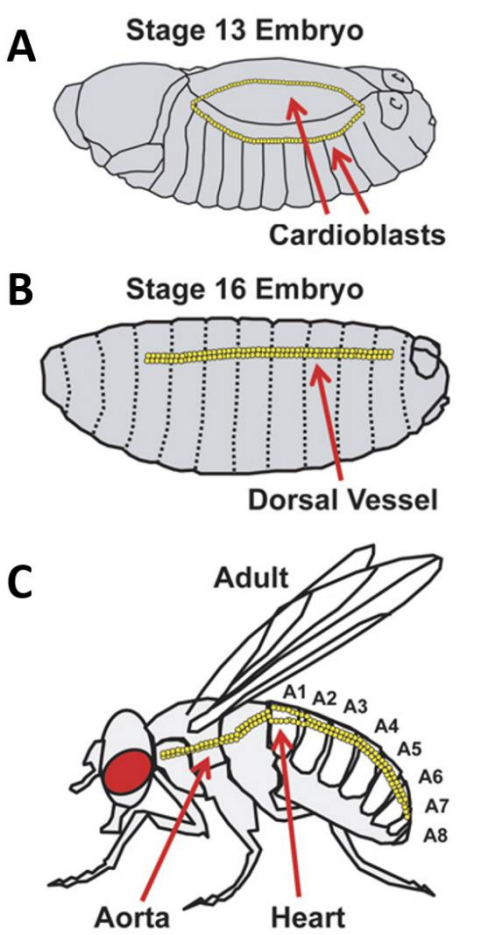


Figure 14. The *Drosophila* heart development.

(A-C) Schematic representation of heart development. Anterior is to the left and the dorsal is up. (A) The heart is formed from two lines of cardioblasts migrating toward the dorsal midline of stage 13 embryo. (B) At the stage 16, the embryonic dorsal vessel is located in the dorsal midline. (C) The adult fly circulatory system is formed of the aorta located in the thoracic part of adult fly and of the heart present in the abdomen. The adult abdominal segments (A1 to A8) are shown (adapted from (Wolf, 2012).

have redundant roles during heart development. Similar to the common role of GATA4 and GATA6 in mammalian heart development, Pnr has critical functions during *Drosophila* cardiac cell specification, and its loss-of-function is associated to an almost complete absence of cardiomyocytes (Klinedinst & Bodmer, 2003).

Interestingly, it was shown that activation of gene expression in cardiac cells is under the control of rat GATA4 and the homeotic box transcription factor, Nkx2-5, which both interact physically to promote the cardiomyocyte program (Durocher et al., 1997). Similarly, in *Drosophila*, it was found that Pnr activates the expression of cardiac cell markers such as *myocyte enhancer factor 2* (*mef2*), which also depends on the presence of the Nkx2-5 protein homologous Tinman (Tin); both Pnr and Tin are able to physically interact in cultured cells (Gajewski et al., 1999, 2001).

c- The gastro-intestinal system

1- Mammalian and Drosophila intestinal development

In mammals, gut morphogenesis begins when the anterior and posterior edges of the endodermal epithelial sheet fold off to form the foregut and the hindgut pockets. Once folded, both pockets continue to migrate towards each other in order to meet in the center of the embryo and to form a closed tube, the primitive gut tube (Spence et al., 2011). The digestive tube is subdivided into three regions: foregut, midgut and hindgut, which will give rise to different adult intestinal structures. The foregut gives rise to structures including the esophagus, liver and pancreas. The midgut yields the stomach and the small intestine, and the hindgut forms the

colon. The small intestine is in turn divided into three regions: duodenum, jejunum, and ileum (Figure 15A) (Spence et al., 2011).

The regions of the mammalian intestine differ in their functions. The jejunum's role is to digest and absorb important nutrients such as fatty acids, amino acids and sugars. The other small intestine regions also participate in the absorption of these nutrients, however, they have additional specific functions, like the absorption of iron in the duodenum and the absorption of vitamin B12 and bile salts in the ileum (Aronson et al., 2014). In contrast, in the colon, the epithelial cells absorb only water and mineral ions. The colon also contains bacteria that allow the fermentation of indigestible materials (Greenwood-Van Meerveld et al., 2017). The intestinal epithelium is formed of repeated crypt-villus units that contain different types of intestinal cells, such as the enterocytes (that play a role in nutrient absorption) and the secretory cells (enteroendocrine, goblet, tuft and Paneth cells). In addition, in the intestinal crypt, there are quiescent intestinal stem cells, rapidly dividing stem cells and proliferating transit-amplifying cells that differentiate into mature intestinal cells as they migrate to the base of or outside the crypt (Figure 15A) (Noah et al., 2011).

Similar to the mammalian digestive tube, the *Drosophila* gut is composed of three regions: two ectodermal-derived regions, namely the anterior gut (foregut) and the posterior gut (hindgut), and an endodermal-derived region called the midgut (Figure 15B). The embryonic midgut is formed from two different primordia, the anterior midgut primordium (that appears at early stages of embryogenesis at the anterior-ventral side of embryonic blastoderm) and the posterior midgut primordium (that is formed at the posterior side of embryonic blastoderm) closely related to the hindgut ectodermal cells (Figure 16A). During embryonic germ band elongation,

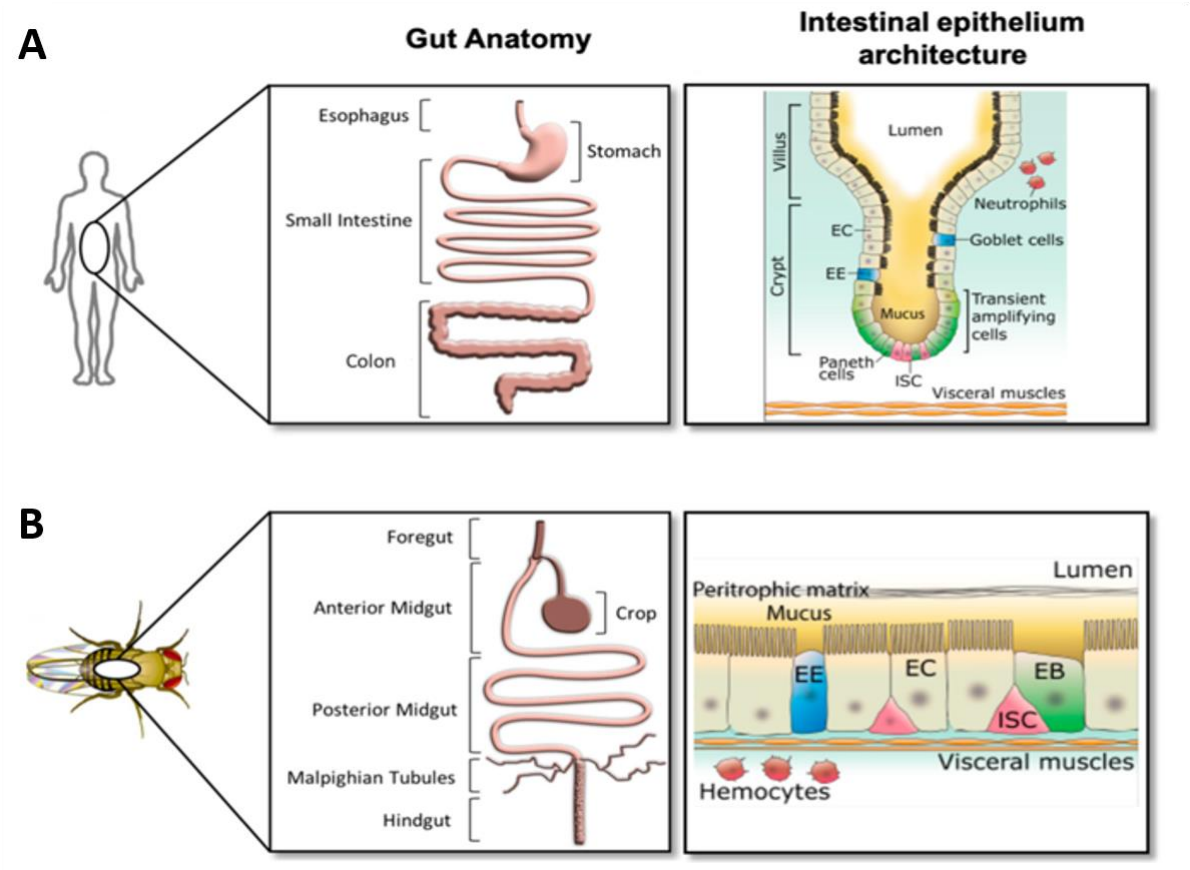


Figure 15. Mammalian and *Drosophila* gut.

(A) The mammalian intestine is divided into four regions: esophagus, stomach, small intestine and colon. The intestinal epithelium is formed of series of crypt-villus structure that are composed of different intestinal cell types: absorptive erythrocytes (EC), secretory cells (enteroendocrine (EE), goblet cells, Paneth cells), intestinal stem cells (ISC) and transit-amplifying cells. The epithelium is surrounded by visceral muscles and also by mucus that prevent direct contact between intestinal epithelial cells and intestinal bacteria located in the lumen. Neutrophils circulate in proximity of the intestinal epithelium (B) The fly gut is divided into foregut, midgut (anterior and posterior midgut) and hindgut. The intestinal epithelium is formed of EC, EE, ISC and erythroblasts (EB). The epithelium is surrounded by visceral muscles and by peritrophic matrix mucus, which is the equivalent of the human mucus. Hemocytes can be found close to the intestinal epithelium. The Malpighian tubules in the fly are the equivalent of the mammalian kidney (Capo et al., 2019).

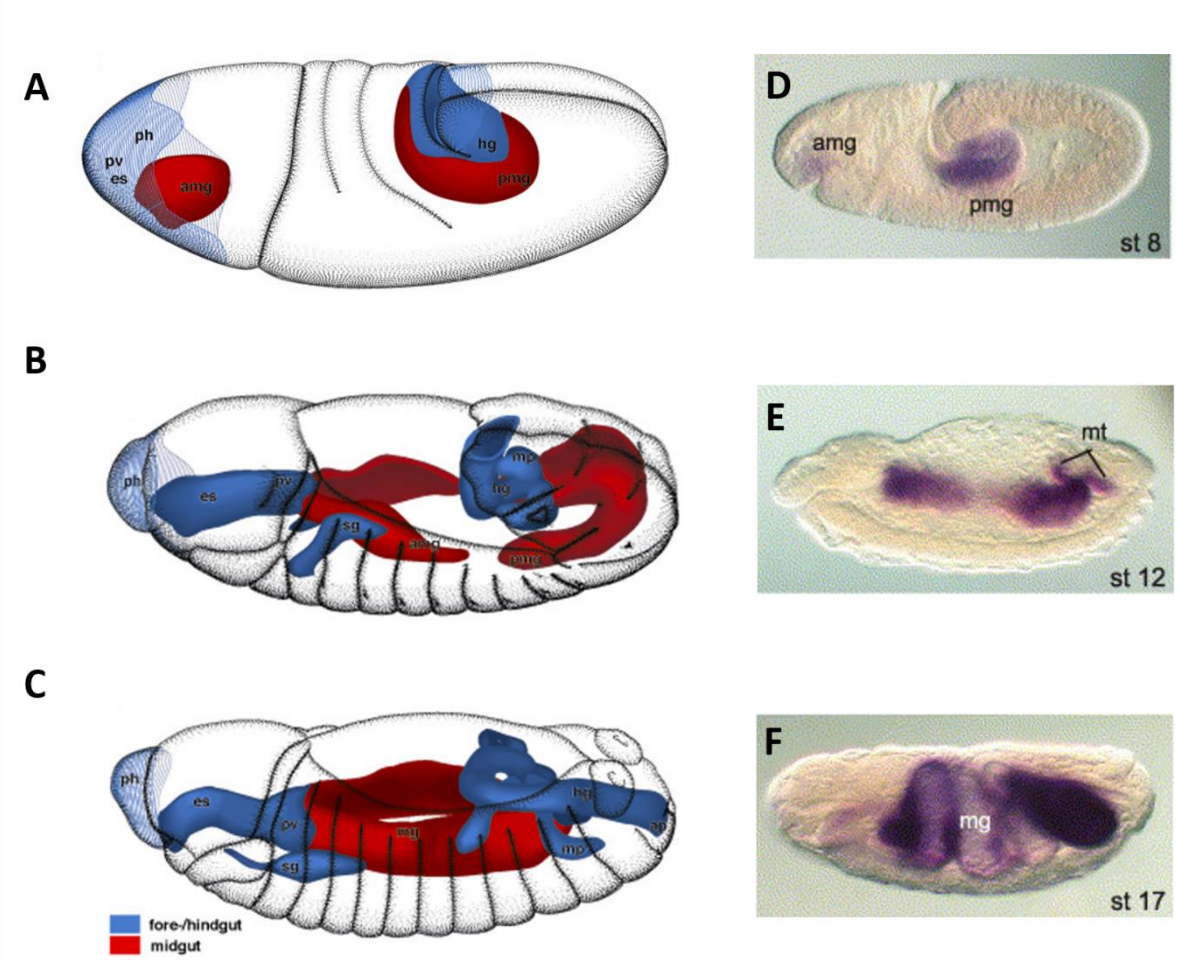


Figure 16. The *Drosophila* gut formation.

(A-C) Schematic representation of embryonic gut formation. The embryonic anterior side is to the left while the dorsal is up. (A) At stage 7 of embryogenesis, the endodermal (red) anterior midgut primordium (amp) and posterior midgut primordium (pmg) are located in the embryonic anterior and posterior parts, respectively. The ectodermal (blue) anlagen that surround the amp and the pmg are at the origin of the foregut and hindgut respectively. (B) At stage 12, the amg and pmg are migrating one toward the other. (C) At stage 13, the amp and the pmg cells fuse to form a continuous gut tube. *ph*: pharynx, *es*: esophagus, *pv*: proventriculus, mp: malpighian tubules, sg: salivary glands (adapted from (Hartenstein, 1993). (D-F) In situ hybridization of whole-mount embryos against *GATAe* transcript shows the amg and the pmg of stage 8 embryo (D), the migrating amp and pmg and developing Malpighian tubues (mt) of stage 12 embryo (E) and the closed multi-chambered midgut at the stage 17 of embryogenesis (F) (adapted from (Okumura et al., 2005).

the anterior and posterior midgut primordia move towards each other, and once they meet in the middle of the embryo, undergo mesenchymal to epithelial transition and fuse to one another, giving rise to the gut tube at the end of embryogenesis (Figure 16B, C). Although constituted of a simple monolayer epithelium, the midgut can be divided, as the mammalian small intestine, into different regions that differ from each other by histological and anatomical characteristics and by gene expression patterning (Buchon et al., 2013; Ryutaro Murakami et al., 1994).

As in vertebrates, the *Drosophila* gastrointestinal tract plays a major role in nutrient digestion and molecule absorption (Lemaitre & Miguel-Aliaga, 2013; Miguel-Aliaga et al., 2018). The adult midgut epithelium contains four cell types, two of them representing mature gut cells, called enterocytes and enteroendocrine cells. Enterocytes have a major role during nutrient absorption and digestion, while enteroendocrine cells are responsible for secretion of peptide hormones to regulate gut physiology and homeostasis (Miguel-Aliaga et al., 2018). The other two cell types are the intestinal stem cells and their daughters (the enteroblasts) that are able to constantly regenerate midgut epithelium and to repair tissue damage (Figure 15B) (Jiang et al., 2011).

2- Role of GATA factors in mammalian and Drosophila intestine

Members of the GATA family of both *Drosophila* and mammals have similar roles during intestine development. In mice, GATA4, GATA5 and GATA6 are expressed in embryonic intestinal cells (Arceci et al., 1993; Fang et al., 2006; E. E. Morrisey, Ip, Tang, Lu, et al., 1997; Edward E. Morrisey et al., 1996). Among them, GATA5 has no essential functions during gut morphogenesis and differentiation (Molkentin et al., 2000). In contrast, both GATA4 and

GATA6 have redundant and crucial roles during gut development (Kuo et al., 1997; Molkentin et al., 1997; Walker et al., 2014). Mice with simultaneous loss-of-function of both GATA factors have altered developing intestinal epithelium and intestinal villus architecture (Walker et al., 2014). Similarly, in *Drosophila*, three GATA factors Srp, Grn and GATAe are exrepssed in developing embryonic midgut primordia. Among these factors, only Srp is required for midgut primordia specification and subsequently for gut formation (Hernández de Madrid & Casanova, 2018; Lin et al., 1995; Okumura et al., 2005; Reuter, 1994; Sam et al., 1996).

In adult mouse small intestine, GATA4 is expressed in the duodenum and in the jejunum but not in the ileum. GATA4 maintains jejunum identity, as loss-of-function of GATA4 in the jejunum causes a loss of jejunal specific gene expression and inadequate expression of ileal specific genes in jujenal enterocytes (Bosse et al., 2006; Fang et al., 2006). This altered gene expression affects the ability of the intestine to absorb cholesterol and dietary fat, which are two jejunal-specific functions, and causes an increase of bile acid absorption, which is an ileal-specific function (Battle et al., 2008; Bosse et al., 2006). Contrary to GATA4, whose expression profile is restricted to the proximal intestine, GATA6 is expressed throughout the small intestine including the jejunum (Fang et al., 2006). Mice with inducible disruption of GATA6 in the small intestine, display reduced expression in the ileal enterocytes of some genes normally expressed in the ileum, and mainly of those involved in lipid metabolism, such as *apolipoprotein CIII*, *apolipoprotein A-I*, and *fatty acid binding protein 6*. In parallel, GATA6 inactivation in the ileum induces the expression of genes normally expressed in the colon (Beuling et al., 2011). Thus, GATA6 is required for activation of ileum specific gene expression and inhibition of colon specific genes transcription in the ileal enterocytes. The role of members of GATA family in regulation of the intestinal regionalization has also been identified in *Drosophila*. Srp acts as a homeotic gene during early stages of embryogenesis in order to

prevent endodermal anterior midgut primordium from adopting ectodermal foregut fate, and to specify endodermal posterior midgut *versus* ectodermal hindgut fate (Reuter, 1994). Similarly, GATAe maintains endodermal identity of the midgut by inhibiting expression of the prospective hindgut determinant *brachyenteron* in the posterior midgut primordium (Okumura et al., 2005). Later, in embryonic developed intestine, GATAe plays important functions in the expression of terminally differentiated pan-midgut markers, namely *innexin* and *CG4781*, and regionalized midgut markers such as *integrin-β_v, midgut expression 1*, *Tetraspanin 29Fa*, *CG5077*, *CG10300*, *lambda-Trypsin*, *CG17633* and *CG18493* (Okumura et al., 2007). In a similar manner, in the *Drosophila* adult midgut, GATAe regulates most of the genes expressed in a regionalized manner in the midgut, including genes encoding proteins involved in digestion such as the Amylase distal, beta-galactosidase and the trypsin family enzymes (theta-Trypsin, iota-Trypsin and CG31269) (Buchon et al., 2013). All these results show that *Drosophila* Srp and GATAe proteins are required, as are the mammalian GATA4 and GATA6, for determining intestinal region identities in the gut, and for regulating expression levels of essential genes in gut enterocytes.

Furthermore, murine GATA6 proteins regulate the proliferation of cells at the bottom of the colonial crypt and their subsequent migration from the bottom to the villus surface. The proliferation of the crypt intestinal stem cells and their subsequent migration and differentiation into mature intestinal cells is essential for colonic epithelial renewal (Beuling et al., 2012). Similarly, GATAe is required for the maintenance of intestinal stem cells, by regulating their proliferation and other stem cells properties, such as their small nuclei size and diploid state (Okumura et al., 2016). By regulating proliferation and stemness of intestinal stem cells and consequently their differentiation into intestinal mature cells, GATAe is identified, like

GATA6, as a crucial factor for the continuous regeneration of the adult midgut during normal fly development, but also during stress conditions (Dutta et al., 2015; Okumura et al., 2016).

Differentiation of mature intestinal cells depends on activation of the Notch signaling pathway in intestinal stem cells. GATAe contributes to Notch signaling pathway activation in *Drosophila* adult intestine by maintaining expression of the Notch receptor ligand Delta in intestinal stem cells. Alteration of GATAe-mediated Notch activation in intestinal stem cells is related to a high reduction rate of both enterocytes and enteroendocrine differentiation (Okumura et al., 2016). Strikingly, the double knockout of GATA4 and GATA6 in mammalian developing intestine is also associated to reduced expression of Delta and of Delta mediated activation of the Notch signaling pathway in intestinal cells. This dis-regulation of the GATA4 and GATA6-mediated Notch activation in mammalian intestine, alters the intestinal epithelial cell populations and provokes a significant reduction in the differentiation rate of enterocytes and enteroendocrine cells (Walker et al., 2014).

Finally, members of GATA family have essential roles in the epithelial-to-mesenchymal transition (EMT), which is a process required for cell migration. It occurs during formation of the fly embryonic intestine, where the anterior and posterior midgut primordia gradually lose their epithelial properties to become more rounded and irregular in shape. They finally adopt mesenchymal properties, allowing them to migrate towards the center of the embryo, where they form a closed gut tube (K. Campbell et al., 2011; Leptin, 1995; Reuter, 1994). Srp is required for fly intestinal EMT (K. Campbell et al., 2011; Reuter, 1994). This role depends on direct repression by Srp of *crumbs* expression in these cells, which results in relocalization of junctional protein dE-cadherin without affecting its expression. Interestingly, it was found that

vertebrate GATA6 factor promotes EMT of mammalian cells in a similar manner to that of Srp, and this by inhibiting expression of the mammalian *crumbs* ortholog *crb2*, and by provoking a downregulation of junctional dE-cadherin proteins without blocking their expression (K. Campbell et al., 2011). Thus, Srp and GATA6 factors acts similarly during EMT in mammalian and *Drosophila* cells.

d- The ovaries

1- Mammalian and Drosophila oogenesis

Mammalian gonads develop from mesonephros when the epithelium overlaying the mesonephros coelomic surface starts to swell and to form the presumptive gonads, the genital ridge. Primordial germ cells are specified at the base of the extra-embryonic allantois. These cells migrate towards the genital ridge to later populate the primordial gonads (DeFalco & Capel, 2009).

Mammalian ovaries contain germ cells representing all the cells that will undergo gametogenesis, ranging from oogonia to eggs, as well as somatic cells that form the structures harboring and helping the primordial germ cells during their development, namely the ovarian follicles. Mammalian ovarian follicles are constituted of granulosa cells, surrounded by steroidogenic cells called theca cells that form a part of the ovarian stroma (J. Wu et al., 2016).

In *Drosophila* as well, the gonads develop from two different cell types: the primordial germ cells that form one cell cluster, and the somatic gonadal precursors that form three cell clusters in the mesoderm of developing embryo (Brookman et al., 1992; Mahowald, 1962). During embryonic germ band elongation, primordial germ cells are passively moved from the most posterior end of the embryo to the dorso-posterior part, and the cells are later pulled in an embryonic dorsal invagination. Germ cells then pass through the invaginating epithelium towards the mesoderm in the direction of the somatic gonadal precursors, which migrate and coalesce to then ensheath the primordial germ cells. Finally, both germline and somatic cells compact together to form a round embryonic gonad (Santos & Lehmann, 2004).

Fly ovaries are composed of egg chambers (Figure 17) that each contain 15 small nurse cells and one larger oocyte. Nurse cells and oocyte all derive from division of ovarian germline stem cells. The egg chamber is surrounded by an epithelial monolayer of follicle cells that provide the oocyte with yolk proteins, a source of nutrients essential for oocyte growth (DiMario & Mahowald, 1987; X. Wu et al., 2008). Yolk proteins are either expressed in the female fly fat body and transported by the hemolymph to the follicle cells, or directly expressed in the ovarian follicle cells (Barnett et al., 1980; Hames & Bownes, 1978).

2- Role of GATA factors in mammalian and Drosophila ovaries

GATA4 is highly expressed in the anterior part of mammalian mesonephros coelomic epithelium, and to a lesser degree in the posterior coelomic epithelium. GATA4 is required for coelomic epithelium thickening that occurs mostly in the anterior part of the mesonephros, in order to constitute the genital ridge and hence, the gonadal primordium (Hu et al., 2013). In

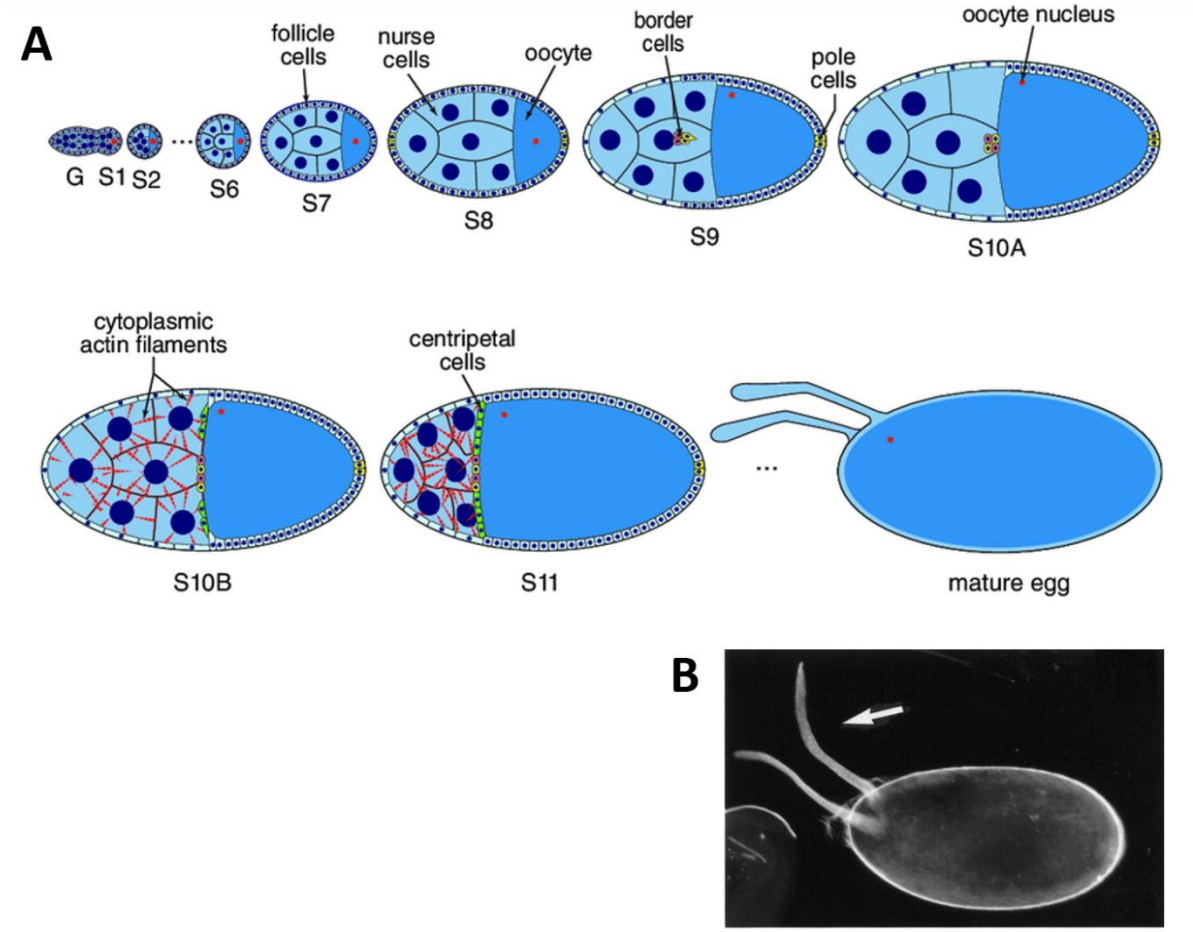


Figure 17. The *Drosophila* oogenesis.

(A) Egg chamber representation from selected stages of oogenesis. Each egg chamber is formed of an oocyte and several supporting nurse cells and is surrounded by a monolayer of somatic cells (follicle cells). The border cells and pole cells form distinct specialized follicular cells. During late stages of oogenesis, dynamic actin remodeling occurs within the nurse cells of the egg chamber, centripetal cells separating the oocyte from nurse cells move inward and the oocyte volume increases significantly and progressively in order to give rise at the end of oogenesis to a mature egg (adapted from (E. Fedorova et al., 2018). (B) Image of a mature wild-type egg with the two dorsal appendages are pointed out by an arrow. Anterior part of the embryo is to the left while the dorsal is up (adapted from (Dong et al., 1999).

contrast, no GATA factors have been identified as being expressed in developing *Drosophila* germline and somatic cells (Lin et al., 1995; Okumura et al., 2005; Rehorn et al., 1996; Winick et al., 1993).

In the developing murine ovaries, GATA4 is expressed in somatic granulosa cells, where it promotes granulosa cell proliferation and ovarian follicle formation (Efimenko et al., 2013; Viger et al., 1998). In addition, loss of GATA4 expression in the ovary leads to a reduction of ovary size, of the number of oocytes released, of the level of estrogen produced, and of the expression of the GATA4 ovarian target genes *Star*, *Cyp11a1* and *Cyp19* (Kyrönlahti et al., 2011). Similar or even stronger phenotypes have been observed in the ovary of mice mutants, with simultaneous loss of GATA4 and GATA6 functions (Bennett et al., 2012; Padua et al., 2014). Thus, GATA4 and GATA6 play essential functions during mammalian gonadogenesis, folliculogenesis and oocyte release. In *Drosophila*, GATA factors Srp and GATAd are expressed in *Drosophila* adult ovaries (Lepesant et al., 2020). Srp plays essential functions in ovarian somatic cells that ensure *Drosophila* oogenesis, as female flies with downregulated Srp in somatic cells lay almost no eggs, and their ovarioles contain degenerating mid-stage egg chambers, revealed by the pycnotic morphology of their nuclei (Lepesant et al., 2020). In addition, it has been suggested that Srp is responsible, in the ovarian follicle cells, for the expression of yolk proteins that are required for oocyte nutrition (Lossky & Wensink, 1995). Thus, *Drosophila* Srp has, as vertebrate GATA4 and GATA6, essential functions during fly oogenesis.

e- The liver and the fat body

1- Mammalian liver and Drosophila fat body formation

The liver derives from the foregut endoderm, where progenitor cells destined to adopt the hepatic cell fate are specified. These cells later converge to generate the epithelial cells of the liver bud (K. D. Tremblay & Zaret, 2005). During gestation, the liver develops into a hematopoietic organ. However, after birth, the liver shifts from a hematopoietic role to being the primary site controlling levels of many metabolites and serum proteins, in the bloodstream and for endotoxin detoxification (Zaret, 2002).

The *Drosophila* fat body is often referred to as the equivalent of mammalian adipose cells and liver. It works as an energy reservoir and nutrient sensor, in order to regulate proper fly development and lifespan (Y. Zhang & Xi, 2015). *Drosophila* fat body development starts when clusters of fat body precursors are specified in the different embryonic mesodermal segments (Figure 18A). Later during development, fat body precursors start to proliferate, and fat body primordia extend in size (Figure 18B). These primordia then coalesce to form a continuous sheet of fat body cells, along the lateral wall of the embryo at the end of embryogenesis (Figure 18C) (Abel et al., 1993; Rehorn et al., 1996; Riechmann et al., 1998; Sam et al., 1996).

2- Role of GATA factors in the mammalian liver and the Drosophila fat body

GATA4 and FOXA proteins bind the regulatory regions of liver-specific genes, such as *albumin*, where they act as pioneer factors, thus facilitating the accession of other transcription factors to regulatory regions, in order to promote liver-specific gene expression. Binding of

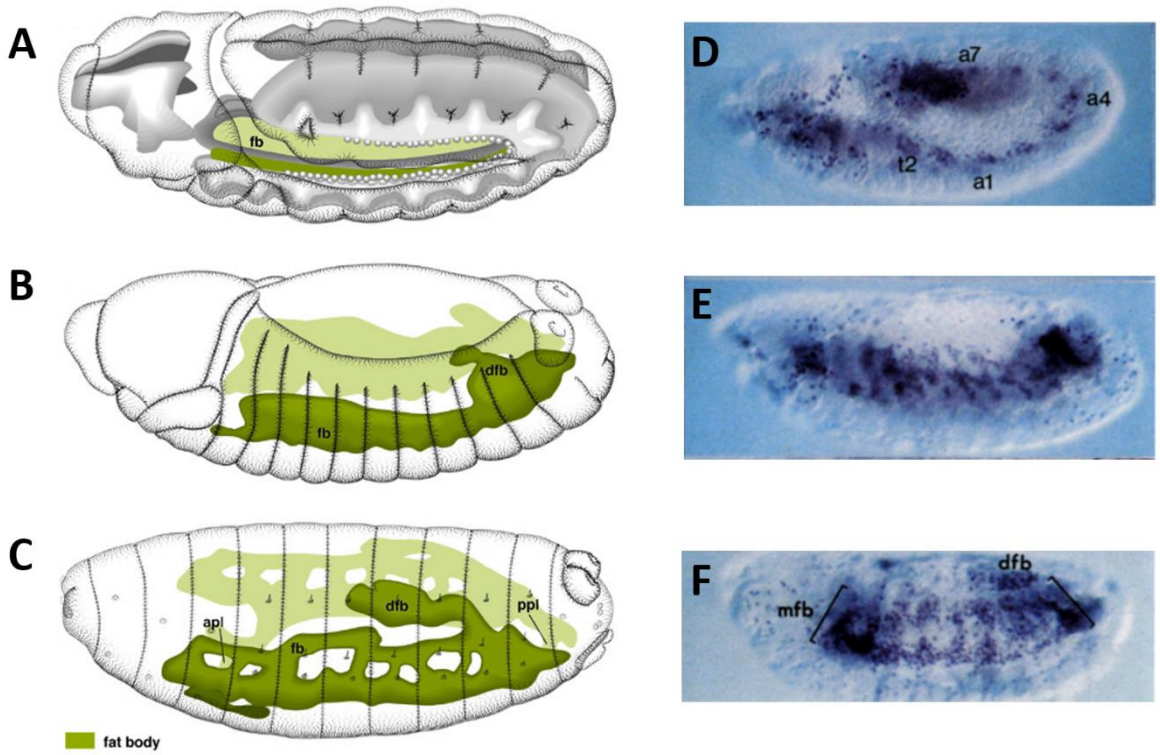


Figure 18. The *Drosophila* fat body formation.

(A-C) Schematic representation of embryonic fat body formation. The anterior pole of the embryo is to the left while the dorsal is up. (A) At stage 12 of embryogenesis, the fat body precursors are specified in the inner layer of mesoderm (shaded gray). (B) At stage 13, the fat body precursors form an elongated sheet of cells, and at the posterior part of this sheet a narrow strip of cells separates from the main mass of the fat body called the dorsal fat body (dfb). (C) At stage 17, large holes and clefts appear in the fat body sheet and supplementary regionally specialized regions such as the anteriorly (apl) and the posteriorly (ppl) horizontal plate are formed (adapted from (Hartenstein, 1993). (D-F) Immunostaining of whole-mount embryos against Serpent protein shows clusters of fat cell progenitors in segments ranging from the thoracic region t2 to abdominal region a7 at stage 12 of embryogenesis (D), the fat body precursor cells sheet at stage 13 of embryogenesis (E) and the mature fat body at the stage 16. At this latter stage, both of the main fat body cells (mfc) and the dorsal fat body cells (dfc) are formed (adapted from (Sam et al., 1996).

GATA4 to liver-specific regulatory regions occurs prior to commitment of the pluripotent endodermal cells to hepatic cell fate, suggesting that GATA4 has a role in liver specification. However, in GATA4 loss-of-function mice, the hepatic cells are still specified in the endoderm. The same results are obtained in mice with GATA6 loss-of-function, suggesting that GATA4 and GATA6 are redundant during liver specification (Watt et al., 2007; Zhao et al., 2005). Although mutant mice with loss-of-function of GATA4 or GATA6 do not alter liver specification, expansion of the primary hepatic rudiment is altered in both mutants, demonstrating the essential roles played by GATA4 and GATA6 factors during liver bud growth and thus liver development (Watt et al., 2007; Zhao et al., 2005). In *Drosophila,* Srp is expressed in fat body precursors since their specification and remains expressed in fat body cells throughout the entire fat body development. Although some fat body precursors are specified at early stages in *srp* loss of function mutant embryos, the number of fat body precursors produced is reduced as compared to wild-type flies. In addition, these fat body precursors are not able to proliferate and die. Premature apoptosis of fat body precursors provokes a subsequent loss of differentiated fat body cells. These results show that GATA4, GATA6 and Srp are required for the expansion and development of both the mammalian liver and the fly fat body.

(D) *Drosophila* as a model system to study GATA functions

GATA proteins are versatile transcription factors. They act as either activators or repressors of gene expression. They modulate chromatin conformation and the DNA packaging in order to facilitate activation of gene transcription (M. Tremblay et al., 2018). In addition, they interact with several other proteins in order to widely regulate their own functions (Jason A. Lowry &

Mackay, 2006). Importantly, during mammalian development, GATA factors are critical for the formation and physiological functions of numerous tissues and organs, such as blood cells, heart, lungs, intestine, liver, pancreas, kidneys, testis, ovaries, breast and nervous system (M. H. Lentjes et al., 2016). Despite the variable modes of action and roles identified for the six mammalian GATA factors, they all have in common the presence of two highly conserved zinc finger domains that are crucial for GATA protein functions (Martin & Orkin, 1990; Trainor et al., 1996; Trainor et al., 2000; Tsang et al., 1997; Yang & Evans, 1992). Accordingly, these two domains have been strictly conserved across evolution. In the fruit fly *Drosophila melanogaster*, five GATA factors are produced, three of them having the canonical two zinc finger domains (Lin et al., 1995; Ramain et al., 1993; Waltzer et al., 2002). The amino acid sequences of these domains are almost identical to those of the mammalian GATA zinc finger domains. Among these amino acids, we mention the sole valine residue of the N-terminal zinc finger domain that is responsible for the interaction of the mammalian GATA1 and GATA4 factors with FOG proteins. This residue is also present in the N-ZnF of *Drosophila* GATA Srp and plays, like in mammals, an important role during the interaction of Srp with the *Drosophila* FOG Ush (J. D. Crispino et al., 1999, 2001; Nancy Fossett et al., 2003).

In addition to this sequence conservation between mammalian and *Drosophila* GATA zinc fingers, it has been shown that GATA proteins display an incredible level of similarity, as they are implicated in the regulation of the same developmental processes, including blood cell differentiation (Rehorn et al., 1996; Takahashi et al., 1997), hematopoietic precursor proliferation and maintenance (Frelin et al., 2013; H. Gao et al., 2016), cardiomyocyte differentiation (Klinedinst & Bodmer, 2003; Zhao et al., 2008), gut formation and regeneration (Reuter, 1994; Walker et al., 2014), gut regionalization (Battle et al., 2008; Beuling et al., 2011; Okumura et al., 2005, 2007), oocyte maintenance (Kyrönlahti et al., 2011; Lepesant et al., 2020)

and finally mammalian liver / *Drosophila* fat body development (Rehorn et al., 1996; Watt et al., 2007; Zhao et al., 2005). They also share similarities at the mechanistic level (same molecular interactors, antagonism mechanisms, transcriptional regulation and GATA switch mode of action, as described above), and they even regulate the same target genes during similar developmental processes, *e.g.*, GATA4 and GATA6 proteins promote Delta expression during gut formation (Walker et al., 2014), while *Drosophila* GATAe is implicated in the expression of Delta in intestinal stem cells (Dutta et al., 2015; Okumura et al., 2016), and both Srp and mammalian GATA6 repress the expression of Crumbs transmembrane proteins in order to promote EMT (K. Campbell et al., 2011).

Throughout the years, *Drosophila* has emerged as a powerful genetic model organism. Numerous studies have been designed in *Drosophila* in order to analyze the relation between human genetic alterations and illnesses, including morphological defects, abnormal organ functioning and cancer (Bier, 2005; Mirzoyan et al., 2019). Furthermore, basic developmental regulatory pathways and processes are conserved between vertebrates and invertebrates, thus establishing *Drosophila* as an ideal model in which to elucidate the mode of action and the role of numerous factors implicated in various developmental processes (Gold & Brückner, 2015; Pitsouli & Perrimon, 2008; Saffman & Lasko, 1999; Zaffran, 2003). Among them, we have described above the mammalian and *Drosophila* hematopoietic development (Gold & Brückner, 2015), cardiac developmental processes (Brade et al., 2013; Vogler & Bodmer, 2015), intestine morphogenesis, ovary formation, and finally mammalian liver and *Drosophila* fat body development (Sam et al., 1996; Zaret, 2002; Y. Zhang & Xi, 2015).

Conservation of developmental processes between mammals and fly, as well as the structural and functional conservation of GATA factors across evolution, led us to consider the fly as an ideal organism model in which to study the roles of GATA zinc finger domains. Among the *Drosophila* GATA factors, Srp provides a unique opportunity to study the functions of the GATA zinc finger domains as, like I have already stated, since Srp proteins are produced as two different isoforms containing either the two zinc finger domains (SrpNC) or only the C-terminal zinc finger domain (SrpC) (Figure 19) (Waltzer et al., 2002). Interestingly, the alternative splicing process at the origin of these two isoforms, occurs in all cells expressing *srp* gene and at the same developmental time (Waltzer et al., 2002). The simultaneous presence of the two Srp isoforms, makes Srp protein very interesting to study in order to understand the role of GATA zinc finger domains *in vivo*.

In order to analyze the role of Srp zinc finger domains during the development of the fly, I took advantage of the CRISPR/Cas9 technique to generate two mutant fly lines (Gratz et al., 2015). The first fly line is characterized by the loss-of-function of SrpNC isoform but not of SrpC, thus allowing us to investigate the role of the N-ZnF domain function during fly development. The second fly line is unable to produce the SrpC isoform but still expresses the SrpNC isoform, a situation similar to all vertebrate contexts. The study of this mutant fly line should help us in understanding of the role played by this single zinc finger-containing SrpC isoform during development. Finally, a mutant fly line with a substitution mimicking the V205G mutation in the N-terminal zinc finger domain, has been engineered in order to identify the functions depending on the interaction of SrpNC with its cofactor of the Friend of GATA family, Ush. Srp proteins are produced in several tissues, where they play essential functions (hemocytes, gut, fat body and the ovaries) (Lepesant et al., 2020; Rehorn et al., 1996; Reuter, 1994; Sam et al., 1996). Although several studies have been carried out to determine the functions of Srp *in*

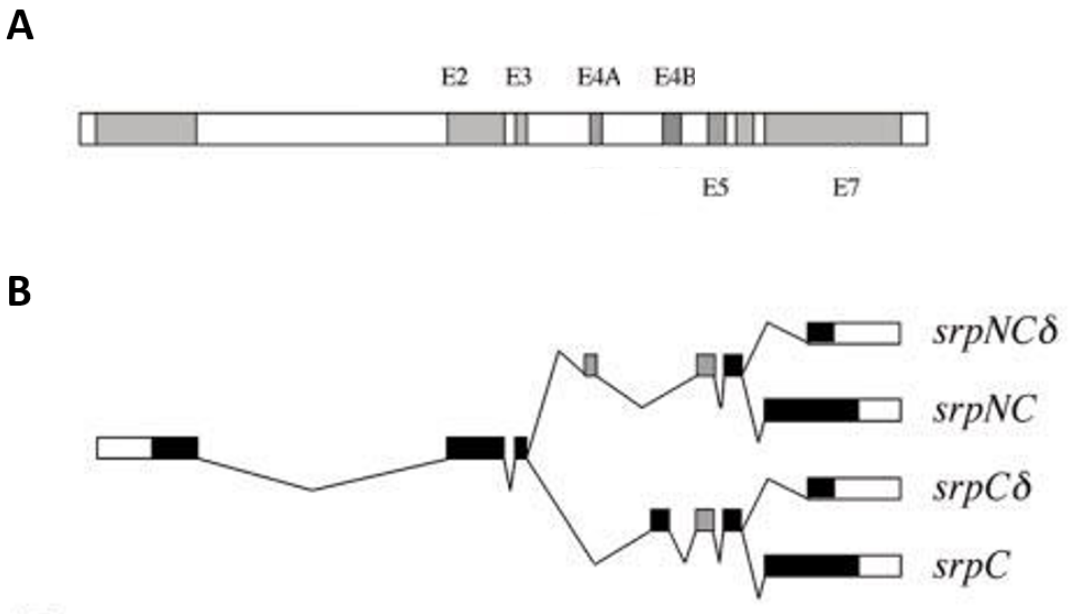


Figure 19. Different isoforms of srp are produced by alternative splicing process.

(A) Schematic representation of the *srp* locus shows the 8 exons (grey boxes) and the 9 introns (white boxes) of *srp* gene. (B) Schematic representation of the alternatively spliced transcripts. Transcripts containing exons 4A and 5 (grey boxes) that encode the N-terminal and the C-terminal zinc finger domains respectively are known as *srpNC* while those containing the exons 4B (black box that is alternatively spliced with the exon 4A) and 5 (grey box) are called *srpC*. The presence of an internal splice acceptor site in the exon 7 increases the numbers of *srp* isoforms to four: *srpNCδ, srpNC, srpCδ, srpC* (adapted from(Waltzer et al., 2002)).

vivo, no study has ever assessed the roles of the zinc finger domains during the establishment of these functions. Characterization of the different mutants I generated, reveals an alternative splicing-mediated sub-functionalization of the *srp* gene. The results of my work are described in the next chapter.

Chapter II. Results

(A) GATA factor Srp is produced as two isoforms, containing either one or two Zinc-Finger domains

The two major isoforms of Srp are produced by alternative splicing of *srp* precursor mRNA (Figure 20A). Both SrpNC and SrpC contain the C-terminal finger domain encoded by the fifth exon. The isoform srpNC results from the inclusion of the alternative exon 4A, which encodes the Srp N-terminal finger domain, and simultaneous exclusion of the exon 4B. In contrast, srpC is obtained by inclusion of exon 4B and simultaneous exclusion of exon 4A. The substitution of exon 4A that encodes the N-terminal finger domain by exon 4B leads to a product that is devoid of the N-terminal finger, but still contains the C-terminal one encoded by exon 5. In order to get more information about the region of SrpC encoded by the alternative exon 4B, we aligned the SrpC amino acid sequence with protein sequences of other single zinc finger GATA factors found in various arthropod species. Two motifs, located almost at both extremities of the sequence encoded by the exon 4B, have been conserved in all the aligned sequences (Figure 21), suggesting that the middle portion encoded by this exon does not support a conserved function. Furthermore, alignment of SrpC sequence with sequences of vertebrate GATA factors that normally contain the two canonical zinc finger domains, showed a conservation of only three amino acids, arginine-arginine-leucine (RRL), located at the most C-terminal extremity of the exon 4B encoded region (Figure 22). In mammals, this RRL motif belongs to the basic region that links the two GATA zinc finger domains, and it is likely required for normal GATA activity, as substitution of GATA4 and GATA5 RRL motif leucine residues to proline, was identified in patients with human heart disease (Bonachea et al., 2014; Reamon-Buettner, 2005). Finally, in order to determine if the amino acid sequence coded by the exon 4B contains any particular protein domains, functional analysis of this sequence by bioinformatics tools was performed, but it did not reveal any known functional domains (see material and methods).

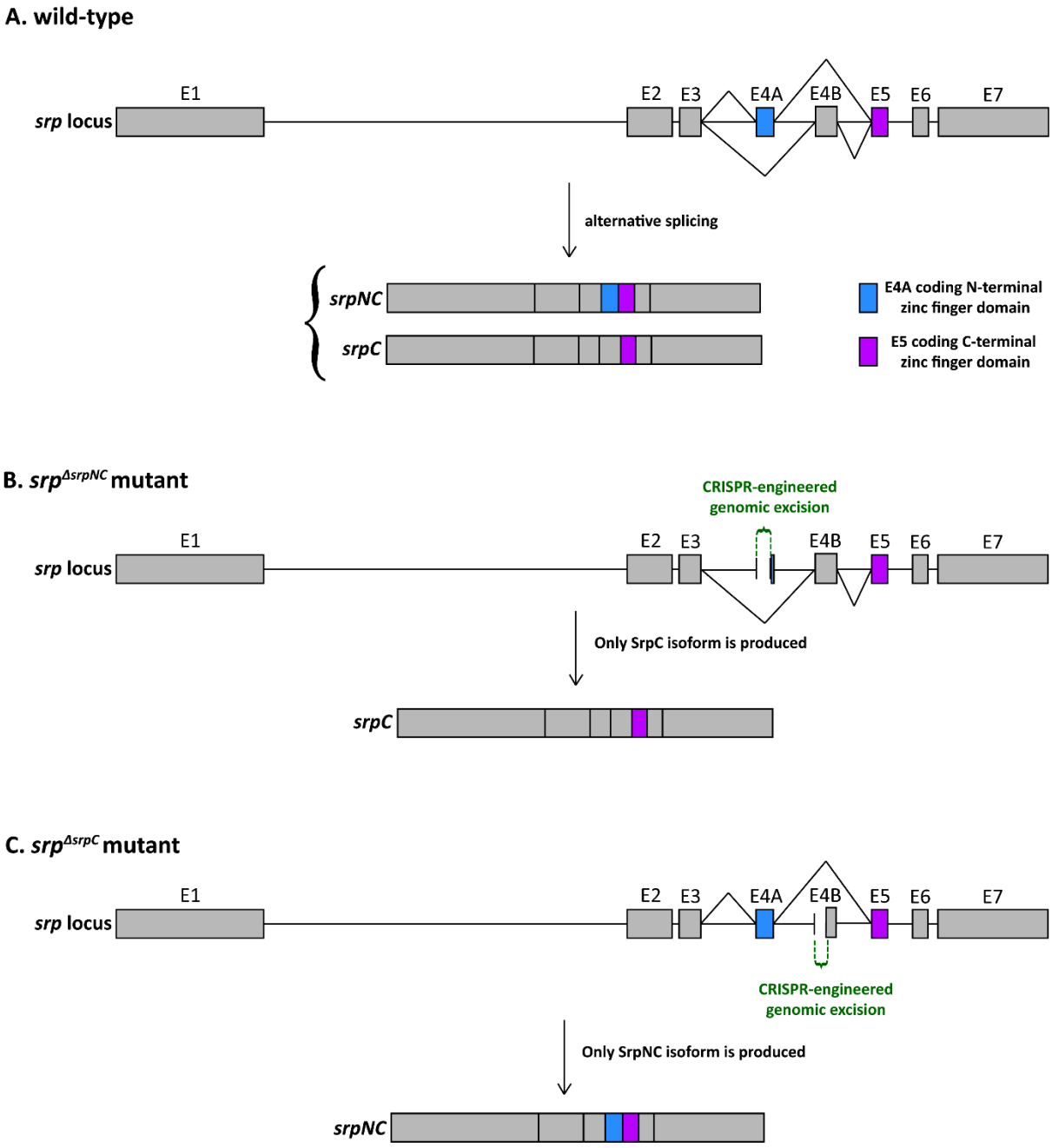


Figure 20. Generation of mutants with SrpNC or SrpC loss-of-function.

(A) Schematic description of *srp* pre-mRNA of wild-type flies showing the mutually exclusive alternative splicing process that occurs between exons E4A and E4B. Inclusion of exon E4A in *srp* transcript gives rise to the *srpNC* isoform, while inclusion of exon E4B in *srp* RNA gives rise to the *srpC* isoform. (B-C) Schematic representation of the *srp* pre-mRNA of mutant fly line $srp^{\Delta srpNC}$ (B) with exon 4A deleted and accompanied by SrpNC loss-of-function, and of the mutant fly line $srp^{\Delta srpC}$ (C) with exon 4B deleted and associated SrpC loss-of-function.

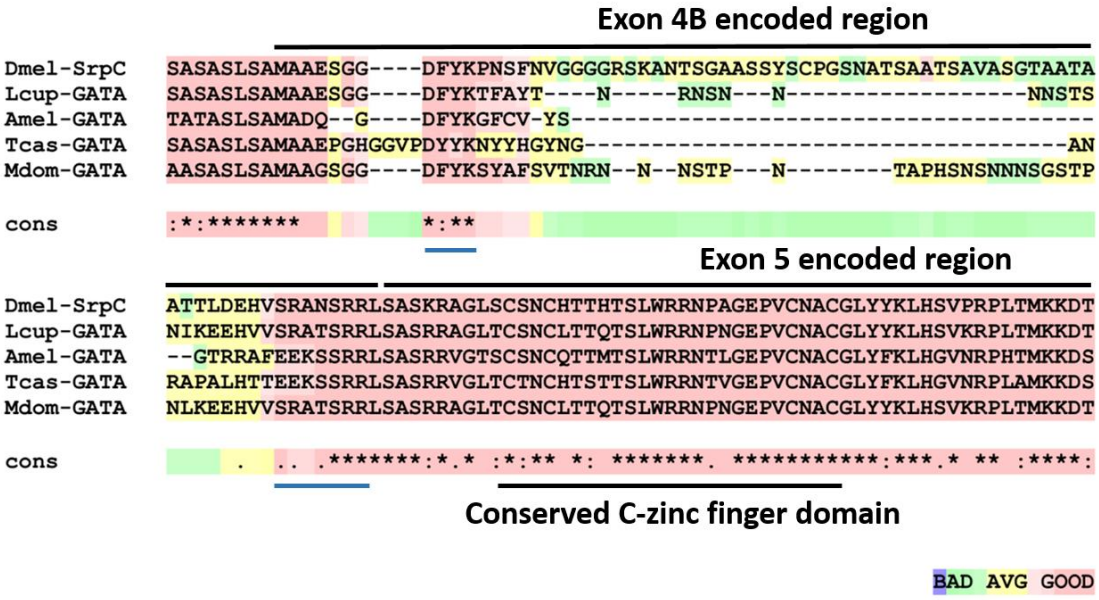


Figure 21. Alignment of SrpC protein sequence with GATA protein sequences of other arthropods.

Some regions (underlined in blue) in the protein sequence encoded by the alternatively spliced exon E4B are conserved with the other insect GATA factors. Dmel: *Drosophila melanogaster*, Lcup: *Lucilia cuprina*, Amel: *Apis mellifera*, Tcas: *Tribolium castaneum*, Mdom: *Musca domestica*. Cons: conservation color scheme.

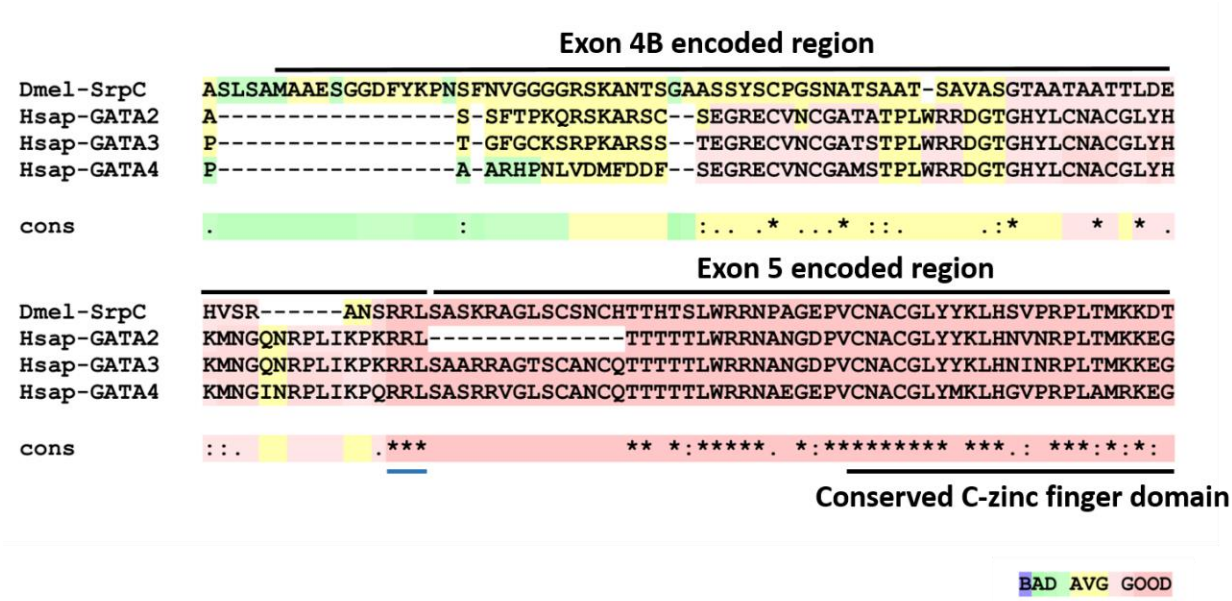


Figure 22. Alignment of SrpC protein sequence with mammalian GATA protein sequences.

RRL motif of the sequence encoded by the alternatively spliced exon E4B, juxtaposed to the C zinc finger domain of SrpC, is conserved with mammalian GATA factors. Dmel: *Drosophila melanogaster*, Hsap: *Homo sapiens*. Cons: conservation color scheme.

Altogether, these data suggest that substitution of exon 4A by exon 4B results in the production of a single zinc finger GATA Srp isoform.

(B) Both isoforms of GATA factor Srp are produced at postembryonic stages

As was already demonstrated at embryonic stage (Waltzer et al., 2002), we found that Srp proteins are produced as two alternatively spliced variants during the *Drosophila* post-embryonic stages. Total RNA of *white-eye* flies at third instar larval stage was extracted and used as matrix for reverse-transcription polymerase chain reaction (RT-PCR) experiments, in order to identify the nature of *srp* isoforms present in the extracts. To do so, forward and reverse primers that recognize exons of *srp* gene located at either side of the *srp* mutually exclusive exons 4A and 4B were used. Migration of RT-PCR products on agarose gel led to the detection of two amplicons corresponding to both *srp* mRNA isoforms: *srpNC* and *srpC* (Figure 23A). Similar results were obtained after analysis of transcripts extracted specifically from fat body larvae (Figure 23B, lane 1), an organ already known to express high *srp* levels (Senger et al., 2006). Furthermore, analyses of mRNAs produced in samples containing adult fly digestive system and Malpighian tubules also showed the presence of the two alternatively spliced *srp* isoforms (Figure 23B, lanes 2, 3). All these results indicate that at different *Drosophila* developmental stages and in different tissues expressing *srp*, similar alternative splicing mechanisms occur, giving rise to two variants of Srp proteins. This prompted us to analyze the role of each of these two isoforms in the fly, by generating mutant fly lines carrying SrpNC or SrpC loss-of-function.

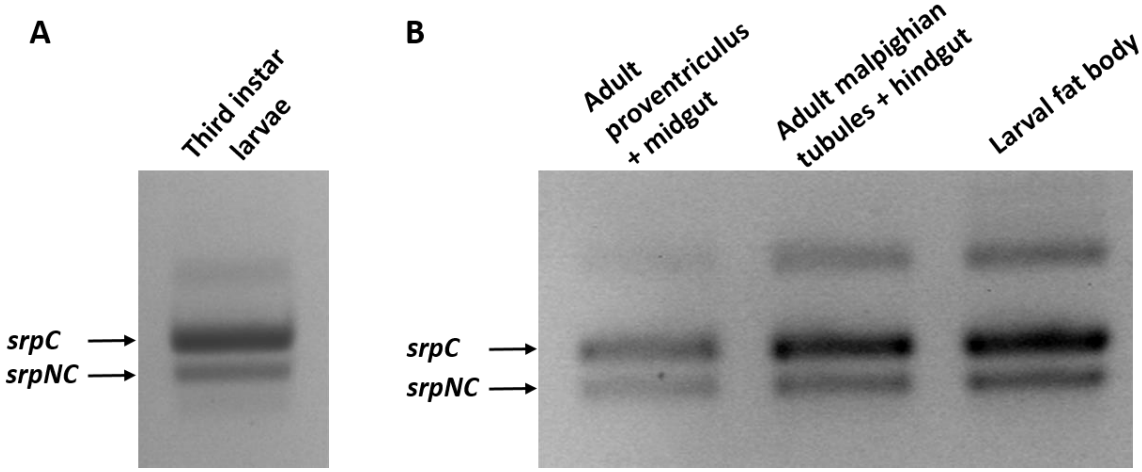


Figure 23. Two isoforms of Serpent are produced in *Drosophila*.

(A) Agarose gel showing the RT-PCR results from total RNA of third instar larvae, indicating the presence of two cDNA bands, corresponding to both isoforms SrpC and SrpNC. Note that the size of the exon included in SrpC is larger than the alternatively spliced exon found in SrpNC, which explains the slower migration of the SrpC corresponding band in comparison to the SrpNC band. (B) Agarose gel showing the RT-PCR results from the RNA extracted from adult proventriculus and midgut (left lane of gel), adult malpighian tubules and hindgut (middle lane) and larval fat body (right lane of gel). Two bands corresponding to SrpNC and SrpC have been identified in these tissues.

(C) Generation of *srp* mutant alleles specifically deprived of either SrpNC or SrpC isoform

In order to determine the role of Srp isoforms in *Drosophila*, we generated mutant flies unable to produce either SrpNC or SrpC (Figure 20B, C). Using the CRISPR/Cas9 system, we generated two fly lines, one containing a deletion of most of exon 4A (thus preventing the production of SrpNC), and the other harboring a deletion in exon 4B, which removes the region containing the exon splice acceptor site (thus devoid of SrpC protein) (Figure 24A, B). We named these generated mutant flies $srp^{\Delta srpNC}$ (loss of SrpNC function) and $srp^{\Delta srpC}$ (loss of SrpC function), respectively.

In order to validate *srpNC* or *srpC* obliteration in these lines, total RNA was extracted from embryos homozygous for exon 4A or exon 4B deletion, respectively, and RT-PCR was carried out with forward and reverse primers located respectively in exons 3 and 5 (see materials and methods for more detail). Analysis on agarose gel of RT-PCR products recovered from total RNA extracted from homozygous $srp^{\Delta srpNC}$ embryos confirmed the specific loss of *srpNC* expression (Figure 25A, lane 2), in contrast to control embryos (Figure 25A, lane 1), and a single cDNA band corresponding to the *srpC* isoform was detected. Conversely, in $srp^{\Delta srpC}$ mutant embryos, only the cDNA band corresponding to s*rpNC* was detected, while no cDNA band corresponding to the isoform *srpC* was found (Figure 25B, lane 2). Of note, in $srp^{\Delta srpC}$ mutant embryos, the cDNA band corresponding to SrpNC was detected at a higher level than in control embryos (Figure 25B, lane 1), suggesting that *srpNC* is expressed at higher levels in $srp^{\Delta srpC}$ mutants. Therefore, in subsequent experiments, we systematically analyzed both $srp^{\Delta srpC}$ homozygous and hemizygous conditions ($srp^{\Delta srpC}$ mutation placed over the *Df(3R)BSC728*

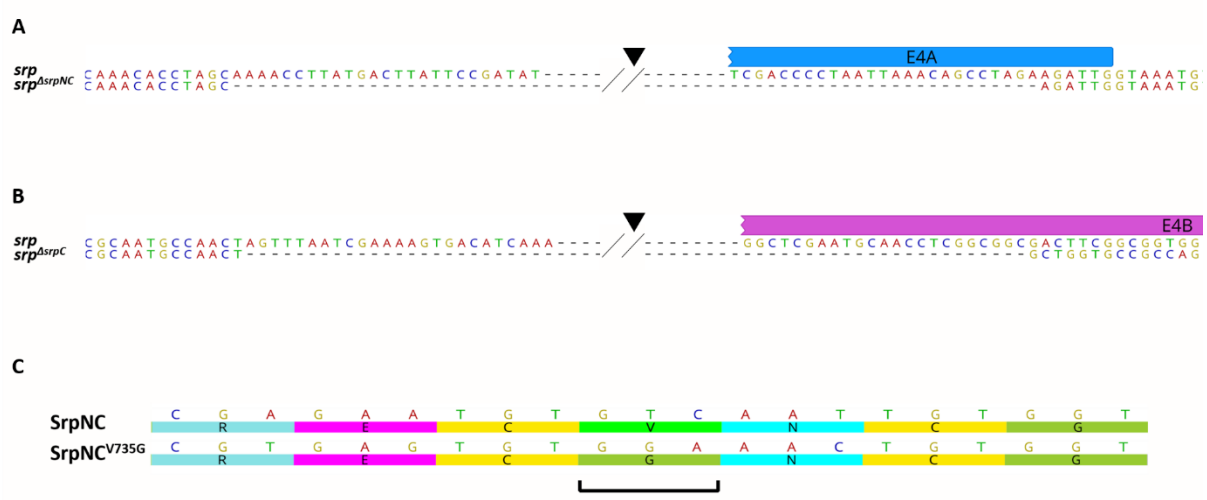


Figure 24. Validation by sequencing of the expected mutations in newly created mutant fly lines.

(A) Flies of the genotype *srp*$^{\Delta srpNC}$ have a deletion of almost the entire exon E4A, including the E4A acceptor splicing site (arrowhead). (B) Flies of the genotype *srp*$^{\Delta srpC}$ have a deletion in exon E4B including the E4B acceptor splicing site (arrowhead). (C) Flies of the genotype *srp*^{V735G}, have a substitution of the nucleotides GTC to GGA, leading to production of protein with substitution of the valine amino acid to glycine (bracket).

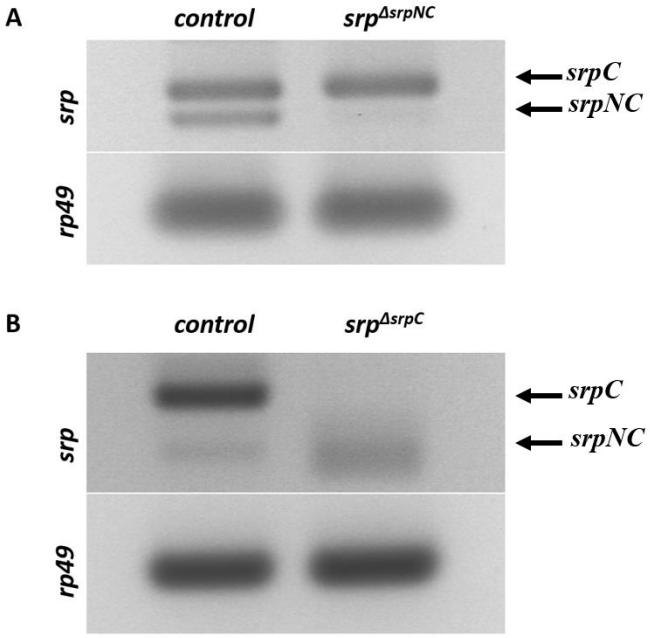


Figure 25. Validation of the expected loss-of-function of srp isoforms in newly created mutant fly lines by RT-PCR.

(B) The generated *srp*$^{\Delta srpNC}$ fly line is characterized by the absence of *srpNC* isoform production. (C) The fly line *srp*$^{\Delta srpC}$ is characterized by the absence of *srpC* isoform production.

deficiency that uncovers the *srp* locus), in order to detect any phenotypes that might be associated to *srpNC* overexpression in these mutants. In addition, we also tested allelic combinations with the amorphic allele srp^{3} and srp^{6G}, or the hemocyte specific allele srp^{AS}. The srp^{6G} allele corresponds to a mutation inducing a premature stop codon in the protein region encoded by *srp* second exon, thus leading to the loss of almost the entire Srp protein, including the two zinc finger domains. The srp^{3} allele carries a missense mutation in *srp* exon encoding the C-terminal zinc finger domain, which most likely inhibits the interaction Srp/DNA. The srp^{AS} allele contains an insertion of a transposable element in a regulatory region that controls *srp* expression specifically in embryonic hematopoietic cells, thus altering *srp* expression in these cells and leading to a complete loss of blood cell development (Rehorn et al., 1996).

(D) Generation of a *srp* mutant allele that specifically abolishes interaction of Srp with its FOG cofactor Ush

Important regulators of GATA factor functions are the FOG cofactors (Chlon & Crispino, 2012). The valine residue in GATA N-terminal zinc finger is required for interaction with FOG, and substitution of this valine to glycine alters the association GATA/FOG (J. D. Crispino et al., 1999). Interestingly, the Srp N-terminal zinc finger amino acid sequence is highly conserved with those of the mammalian GATA factors (Figure 26), and the knock-in replacement of valine to glycine alters the interaction of SrpNC with Ush (Nancy Fossett et al., 2003). In order to study the role of SrpNC/Ush interaction in *Drosophila*, we generated a mutant fly line harboring a substitution of this valine to glycine (V735G) (Figure 24C). The position of this valine in the N-finger was referred to the *srp* transcript variant B, annotated by the National Center for Biotechnology and Information (NCBI) Reference Sequence (Refseq) database (protein-id: NP_732098.1). Flies of this line are named srp^{V735G}.

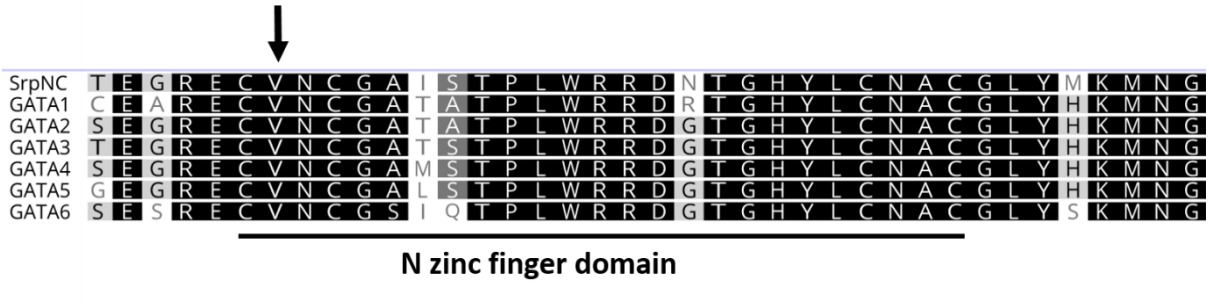


Figure 26. Alignment of SrpNC protein sequence with mammalian GATA protein sequences of other arthropods.

The N zinc finger domain of SrpNC is highly conserved with those of the mammalian GATA transcription factors, including the valine amino acid (arrow) mutated in SrpNC of *srp*^{V735G} flies (figure S1C). The NCBI accession number of each GATA protein sequence is present in the material and methods section, part Bioinformatic tools.

(E) SrpC, but not SrpNC, is required for *Drosophila* viability

Loss of *srp* function affects the ability of the fly to reach adulthood (Rehorn et al., 1996). All embryos homozygous for the amorphic allele srp^{6G} die before hatching (Figure 27). In order to determine whether this lethality is due to the absence of SrpC, of SrpNC, or to the simultaneous loss of both isoforms, the viability of $srp^{\Delta srpC}$ and $srp^{\Delta srpNC}$ embryos was assessed.

About 60% of $srp^{\Delta srpC}$ homozygous embryos were able to develop until the first instar larval stage (Figure 27), then most of them died before reaching third instar larval stage (only 13.88% of these larvae develop into third instar larvae). The surviving larvae were able to form pupae, but only a few escapers emerged as adults and then died immediately. As a consequence, no $srp^{\Delta srpC}$ adult flies were observed, indicating that the SrpC isoform is required for fly viability.

In order to confirm that the lethality observed in $srp^{\Delta srpC}$ mutant flies is due to the loss of SrpC function, and not to the genetic background of $srp^{\Delta srpC}$ flies, or to the overexpression of SrpNC isoform in these mutants, we analyzed the lethality of flies with the $srp^{\Delta srpC}$ mutation placed over the Df(3R)BSC728 deficiency, which removes several genes including *srp*. Interestingly, hemizygous embryos of ($srp^{\Delta srpC}$/Df(3R)BSC728) also failed to reach larval stage (Figure 28B), indicating that the lethality of $srp^{\Delta srpC}$ mutants is not a consequence of their genetic background. Similarly, $srp^{\Delta srpC}/srp^{6G}$ trans-heterozygous individuals died before reaching third instar larval stage (Figure 28A,B). All these results show that developmental defects found in mutant $srp^{\Delta srpC}$ are specifically due to the loss-of-function of Srp, and that SrpC is essential for fly development. In addition, the fact that the lethality phenotype is more severe in embryos hemizygous for $srp^{\Delta srpC}$ than in homozygous embryos indicates that $srp^{\Delta srpC}$ is an hypomorphic

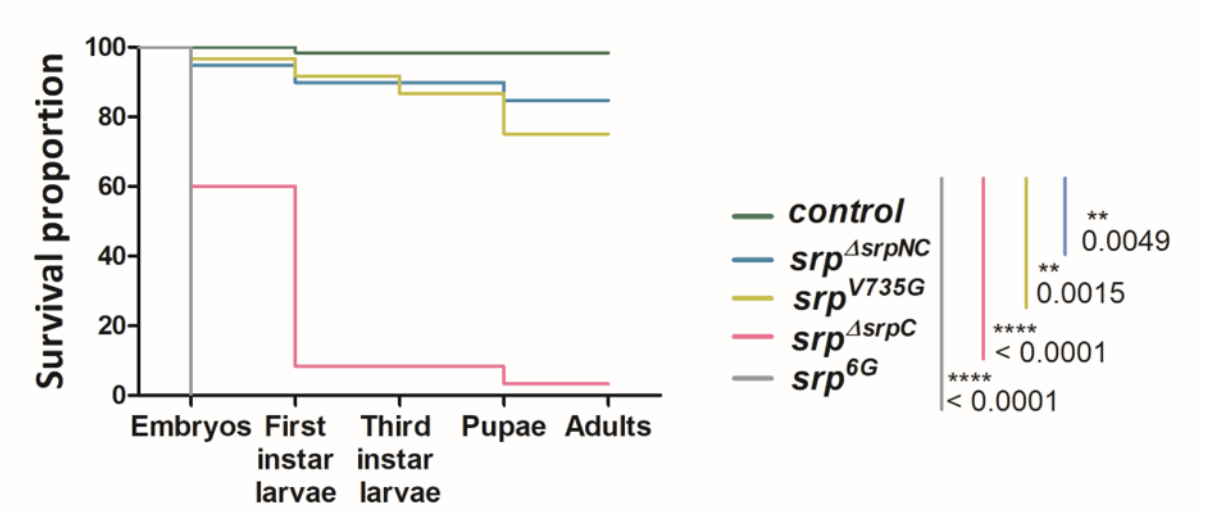
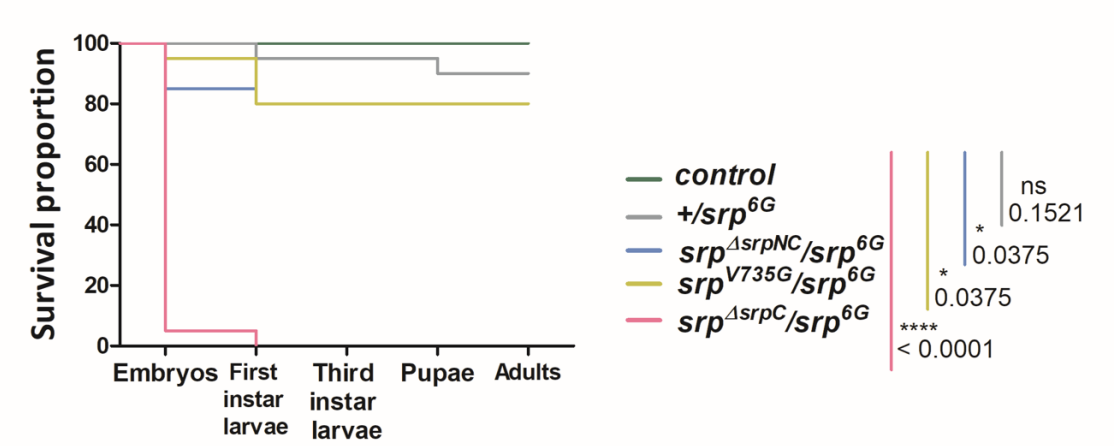


Figure 27. Survival analysis of different generated srp mutants.

Survival analysis shows that the loss-of-function of both Srp isoforms (srp^{6G}) causes *Drosophila* embryonic letality (grey curve), the loss-of-function of SrpC isoform ($srp^{\Delta srpC}$) alters the fly's ability to develop until adult stage (pink curve), the loss-of-function of SrpNC isoform ($srp^{\Delta srpNC}$) or the loss of the interaction of SrpNC/Ush (srp^{V735G}) reduces only slightly the fly's ability to develop until adult stage (blue and yellow curves). The survival analysis of *srp* mutants was repeated three times with similar results. Data were analyzed using the Gehan-Breslow-Wilcoxon test.

A



B

Cross	Percentage of no fluorescent expressing third instar larvae
+/TM3$^{twist>GFP}$ x +/TM3$^{twist>GFP}$	42 (n= 60)
srp$^{\Delta srpC}$/TM3$^{twist>GFP}$ x *Df(3R)BSC728*/TM3$^{twist>GFP}$	0 (n= 143)
srp$^{\Delta srpC}$/TM3$^{twist>GFP}$ x *srp*^{6G}/TM3$^{twist>GFP}$	0 (n= 89)
srp$^{\Delta srpC}$/TM3$^{twist>GFP}$ x *srp*^{3}/TM3$^{twist>GFP}$	0 (n= 46)

Figure 28. Validation of the survival analysis data.

(A) Survival analysis shows that loss of only one copy of Srp isoforms (*srp*^{6G}) has no effect on the ability of the fly to develop until adult stage. Flies trans-heterozygous for the loss-of-function of SrpNC and of both Srp isoforms (*srp*$^{\Delta srpNC}$/ *srp*^{6G}), or of the interaction of SrpNC with Ush and of both Srp isoforms (*srp*^{V735G}/*srp*^{6G}), have only a slight reduction in the ability to develop until adult stage. Flies trans-heterozygous for the loss-of-function of SrpC and of both Srp isoforms (*srp*$^{\Delta srpC}$/ *srp*^{6G}) are not able to develop until larval and adult stages. Data were analyzed using the Gehan-Breslow-Wilcoxon test. (B) Flies hemizygous for *srp*$^{\Delta srpC}$ mutation (*srp*$^{\Delta srpC}$/*Df(3R)BSC728*) and flies trans-heterozygous for the mutations *srp*$^{\Delta srpC}$ and *srp*^{6G} or *srp*^{3} are not able to develop until third instar larval stage. Like *srp*^{6G}, *srp*^{3} is a *srp* amorphic allele.

allele and that the presence of SrpNC in $srp^{\Delta srpC}$ embryos slightly compensate SrpC loss-of-function. Thus, SrpNC and SrpC have partially redundant functions during the fly development.

Furthermore, GATA zinc finger domains interact with DNA but also with other proteins. We analyzed thus the viability of embryos having $srp^{\Delta srpC}$ allele placed over the DNA binding mutant allele srp^{3} in order to understand how SrpC acts. Interestingly, embryos of the genotype $srp^{\Delta srpC}/srp^{3}$ die at embryonic stage (Figure 28B). In embryos $srp^{\Delta srpC}/srp^{3}$, srp^{3} allele produces mutant SrpC and SrpNC proteins that are unable to bind DNA while $srp^{\Delta srpC}$ allele will generate normal SrpNC protein that will bind and regulate the expression of its proper target genes and of target genes that are common to both SrpC and SrpNC. As $srp^{\Delta srpC}/srp^{3}$ embryos are not able to pass embryogenesis, we conclude that there are some SrpC specific target genes that are not recognized by SrpNC and that are required for *Drosophila* viability.

Contrarily to $srp^{\Delta srpC}$ mutants, flies with a specific loss of SrpNC function were able to develop until the adult stage (Figure 27). About 95% of $srp^{\Delta srpNC}$ homozygous embryos developed into first instar larvae. The majority of those developed and reached the third instar larval and pupal stages, and about 94.34% of $srp^{\Delta srpNC}$ pupae gave rise to adult flies. Likewise, almost 80% of $srp^{\Delta srpNC}/srp^{6G}$ flies reached adulthood (Figure 28A), showing that SrpNC is mostly dispensable for *Drosophila* viability (although the observed survival rate was slightly, yet significantly, lower than that in control conditions).

Similar to the results obtained with SrpNC loss-of-function, srp^{V735G} homozygous embryos were able to develop until adult stage, as about 96.67% of these embryos were able to hatch, and 77.58% of hatched larvae developed into adult flies (Figure 27). Accordingly, about 80% of embryos of the genotype srp^{V735G}/srp^{6G} were able to develop until adult stage (Figure 28A).

These results show that the complex SrpNC/Ush is not required for fly viability. Analysis of both $srp^{\Delta srpNC}$ and srp^{V735G} embryos shows that the N-terminal finger domain is dispensable for the essential steps of *Drosophila* development and viability. Finding that the two isoforms SrpC and SrpNC have different effects on the fly development, suggests that these two isoforms have different functions in the fly, and so we wanted to establish which *srp*-dependent functions are attributed to each isoform during *Drosophila* development.

(F) Either SrpNC or SrpC is sufficient for normal gut development

In *Drosophila*, one of the functions mediated by Srp occurs during early gut development (K. Campbell et al., 2011; Reuter, 1994). In the developing gut, srp^{6G} loss-of-function mutants showed no expression of *GATAe* and *grain* (Figure 29E, J, O, T), two markers normally expressed in midgut primordia (Figure 29A, K) and in some regions of terminally developed intestine (Figure 29F, P). Both Srp isoforms were produced in embryonic midgut progenitors and mature cells. However, neither the loss of SrpNC (Figure 29B, G, L, O), the loss of SrpC (Figure 29D, I, N, S), nor the disruption of the SrpNC/Ush interaction (Figure 29C, H, M, R) were associated to gut developmental defects, since gut markers were correctly expressed, and the gut intestine was shaped normally in all these mutant contexts. Taken together, these results show that the gut development is not altered by the loss of either SrpC or SrpNC, and that the presence of only one Srp isoform is sufficient for normal embryonic gut formation.

(G) SrpC, but not SrpNC, is required for plasmatocyte differentiation and crystal cell development

In addition to its function in gut development, Srp has essential roles during embryonic hematopoiesis in *Drosophila* (Rehorn et al., 1996). Consistently, in srp^{6G} complete loss-of-

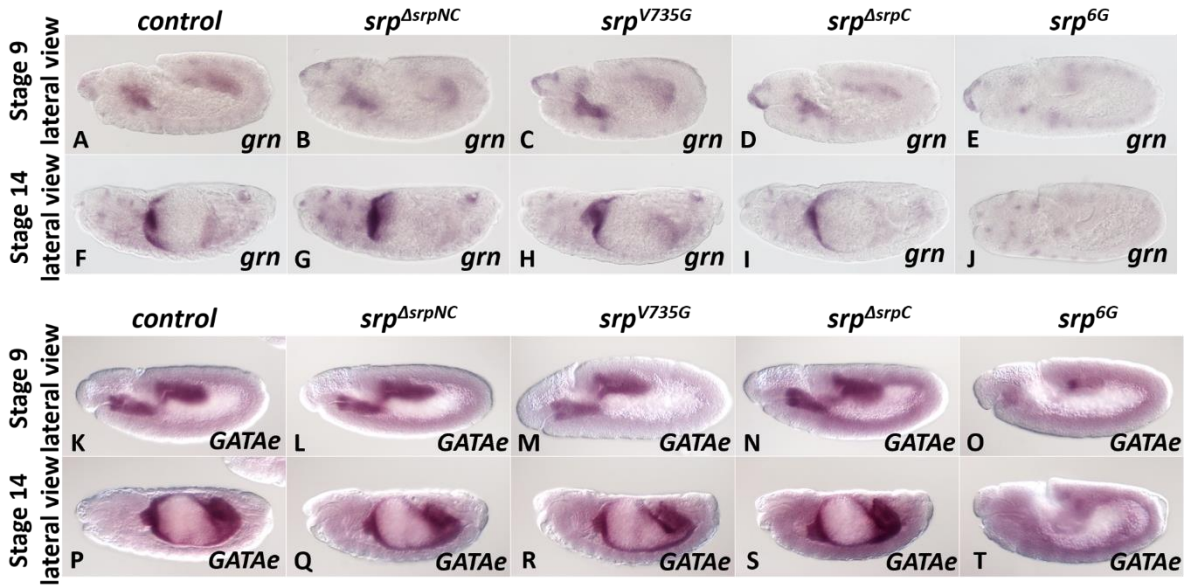


Figure 29. Both SrpNC and SrpC are sufficient for the *Drosophila* embryonic intestine development.

The loss-of-function of Srp (srp^{6G}) alters embryonic intestine development, as the embryonic intestine markers, *grn* and *GATAe*, normally expressed in the anterior and posterior midgut primordia (A and K) or in the completely developed intestine (F and P), are not expressed in srp^{6G} mutants (E, J, O, T). The loss-of-function of only SrpNC (B, G, L, Q) or only SrpC (D, I, N, S) has no effect on intestine development or *grn* and *GATAe* expression. Also, loss of the interaction SrpNC/Ush (C, H, M, R) has no effect on this process. *grn*: *grain*. Each experiment was repeated three times with similar results.

function mutants, no plasmatocytes expressing the *croquemort* (*crq*) or *viking* (*vkg*) markers are detected (Figure 30E, J), in contrast to wild-type embryos (Figure 30A, F). Interestingly, plasmatocytes expressing *crq* and *vkg* are visible in both $srp^{\Delta srpC}$ (Figure 30D, I) and $srp^{\Delta srpNC}$ (Figure 30B, G) specific loss-of-function mutants, as well as in srp^{V735G} mutants (Figure 30C, H). These results show that plasmatocytes are formed even when only one isoform of Serpent is present. However, we found that the staining for the plasmatocyte marker *Peroxidasin* (*Pxn*) is almost abolished in $srp^{\Delta srpC}$ embryos (Figure 31D, I) contrary to wild-type embryos (Figure 31A, F), $srp^{\Delta srpNC}$ embryos (Figure 31B, G) and srp^{V735G} embryos (Figure 31C, H), indicating that the loss-of-function of SrpC, but not that of SrpNC or disrupted SrpNC/Ush interaction, impairs *Pxn* expression in plasmatocytes and that SrpNC is unable to compensate, even slightly, SrpC loss-of-function during *pxn* expression.

In order to confirm the role of SrpC in *pxn* expression, we analyzed $srp^{\Delta srpC}/srp^{AS}$ trans-heterozygous embryos. Interestingly, they display a reduction in *Pxn* expression (Figure 32B) in comparison to plasmatocytes of wild-type embryos (Figure 32A). This phenotype is similar to that obtained in $srp^{\Delta srpC}$ homozygous embryos, indicating that the loss of *Pxn* expression in the mutant $srp^{\Delta srpC}$ is due to the loss-of-function of SrpC in the hemocytes. Taken together, these results demonstrate that the SrpC isoform is required for *Peroxidasin* expression in the embryonic plasmatocytes, while the SrpNC isoform is dispensable in this process.

In addition to plasmatocytes that constitute 95% of the total hemocytes, the crystal cells, which are the equivalent of mammalian granulocytes, form the remaining 5% of the hemocytes. No expression of the crystal cell fate determinant coding gene *lozenge* (*lz*) or of the crystal cell-specific differentiation marker *Prophenoloxidase 2* (*PPO2*) is detected in srp^{6G} null mutants (Figure 31O, T) as compared to wild type embryos (Figure 31K, P). Interestingly, we found that

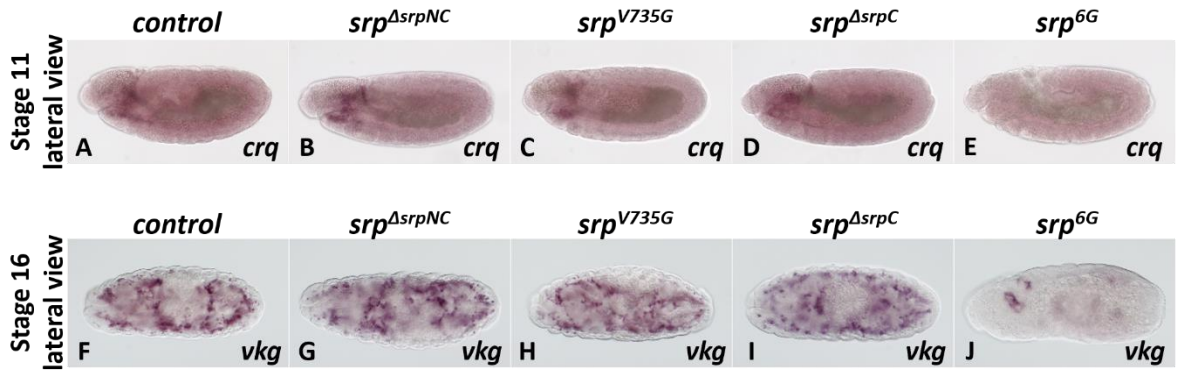


Figure 30. Both SrpNC and SrpC are sufficient for *Drosophila* embryonic plasmatocyte development.

Analysis of embryonic plasmatocyte marker expression, crq (A-E) and vkg (F-J), shows that Srp loss-of-function alters embryonic plasmatocyte formation, compare (E, J) to (A, F), where expression of *crq* and *vkg* in plasmatocytes normally migrating in head mesoderm during stage 11 of embryogenesis (A) and circulating throughout the embryo during stage 16 of embryogenesis (B) is absent in srp^{6G} mutants (E, J). The loss-of-function of only SrpNC (B, G) or only SrpC (D, I) has no effect on plasmatocyte formation. Loss of SrpNC/Ush interaction (C, H) has no effect on this process. *crq*: *croquemort*, *vkg: viking*. Each experiment was repeated three times with similar results.

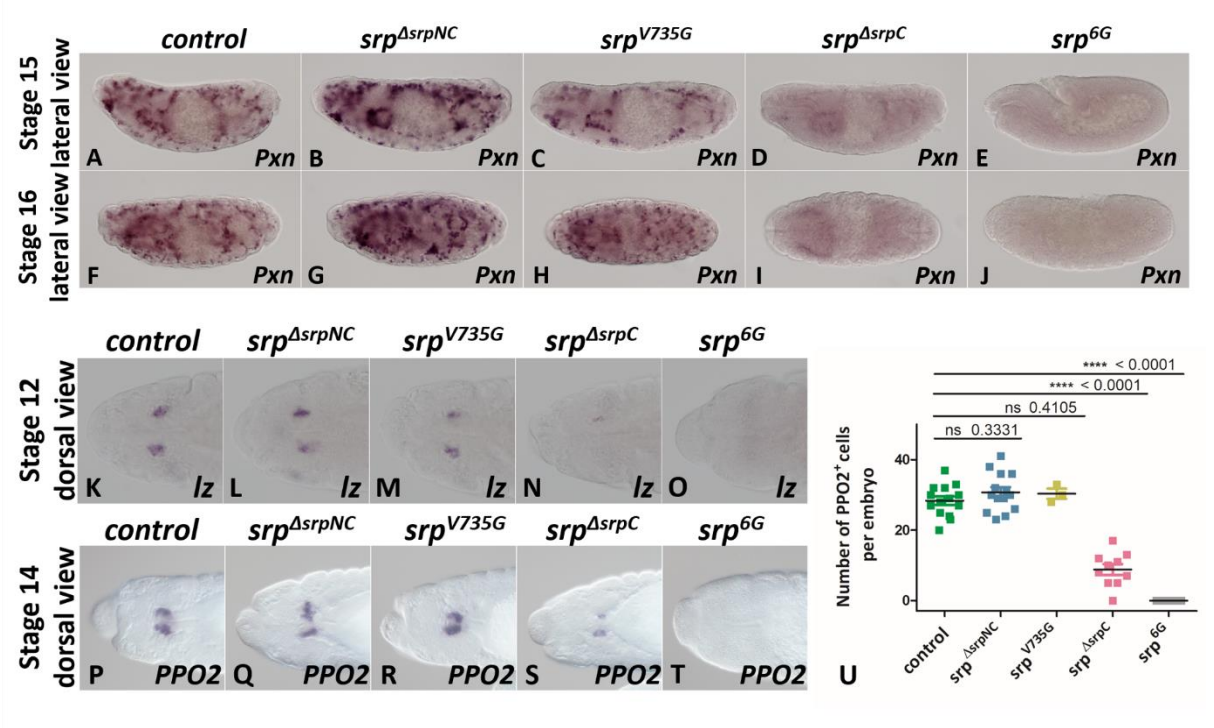


Figure 31. SrpC, but not SrpNC, is requiered for the expression of *pxn* and *lz* in *Drosophila* embryonic hemocytes.

Loss-of-function of both Srp isoforms (E, J) or only SrpC (D, I) alters the presence of *pxn* expressing plasmatocytes in *Drosophila* embryos, normally circulating throughout embryos during stages 15 (A) and 16 (F). Loss-of-function of SrpNC (B, G) or of the SrpNC/Ush interaction (C, H) has no effect on this process. Loss-of-function of both Srp isoforms (O, T) totally abolishes crystal cell formation, normally visualized by expression of both *lz* (K) and *PPO2* (P) genes. Loss-of-function of SrpC (N, S) strongly reduces the number of formed crystal cells in the Drosophila embryo. Loss-of-function of SrpNC (L, Q) or of the SrpNC/Ush interaction (M, R) has no significant effect on this process. The number of formed crystal cells in the different genotypes has been quantified (U) and data were analyzed using the Mann-Whitney test. pxn: peroxidasin, lz: lozenge, *PPO2*: *Prophenoloxidase 2*. Each experiment was repeated three times with similar results.

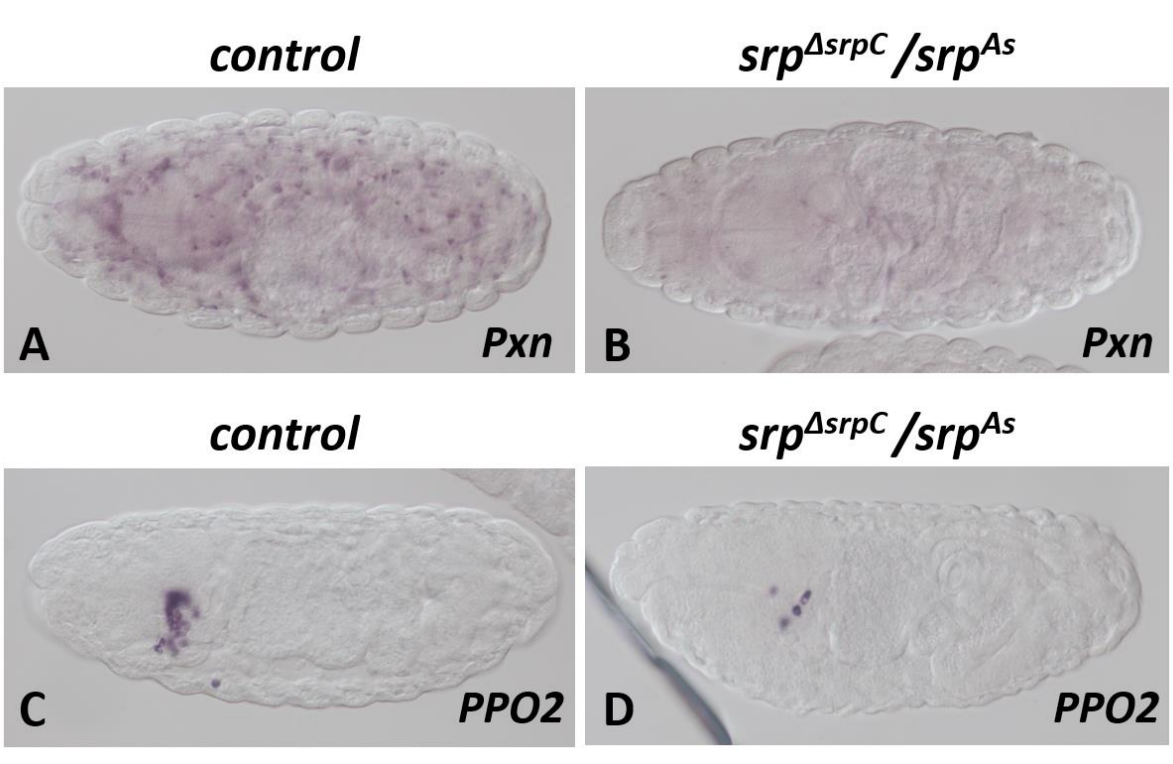


Figure 32. Validation of the role of SrpC in the expression of *pxn* and *lz* in *Drosophila* embryonic hemocytes.

Embryos with loss-of-function of SrpC in embryonic hemocytes ($srp^{\Delta srpC}/srp^{As}$) (B, D) have a strong reduction of *pxn* expression in plasmatocytes (B) and the number of *PPO2* expressing crystal cells (D) in comparison to the control (A, C). srp^{As} is a *srp* allele that abolishes *srp* expression in embryonic hemocytes (Rehorn et al., 1996). *pxn*: *peroxidasin, PPO2*: *Prophenoloxidase 2*. Each experiment was repeated at least three times with similar results.

mutants with loss-of-function of SrpC display a significant reduction in the number of progenitor crystal cells expressing *lz* (Figure 31N, S) and consequently of mature crystal cells (Figure 31S) in comparison to wild-type embryos (Figure 31K, P), $srp^{\Delta srpNC}$ embryos (Figure 31L, Q) and srp^{V735G} embryos (Figure 31M, R). Quantification of crystal cells in each mutant is represented (Figure 31U). The fact that the absence of both *srp* isoforms in srp^{6G} embryos (Figure 31O, T, U) has more severe effect on crystal cell development than the absence of only SrpC isoform (Figure 31N, S, U) indicates that SrpNC can partially compensate SrpC loss-of-function during crystal cell formation.

To corroborate that alteration of crystal cell formation in $srp^{\Delta srpC}$ mutants is due to the loss-of-function of SrpC and not to the mutant genetic background, we analyzed crystal cell formation in $srp^{\Delta srpC}/srp^{AS}$ trans-heterozygous embryos. As expected, these embryos display a strong reduction in crystal cells numbers (Figure 32D) in comparison to wild-type embryos (Figure 32C). This phenotype is very similar to what is seen in $srp^{\Delta srpC}$ mutants, demonstrating that alteration of crystal cells formation in these mutants is specifically associated to the loss of SrpC function.

Altogether, these results show that SrpC, but neither SrpNC nor the SrpNC/Ush complex, is required for crystal cell formation.

Interestingly, although these $srp^{\Delta srpC}/srp^{As}$ embryos exhibit a notable reduction of *pxn* expression in plasmatocytes and a strong reduction in the number of crystal cells formed, we found that these embryos are able to develop until the adult stage, contrary to $srp^{\Delta srpC}$ mutants that die before reaching adult stage (Figure 33). These results indicate that the loss-of-function of SrpC in embryonic hemocytes is not sufficient to cause fly lethality and that SrpC might have essential functions in other cell types.

(H) SrpC, but not SrpNC, is required for fat body development

As already published, mutants with *srp* loss-of-function are unable to develop mature fat body cells (Rehorn et al., 1996; Sam et al., 1996) that express the markers *Alcohol dehydrogenase* (*Adh*) (Figure 34E, J), *Glutactin* (*Glt*) (Figure 34O, T) and *Tiggrin* (*Tig*) (Figure 34Y, D') contrary to wild-type embryos whose fat body cells are expressing those markers (Figure 34A, F, K, P, U, Z). Interestingly, the $srp^{\Delta srpC}$ mutation alters expression of all three fat body markers (Figure 34D, I, N, S, X, C'), which might result from defective fat body cell maturation. In contrast, $srp^{\Delta srpNC}$ (Figure 34B, G, L, Q, V, A') or srp^{V735G} (Figure 34C, H, M, R, W, B') embryos have shown fat body formation, since all markers are expressed normally. These results demonstrate that SrpC, but not SrpNC or the complex SrpNC/Ush, is required for fat body formation.

In order to confirm that the fat body phenotypes observed in $srp^{\Delta srpC}$ mutants are due to the loss of SrpC, we analyzed *Glt* expression in $srp^{\Delta srpC}/srp^{6G}$ trans-heterozygous embryos. Interestingly, we found that $srp^{\Delta srpC}/srp^{6G}$ embryos (Figure 35C, G) display a lower level of *Glt* expression than $srp^{\Delta srpC}$ mutants (Figure 35B, F), which is more like the phenotype observed in srp^{6G} mutants (Figure 35D, H). The lower *Glt* expression level found in $srp^{\Delta srpC}/srp^{6G}$ trans-

Cross	Percentage of adults flies devoid of TM3
+/TM3$^{twist>GFP}$ x +/TM3$^{twist>GFP}$	42 (n= 60)
$srp^{\Delta srpC}$/TM3$^{twist>GFP}$ x $srp^{\Delta srpC}$/TM3$^{twist>GFP}$	0 (n= 64)
$srp^{\Delta srpC}$/TM3$^{twist>GFP}$ x srp^{AS}/TM3$^{twist>GFP}$	45 (n= 80)

Figure 33. Loss-of-function of SrpC in *Drosophila* embryonic hemocytes has no effect on the embryos ability to develop until adult stage.

Crossing of flies with wild-type chromosome placed over third multiply balancer chromosome 3 (TM3) gives rise to progeny devoid of TM3 (+/+) that are able to develop until adult stage. Crossing of flies with loss of function of SrpC ($srp^{\Delta srpC}$) placed over TM3 gives rise to progeny devoid of TM3 ($srp^{\Delta srpC}/srp^{\Delta srpC}$) that die before reaching adult stage. Crossing of female flies of the genotype $srp^{\Delta srpC}$/TM3 with male flies of the genotype srp^{AS}/TM3 gives rise to progeny devoid of TM3 ($srp^{\Delta srpC}/srp^{AS}$) that can reach adult stage.

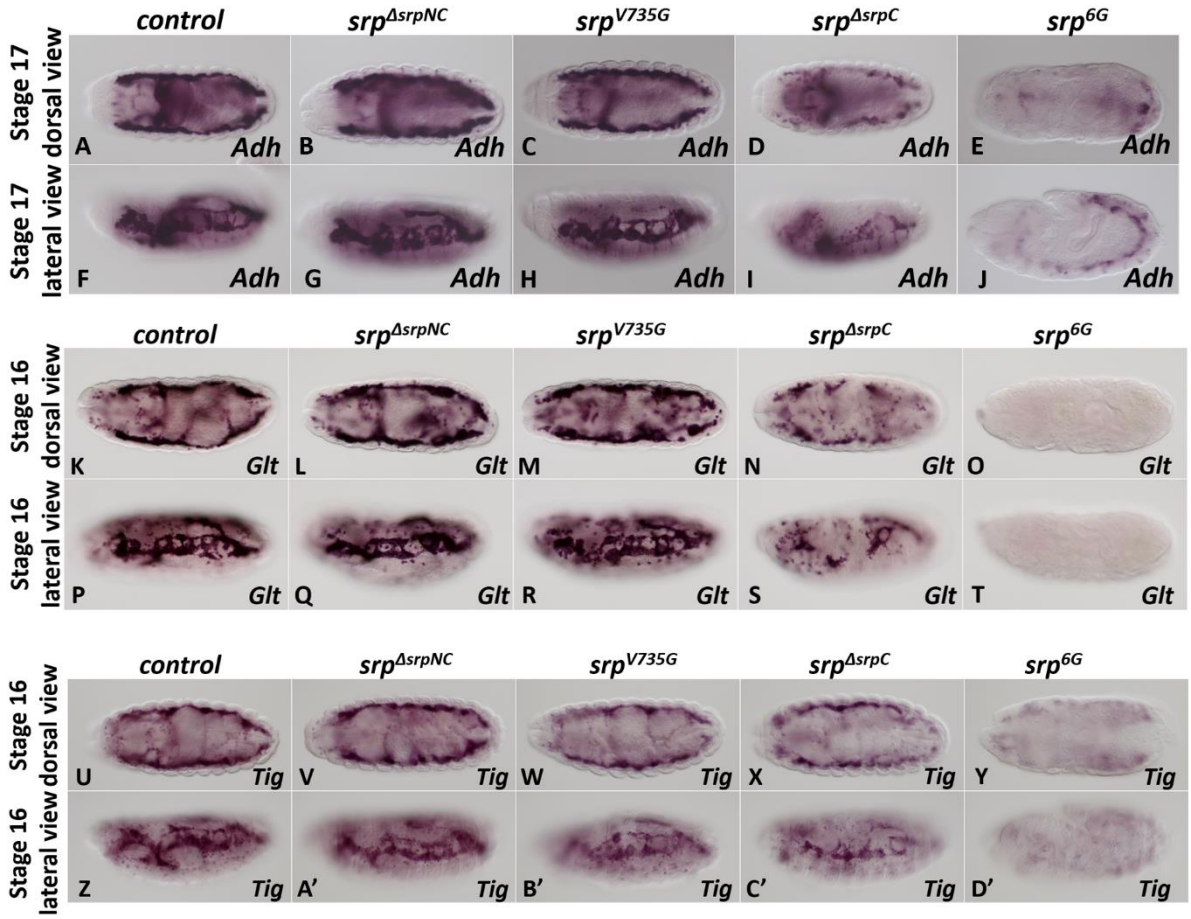


Figure 34. SrpC, but not SrpNC, is required for normal fat body development and/or fat body marker expression in *Drosophila* embryos.

Loss-of-function of both Srp isoforms completely alters the formation of normal fat body cells expressing the fat body markers *Adh* (E, J), *glt* (O, T) and *tig* (Y, D') that are normally arranged as a fat sheet at stage 16 of embryogenesis (A, F, K, P, U and Z). Loss-of-function of SrpC reduces the number of fat body cells expressing the fat body markers *Adh* (D, I), *glt* (N, S) and *tig* (X, C'). The loss-of-function of SrpNC (B, G, L, Q, V and A') or of the interaction SrpNC/Ush (C, H, M, R, W and B') has no effect on this process. *Adh*: *alcohol dehydrogenase, glt*: *glutactin, tig*: *tiggrin.* Each experiment was repeated three times with similar results.

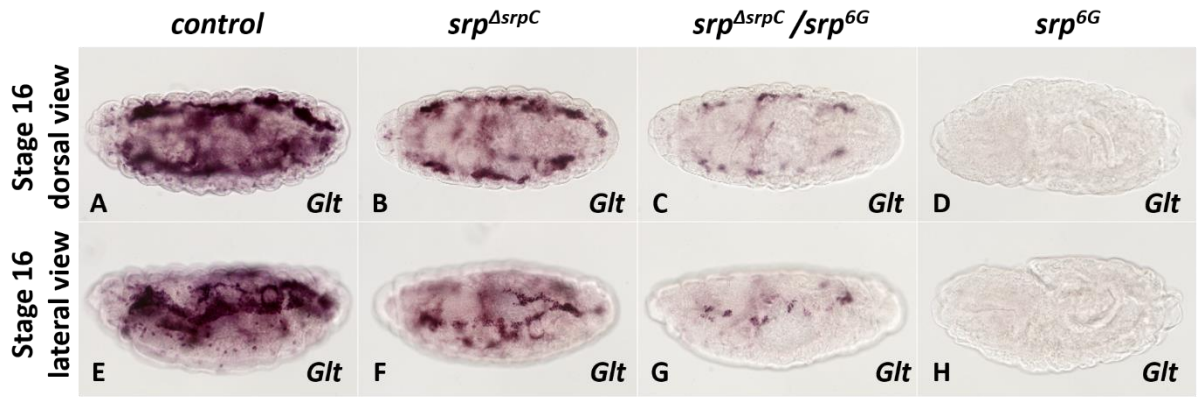


Figure 35. Validation of the association of fat body development and/or fat body marker expression alteration to Srp isoforms loss-of-function.

Embryos with loss-of-function of SrpC ($srp^{\Delta srpC}$) have altered formation of the fat body sheet expressing the fat body marker *glt* (compare B and F to A and E). srp^{6G} is a *srp* amorphic allele that abolishes fat body development (D, H). Fat body development is highly affected in embryos trans-heterozygous for $srp^{\Delta srpC}$ and srp^{6G} (D, H). *glt*: *glutactin*.

heterozygous embryos than in those homozygous for $srp^{\Delta srpC}$ mutation might be due to the fact that the srp^{6G} allele eliminates both SrpC and SrpNC isoforms, where the later could play some residual activity during *Glt* expression not abolished in $srp^{\Delta srpC}$ homozygous embryos still expressing SrpNC. In contrast, $srp^{\Delta srpC}/srp^{AS}$ allelic combinations do not display any obvious fat body defect (Figure 36), indicating that there is no dominant effect due to SrpNC overexpression when SrpC is removed. Thus, these results confirm that it is the loss of SrpC that affects fat body development and that SrpNC may only very partially compensate for it.

After identification of the role played by SrpC during fat body development, we wondered if the defects found in $srp^{\Delta srpC}$ mutant fat body are causing the lethality of these flies. We therefore set out to assess the effects of SrpC downregulation on fly viability, using double stranded RNA (dsRNA) interfering with SrpC expression specifically in the fat body. As all the *srp* dsRNA already available recognize *srp* RNA regions common to *SrpC* and *SrpNC* (Figure 37), we generated transgenic fly lines harboring a *UAS-shRNA* construct specific of the SrpC isoform. This short hairpin RNA recognizes 21 nucleotides of exon 4B, which is included in SrpC protein but excluded from SrpNC (see material and methods). Interestingly, we found that ubiquitous downregulation of *srpC* provokes lethality of the fly during early stages of development (Figure 38A). This phenotype is similar to that seen in the $srp^{\Delta srpC}$ mutant, and thus confirms the essential role played by SrpC in fly viability. However, *srpC* downregulation in the fat body was perfomed under the control of the fat body specific driver (VDRC: VT008145) that is active in fat body cells of late embryonic stages, and it remains activated in larval and pupal fat body cells (data not shown). The srpC silencing in the fat body caused lethality of the flies at pupal stage (Figure 38A, B), suggesting that the role of SrpC in the fat body is required during pupariation, while SrpC also plays additional essential functions in other tissues at earlier developmental stages. Testing the effect of *srpC* downregulation under the control of Vienna

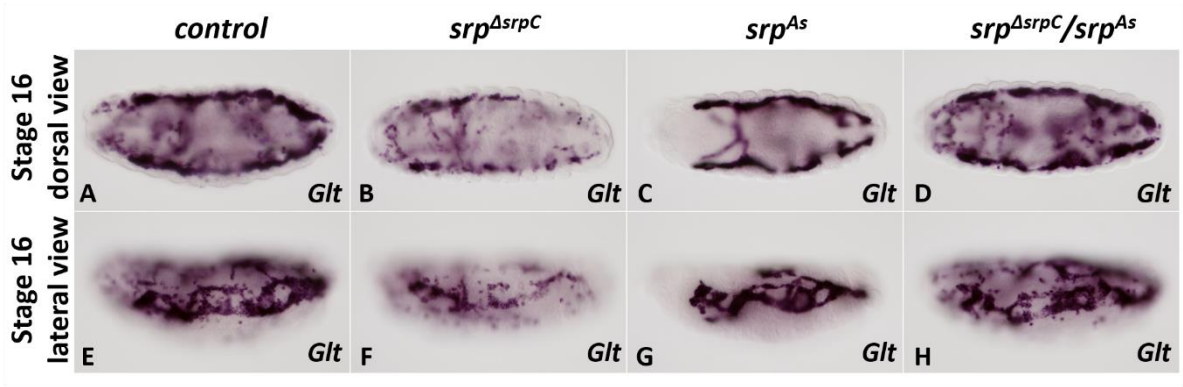


Figure 36. Validation of the dependence of normal fat body development and/or fat body marker expression on the presence of normal expression of Srp isoforms.

Embryos with loss-of-function of SrpC ($srp^{\Delta srpC}$) have altered formation of the fat body sheet expressing the fat body marker *glt* (compare B and F to A and E). srp^{As} is a *srp* allele that abolishes *srp* expression only in embryonic hemocytes but not in the embryonic fat body (C, G). Fat body development is normal in embryos of genotype $srp^{\Delta srpC}/srp^{As}$ (D, H). *glt*: *glutactin*.

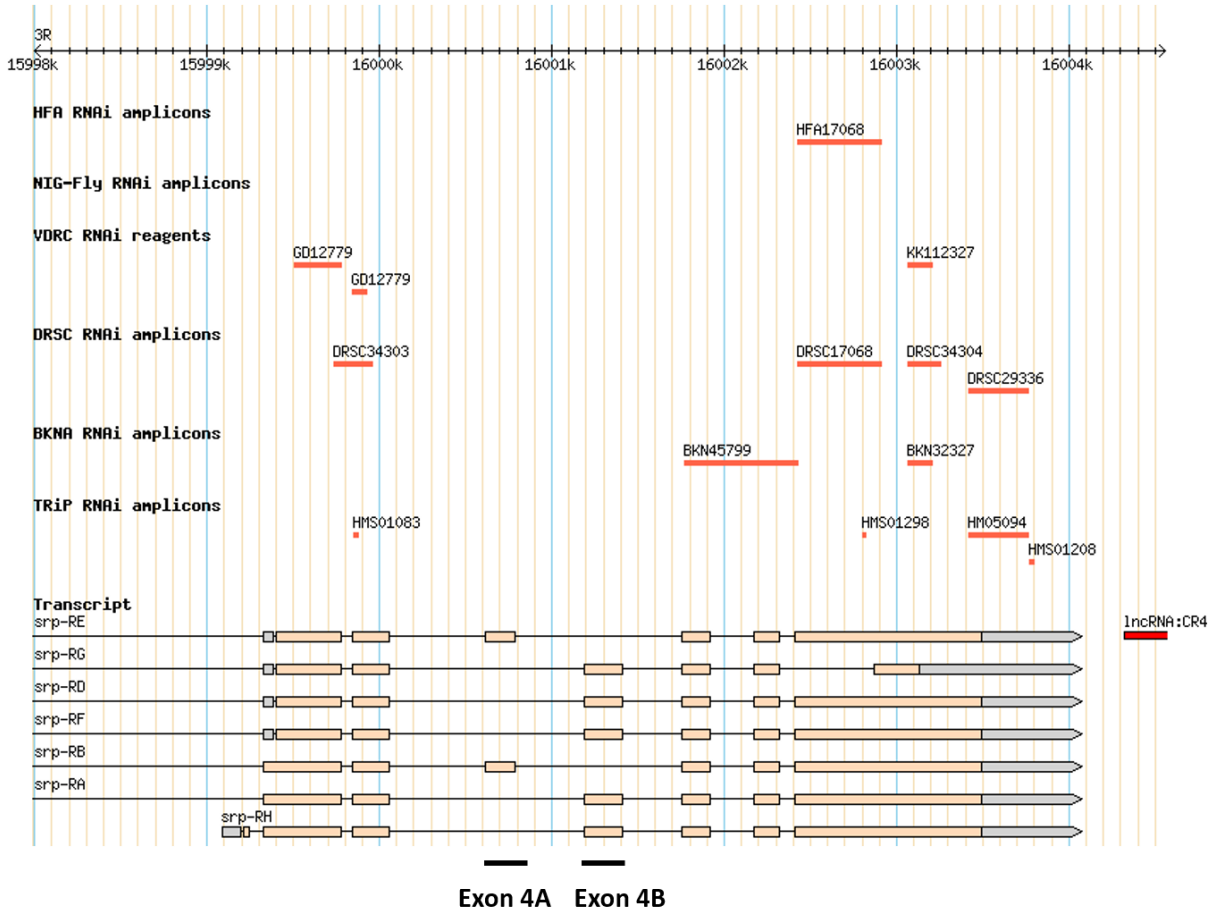


Figure 37. The double strand RNAs interfering with Srp expression available in databases recognize srp cDNA regions common to both SrpNC and SrpC.

Seven different *srp* transcripts (srp-RE, srp-RG, srpRD, srpRF, srp-RB, srp-RA, srp-RH) are schematized. The transcripts (srp-RE and srp-RB) contain exon 4A but not exon 4B. Inclusion of exon 4A in transcript results in production of SrpNC isoforms. The transcripts (srp-RG, srpRD, srpRF, srp-RA, srp-RH) contain exon 4B but not exon 4A. Inclusion of exon 4A in transcript results in production of SrpC isoforms. Six different RNAi reagents and data sources (HFA RNAi amplicons, NIG-Fly RNAi amplicons, VDRC RNAi reagents, DRSC RNAi amplicons, BKNA RNAi amplicons and TRiP RNAi amplicons) contain several RNAi (red) that all recognize regions located outside of exons E4A and E4B that, respectively, determine the type of Srp isoforms as SrpNC and SrpC.

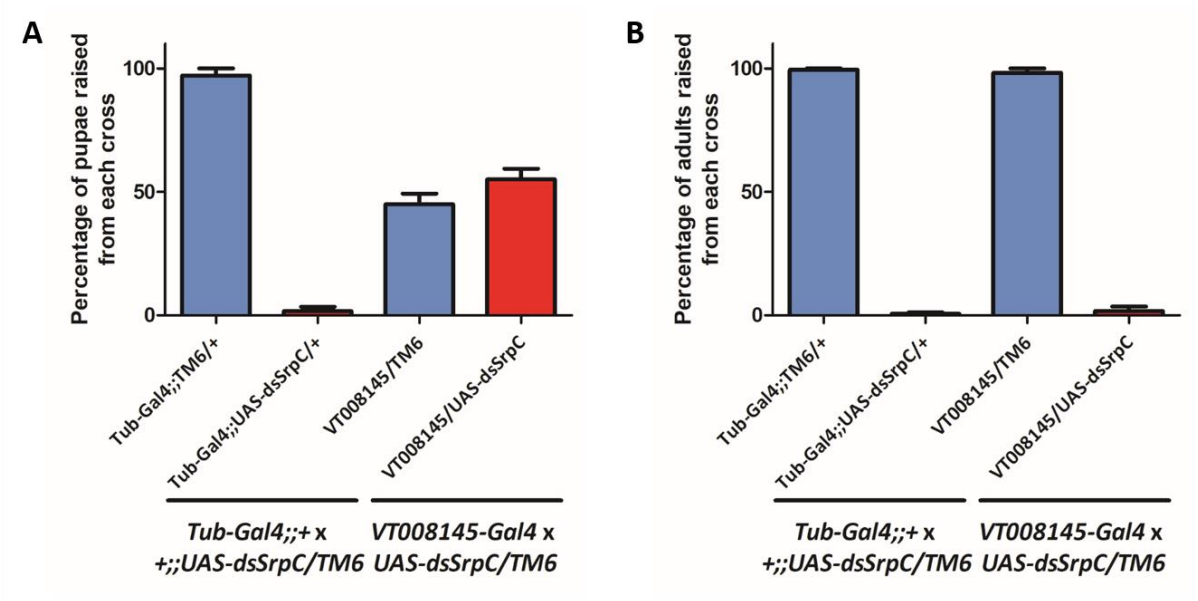


Figure 38. The down-regulation of SrpC in fat body alters the ability of the fly to pass through pupal stage.

Quantification of the percentage of pupae (A) and of adults (B) raised after crossing of flies of the genotypes *dsSrpC/TM6* with ubiquitous driver Tub-Gal4 and fat body specific driver VT008145-Gal4. The downregulation of SrpC by Tub-Gal4 (Tub-Gal4;;UAS-dsSrpC/+) affects the ability of the fly to develop until both pupal and adult stages, while its downregulation in the fat body (VT008145-Gal4/UAS-dsSrpC) alters its ability to pass through pupal stage.

Tiles drivers corresponding to different transcriptional regulatory regions located upstream of *srp* locus (and thus in tissues more likely expressing *srp* gene), on fly viability could help in identifying of additional territories that are responsible for SrpC dependent fly development.

(I) SrpNC/Ush complex maintains larval lymph gland integrity

Although no role for the isoform SrpNC has been identified in embryonic tissues expressing *srp* or regarding fly viability, we wondered if there was any role for SrpNC in *srp* expressing tissues at post-embryonic stages. *srp* is expressed at different stages of larval development in cells that constitute the lymph gland, which is the main site of post-embryonic hematopoiesis (Jung, 2005). This prompted us to determine whether SrpNC or the SrpNC/Ush complex has a role in this hematopoietic organ, and we quantified the integrity of the lymph glands in various conditions, by defining phenotypic categories (lymph glands are either intact, Figure 39A, B, partially dispersed Figure 39C, or completely dispersed, Figure 39D). Interestingly, while lymph glands of control larvae are mostly intact (Figure 39E), those of $srp^{\Delta srpNC}$ larvae are globally more dispersed, and the lymph gland bursting phenotype is even more pronounced in srp^{V735G} mutants with a disrupted SrpNC/Ush interaction, as almost 30% of the analyzed lymph glands are completely ruptured (Figure 39E). Note that all the larvae analyzed during this work are mid third instar larval stage females; finding intact lymph glands in $srp^{\Delta srpNC}$ and srp^{V735G} mutants at this stage allows us to analyze the effect of the loss-of-function of SrpNC and of the SrpNC/Ush interaction, respectively, on the formation of all of the lymph gland hematopoietic cell types, including the posterior signaling center (PSC) cells, progenitor cells, plasmatocytes and crystal cells.

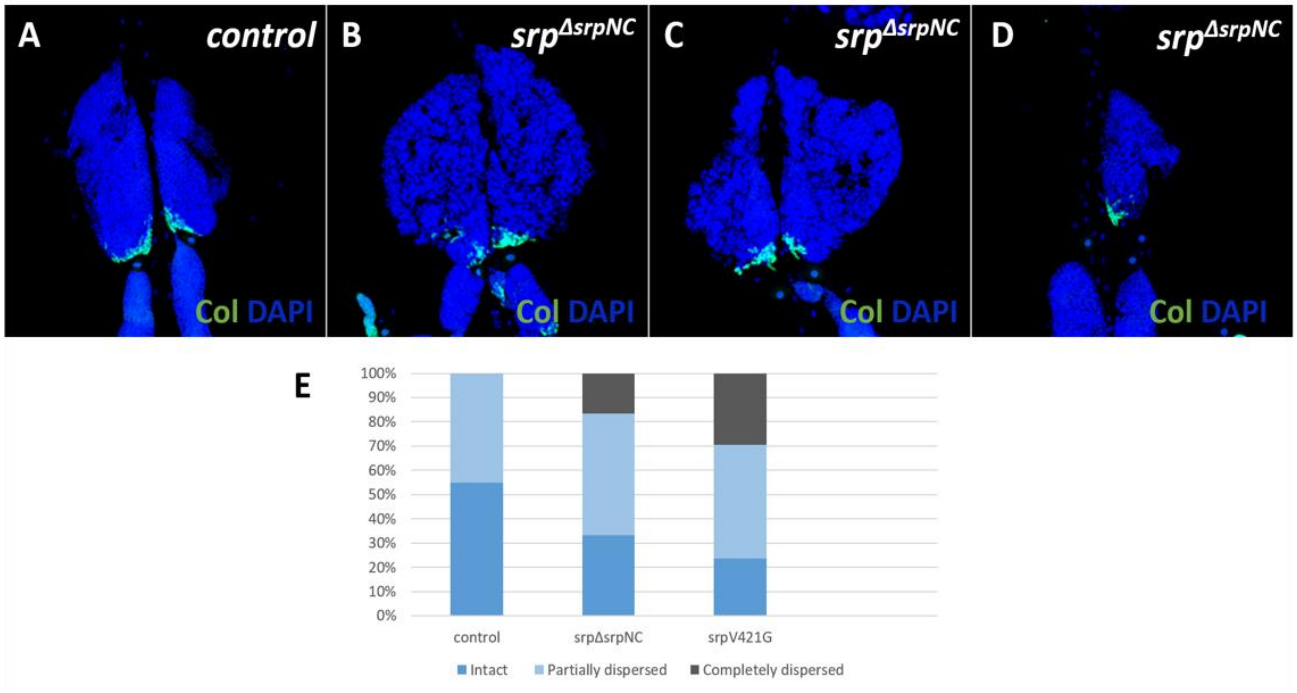


Figure 39. Mutations in SrpNC isoform affect lymph gland integrity.

Lymph glands of control (A) or mutant $srp^{\Delta srpNC}$ flies (B-D) stained with anti-Collier (Col) to mark posterior signaling center and with DAPI to mark nuclei. (E) Stacked histogram representing the percentage of control, $srp^{\Delta srpNC}$ and srp^{V735G} having intact (dark blue, B), partially dispersed (light blue, C) and completely dispersed (grey, D) lymph glands.

(J) SrpNC/Ush complex regulates lymph gland prohemocyte proliferation and differentiation

One of the markers used during lymph gland constituent characterization is Collier (Col). Col is expressed in lymph gland progenitor cells as well as in the PSC (Figure 41C) (Benmimoun et al., 2015; Crozatier et al., 2004). In agreement with the dispensability of Srp during the PSC formation already described (Crozatier et al., 2004), we found that in $srp^{\Delta srpNC}$ (Figure 40B) as well as in srp^{V735G} mutants (Figure 40C), PSC cells expressing Col are specified and maintained as in wild-type lymph glands (Figure 40A).

The progenitor cells are located in the lymph gland anterior lobe medullary zone, as well as in the lymph gland posterior lobes (Jung et al., 2005). Staining of the medullar zone in the $srp^{\Delta srpNC}$ mutant with an α-Col antibody shows a strong reduction of the medullary zone size (Figure 41D) in comparison to the control (Figure 41C). In addition, in lymph gland posterior lobes, we observed some progenitor cell differentiation into lamellocytes, which specifically initiate expression of the *misshapen-mCherry* transgene (*msn-mCherry*) (Figure 42H). However, in addition to developing a lamellocyte fate, posterior lobe progenitor cells also differentiate into crystal cells, as in about 83% of $srp^{\Delta srpNC}$ larvae, the crystal cell marker Prophenoloxidase 1 (PPO1) is also detected (Figure 42I). Taken together, these results show that SrpNC is required for the maintenance of hematopoietic progenitor cell identity and for inhibition of their differentiation into mature hemocytes. Interestingly, similar results have been observed in the posterior lobes of the mutant srp^{V735G} (Figure 42M), indicating that the interaction SrpNC/Ush is required to inhibit differentiation of lymph gland progenitor cells in the posterior lobes.

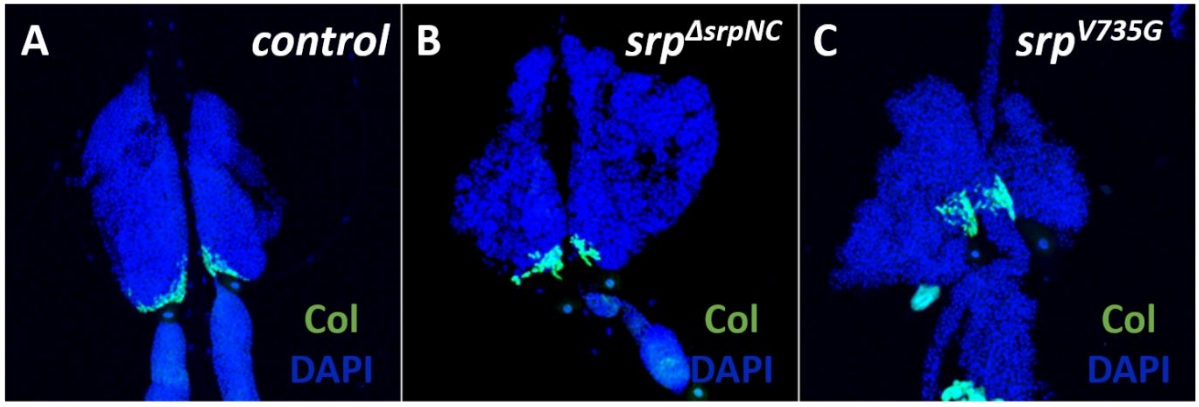


Figure 40. SrpNC is dispensable for lymph gland posterior signaling center formation.

Lymph glands of control (A), $srp^{\Delta srpNC}$ (B) and srp^{V735G} larvae (C) stained with anti-Collier (Col) (green) to mark posterior signaling center and with DAPI (blue) to mark nuclei. Collier is produced in the posterior signaling center in all contexts.

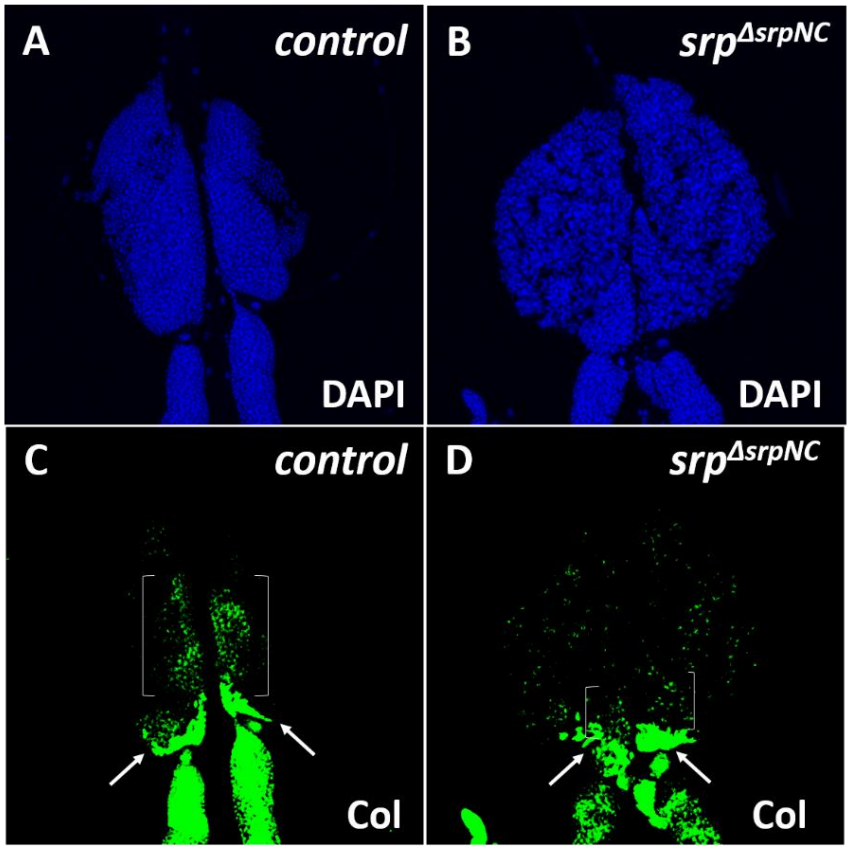


Figure 41. SrpNC is required for lymph gland prohemocyte maintenance.

Lymph glands of control (A, C) and mutant *srp*$^{\Delta srpNC}$ flies (B, D) stained with anti-Collier (Col) (C, D) to mark posterior signaling center and with DAPI (A, B) to mark nuclei. In the anterior lobes, Collier is highly expressed in the posterior signaling center (arrows), while it is slightly expressed in the medullary zone (bracket) of control lymph glands. Collier expression in the medullary zone of *srp*$^{\Delta srpNC}$ mutants is altered.

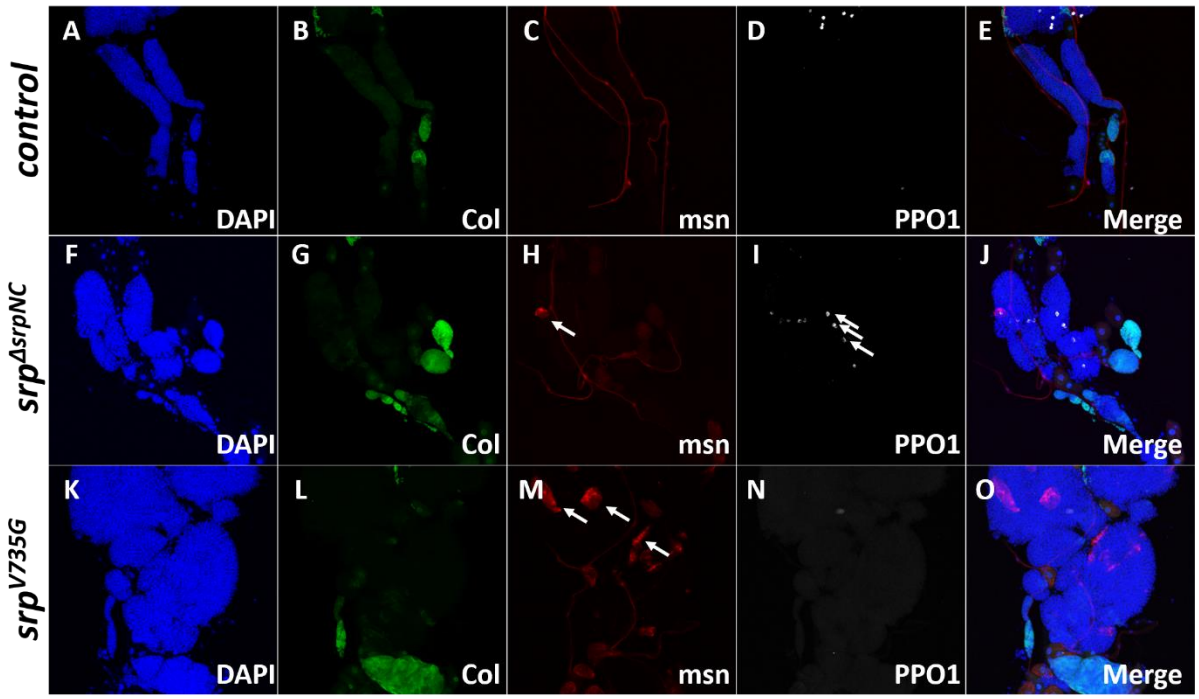


Figure 42. SrpNC is required for regulation of lymph gland prohemocyte proliferation and differentiation.

Lymph glands of control (A-E), *srp*$^{\Delta srpNC}$ (F-J) and *srp*^{V735G} (K-O) flies stained with DAPI (A, F, K) to mark nuclei, with anti-Collier (Col) (B, G, L) to mark posterior lobes and with Prophenoloxidase1 (PPO1) (D, I, N) to mark crystal cells. Both control and mutant larvae have the transgene misshapen (msn)-mCherry that activates mCherry expression under the control of msn regulatory region that is activated in lamellocytes but no other hematopoietic cell types. Posterior lobes of *srp*$^{\Delta srpNC}$ contain differentiated crystal cells and lamellocytes. Posterior lobes of *srp*^{V735G} are hypertrophic and contain lamellocytes.

In addition to regulating posterior lobe hematopoietic progenitor cell differentiation, we found that the SrpNC/Ush complex also regulates the size of the posterior lobes, as lobes of srp^{V735G} larvae are significantly hypertrophic in comparison to those of control larvae (Figure 42O, E). This result indicates that the SrpNC/Ush interaction is either required for regulation of progenitor cell proliferation or to control their survival rate. Analysis of the number of progenitor cells expressing the cellular division marker PHH3, or the anti-apoptotic marker p53, should give more information about this phenotype. Accordingly, a role for Ush in the control of lymph gland progenitor cell proliferation and lymph gland size has already been reported (H. Gao et al., 2009, 2016; R. P. Sorrentino et al., 2007). Note that the lymph gland hypertrophic phenotype seen in the mutant $srp^{\Delta srpNC}$ (Figure 42J) is not as pronounced as in the mutant srp^{V735G} (Figure 42O), suggesting that SrpNC and SrpNC/Ush complex ensure different functions in lymph gland progenitor cell regulation.

(K) SrpNC/Ush interaction is required for lymph gland crystal cell formation

According to what we found at embryonic stages, SrpNC is dispensable for larval lymph gland plasmatocytes and crystal cell formation, as in $srp^{\Delta srpNC}$ mutants both plasmatocytes expressing the P1 marker (Figure 43B) and crystal cells expressing the PPO1 marker (Figure 43D) are detected at mid-third instar larvae. However, in the lymph gland of the srp^{V735G} mutant, both the crystal cell progenitors expressing Lz (Figure 44A) and the mature crystal cells expressing PPO1 (Figure 44B) are not detected, suggesting that although SrpNC is not essential for the activation of the lymph gland crystal cell differentiation program, its presence free of the interaction with Ush blocks this process. In contrary, formation of plasmatocytes is not affected by loss of the interaction SrpNC/Ush (Figure 45B). These results show that the SrpNC/Ush

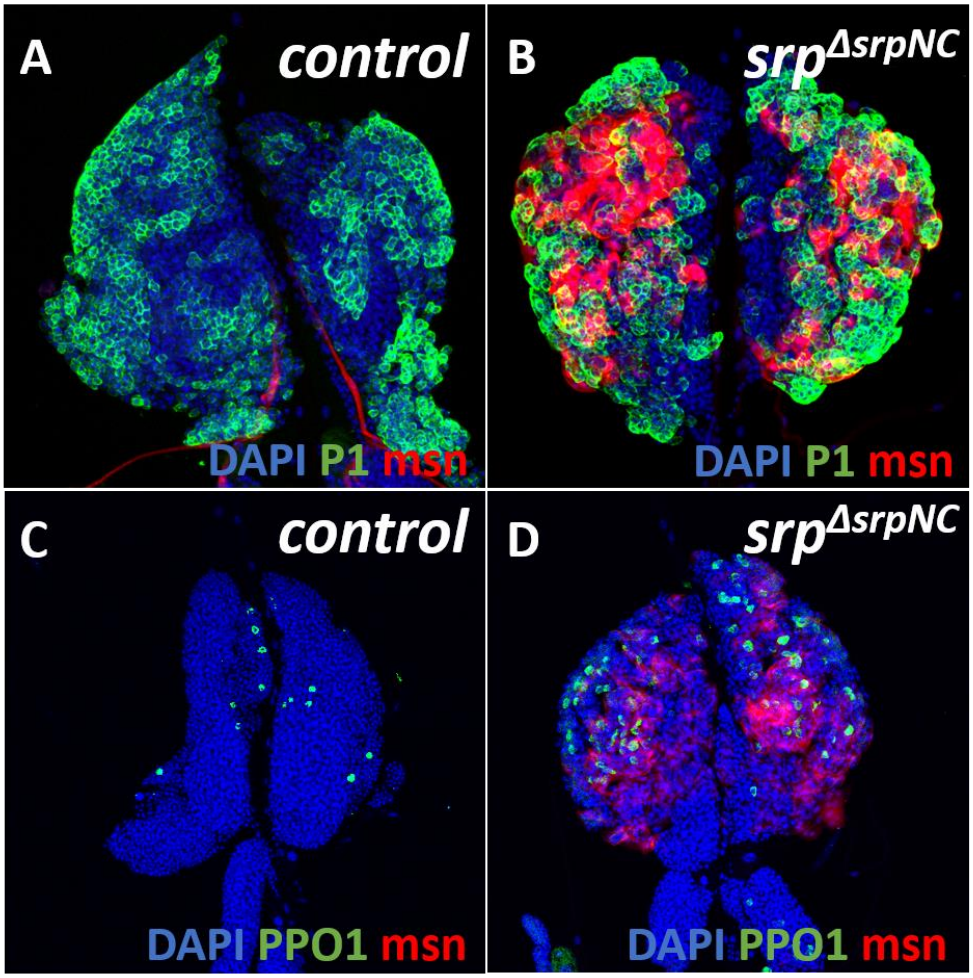


Figure 43. The isoform SrpNC is essential to inhibit lamellocyte formation.

In the lymph gland of larvae with loss-of-function of SrpNC (B, D), as in control larvae (A, C), plasmatocytes express P1 (A, B) and crystal cells express PPO1 (C, D). Lamellocytes activating *msn* expression, which is normally absent in lymph gland (A, C), become highly produced in mutants with loss-of-function of SrpNC (B, D). PPO1: Prophenoloxidase 1, *msn*: *misshapen.* Each experiment was repeated more than three times with similar results.

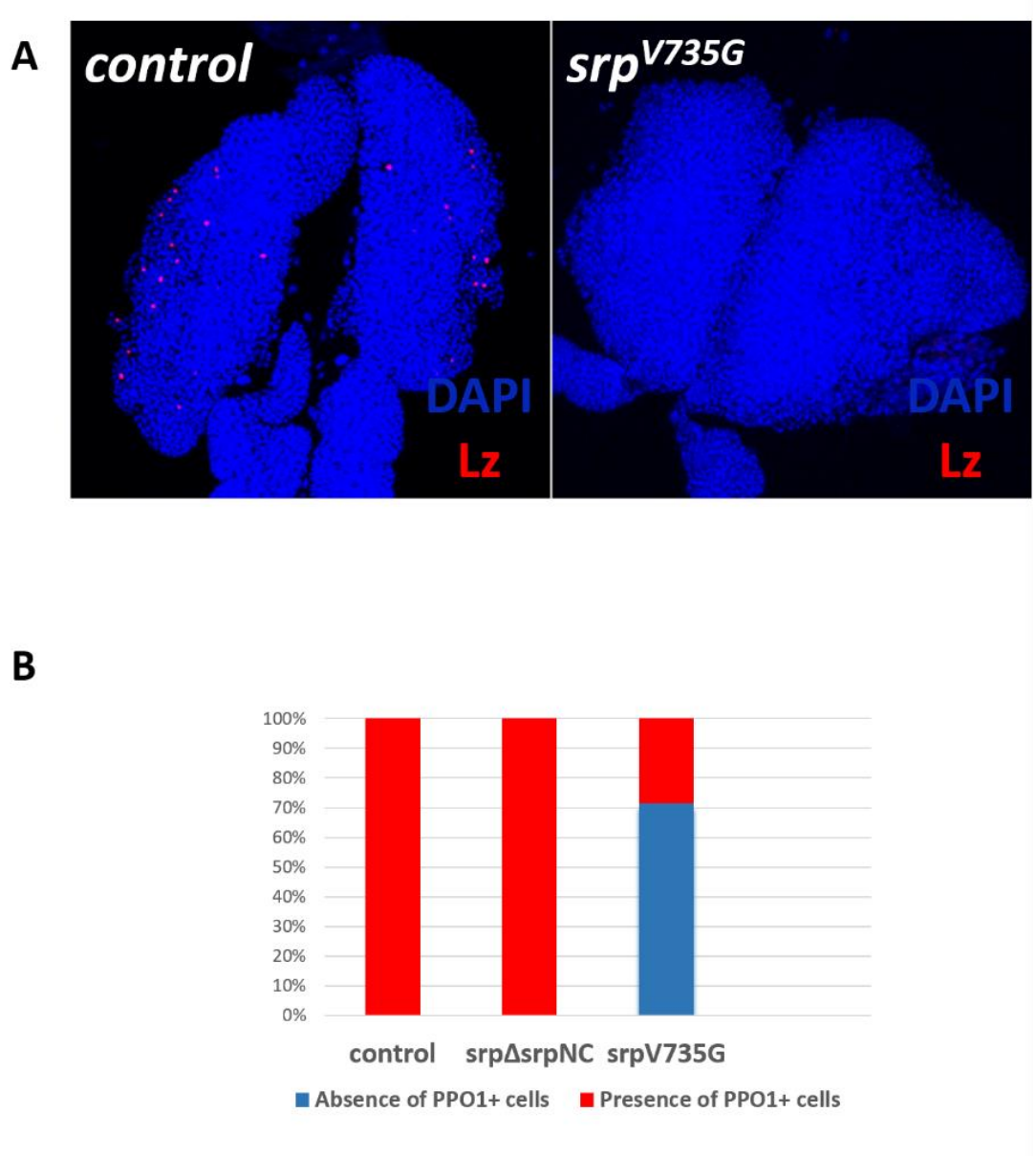


Figure 44. SrpNC/Ush complex is required for regulation of crystal cell formation.

(A) Lymph glands of control and of *srp*^{V735G} mutant flies stained with DAPI to mark nuclei and with anti-Lozenge (Lz) to mark crystal cell progenitors. No staining for Lz has been detected in *srp*^{V735G} lymph gland. (B) Stacked histogram representing the percentage of control, *srp*$^{\Delta srpNC}$ and *srp*^{V735G} larvae containing mature Prophenoloxidase 1 (PPO1) expressing crystal cells (red) or devoiding of mature crystal cells (blue). The majority of *srp*^{V735G} larval lymph gland don't contain PPO1 expressing crystal cells.

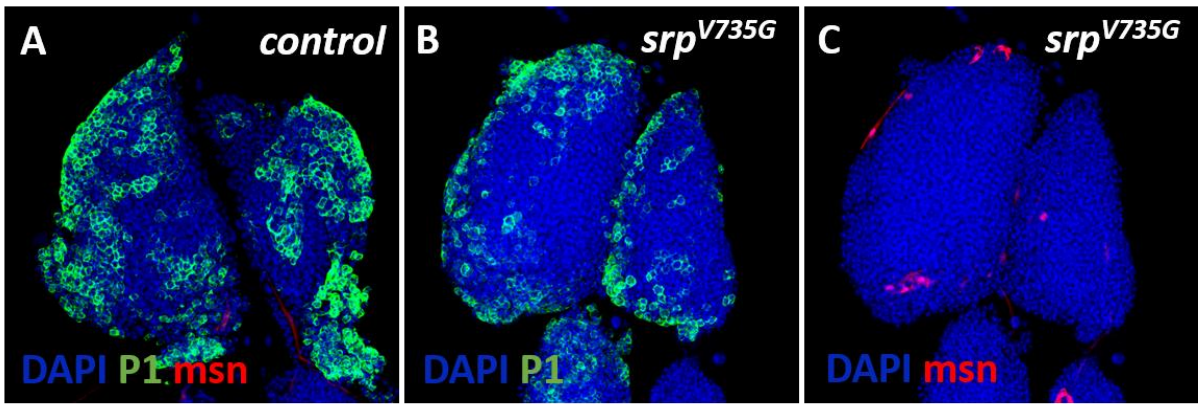


Figure 45. The complex SrpNC/Ush is essential to inhibit lamellocyte formation.

In the lymph gland of larvae with loss of the SrpNC/Ush interaction (B, C), as in control larvae (A), plasmatocytes expressing P1 (A, B) are formed. Lamellocytes activating msn expression are not detected in the control lymph gland (A) but they are produced in mutants with loss of SrpNC/Ush interaction. *msn*: *misshapen.* This experiment was repeated three times with similar results.

complex is required for crystal cell but not plasmatocyte formation. The different effects promoted by SrpNC alone and the SrpNC/Ush complex on crystal cell formation suggest an antagonistic role played by Ush and SrpNC during this process.

(L) SrpNC/Ush is required for inhibition of lamellocyte formation

Although the loss of SrpNC function has no effect on plasmatocyte and crystal cell formation, $srp^{\Delta srpNC}$ mutant lymph glands reveal an inappropriate production of lamellocytes (as shown by their expression of the *msn-mCherry* reporter gene) at mid-third instar larval stage (Figure 43B, D), while these cells are normally absent in wild-type lymph glands (Figure 43A, C). Thus, SrpNC is required for inhibition of lamellocyte formation. Consistent with premature differentiation of lamellocytes and loss of integrity in the $srp^{\Delta srpNC}$ mutant lymph glands, lamellocytes were also detected in the hemolymph of $srp^{\Delta srpNC}$ larvae by the end of the mid-third instar larval stage (Figure 46B, G) and not in control larvae (Figure 46A, G). Interestingly, similar results were obtained in the srp^{V735G} mutant with loss of the interaction SrpNC/Ush, as lamellocytes were detected both in lymph glands and in hemolymph of homozygous srp^{V735G} (Figure 45C and 46D, G) indicating that the interaction SrpNC/Ush is required for inhibition of lamellocyte production. These results are consistent with the already published role for Ush in the inhibition of lamellocyte formation (Avet-Rochex et al., 2010; R. P. Sorrentino et al., 2007) (Figure 46F, G).

(M) SrpNC/Ush, but not SrpC, has a dosage sensitive effect on lamellocyte formation

Following identification of a role for the SrpNC/Ush complex in the inhibition of lamellocyte

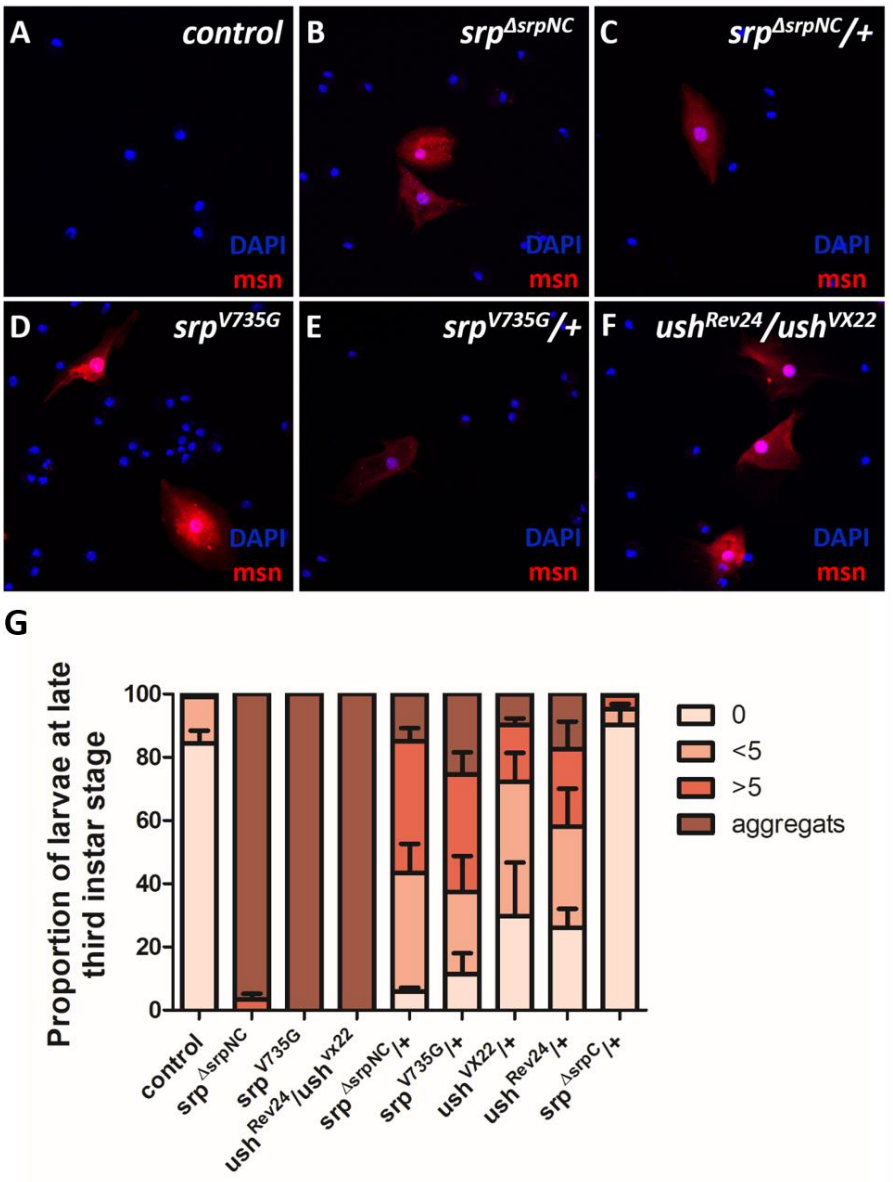


Figure 46. SrpNC/Ush complex, but not SrpC, has a dominant effect on lamellocyte formation.

Lamellocytes activating *msn* expression are produced in larvae homozygous (B), heterozygous (C) for SrpNC loss-of-funciton, homozygous (D) and heterozygous (E) for SrpNC/Ush interaction loss and trans-heterozygous for Ush loss-of-function (E), but not in control larvae (A). (G) Quantification of the percentage of larvae producing lamellocytes shows that almost all larvae homozygous for SrpNC loss-of-function ($srp^{\Delta srpNC}$), for SrpNC/Ush interaction loss (srp^{V735}) and for Ush loss-of-function (ush^{VX22}/ush^{rev24}) produce lamellocytes in their body. The phenotype is less severe in larvae heterozygous for SrpNC loss-of-function ($srp^{\Delta srpNC}/+$), for SrpNC/Ush interaction loss ($srp^{V735}/+$) and for Ush loss-of-function ($ush^{VX22}/+$ and $ush^{rev24}/+$), whereas in almost all larvae heterozygous for the loss-of-function of SrpC ($srp^{\Delta srpC}/+$), as in control larvae, no lamellocytes are produced. *msn*: *misshapen.* Quantification of lamellocyte formation penetrance in all genotypes has been repeated three times with similar results.

formation, we analyzed the effect of the loss-of-function of only one copy of SrpNC or of Ush. Lamellocytes are detected in the hemolymph of at least 70% of flies in the different heterozygous background for the $srp^{\Delta srpNC}$, srp^{V735G}, ush^{Rev24}, ush^{VX22} mutations (Figure 46G), indicating that control of lamellocyte production depends on the dose of SrpNC and Ush. Of note, more than 80% of $srp^{\Delta srpC}$/+ heterozygous larvae are completely devoid of lamellocytes like wild-type larvae (Figure 46G), indicating that, contrary to the complex SrpNC/Ush, the isoform SrpC doesn't display a dosage sensitive effect on lamellocyte formation. In order to more analyze the role of SrpC during lamellocyte formation, I analyzed the hemolymph of larvae having downregulation of SrpC expression in all hemocytes using the *Collagen-Gal4* (*Cg-Gal4*) driver. Only few of these larvae have lamellocytes activating *misshapen* (*msn*) expression in their hemolymph indicating that SrpC might contribute but slightly in the inhibition of lamellocyte differentiation (Figure 47).

(N) SrpNC/Ush complex is dispensable for ovogenesis

A recent study demonstrated that *srp* is expressed and plays essential functions in adult ovaries also (Lepesant et al., 2020). Interestingly, we found that although $srp^{\Delta srpNC}$ and srp^{V735G} flies develop apparently normally until adult stage, adult female mutants are sterile. Thus, the SrpNC/Ush complex is required for female fertility. In order to identify the defects responsible for the $srp^{\Delta srpNC}$ and srp^{V735G} female sterility, we checked for ovary formation in both mutant fly lines, but no noticeable defect in general morphology of the ovaries was seen, as in both wild-type and $srp^{\Delta srpNC}$ female flies ovarioles are formed and the linear sequences of egg chambers that constitute the ovariole are also formed (Figures 48B, C, 49B, C, 50B, C).

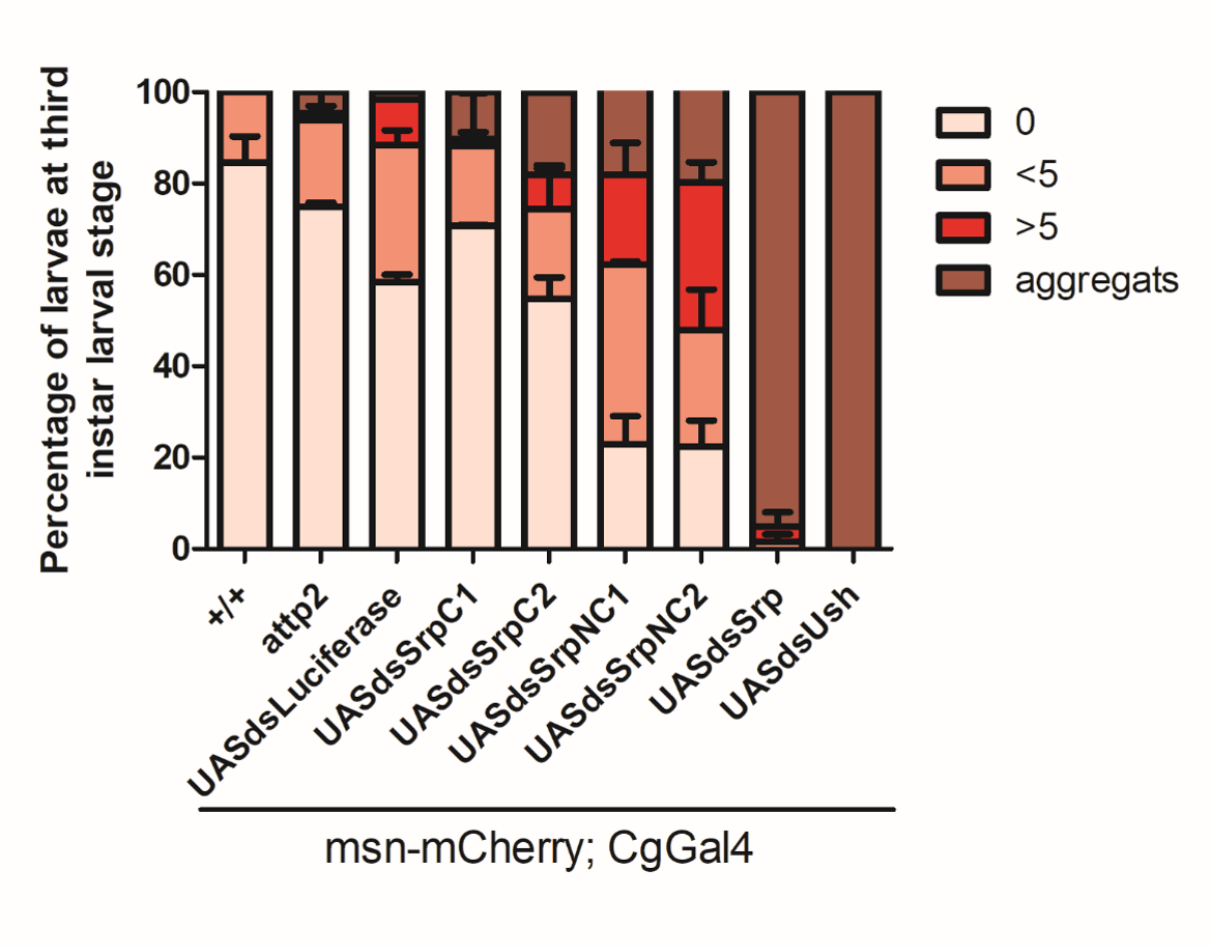


Figure 47. SrpC might regulate, but slightly, lamellocyte formation.

Quantification of the percentage of larvae producing lamellocytes activating *misshapen* regulatory region (msn-mCherry) shows that almost all larvae having Srp (UASdsSrp) or Ush (UASdsUsh) downregulation under the control of *Collagen-Gal4* (*Cg-Gal4*) driver have aggregates of lamellocytes in their hemolymph. About 80% of larvae having SrpNC downregulation (UASdsSrpNC1 and UASdsSrpNC2) produce lamellocytes in their hemolymph. The lamellocyte production is less important by inducing expression of interfering RNA against SrpC (UASdsSrpC1 and UASdsSrpC2) and against *Luciferase* (control). In the control larvae (+/+ and attp2), lamellocytes are rarely detected.

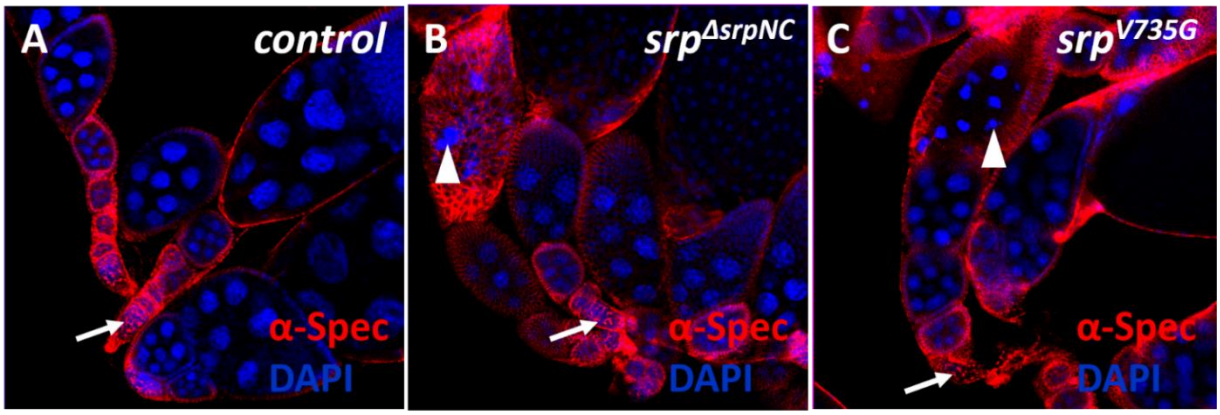


Figure 48. SrpNC is dispensable for ovarian germline cell formation.

Ovaries stained with anti-alpha Spectrin (α-Spec) to mark cytoplasmic structures between germline cells and with DAPI to mark nuclei of control (A), $srp^{\Delta srpNC}$ (B) and srp^{V735G} (C) flies. No differences of Spectrin coloration were found between *srp* mutants and control ovaries (arrows). In some $srp^{\Delta srpNC}$ and srp^{V735G} egg chambers, nuclei of nurse cells have pyknotic morphology (arrowheads).

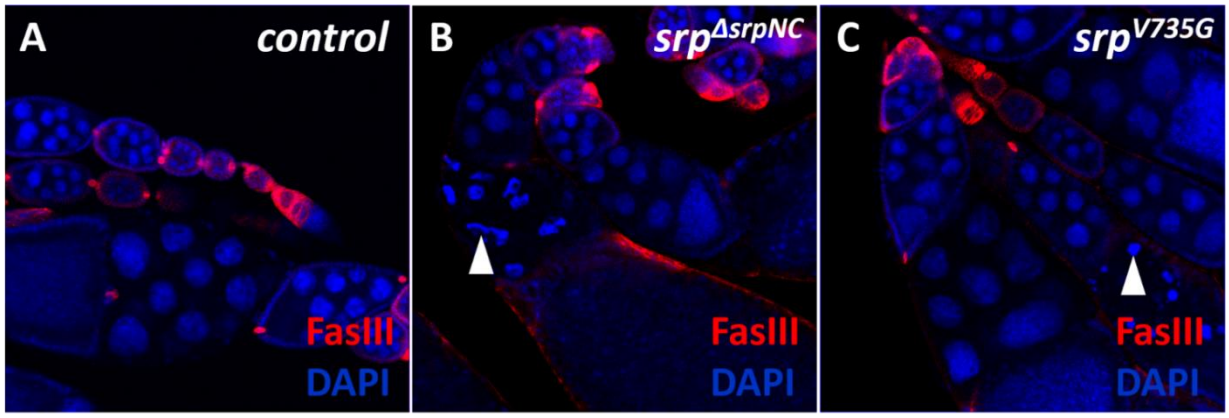


Figure 49. SrpNC is dispensable for ovarian early stages follicle cell formation.

Ovaries stained with anti-Fasciclin III (FasIII) to mark follicular cells at early stages of oogenesis and with DAPI to mark nuclei of control (A), $srp^{\Delta srpNC}$ (B) and srp^{V735G} (C) flies. No differences of Fasciclin coloration were found between *srp* mutants and control ovaries. In some $srp^{\Delta srpNC}$ and srp^{V735G} egg chambers, nuclei of nurse cells have pyknotic morphology (arrowheads).

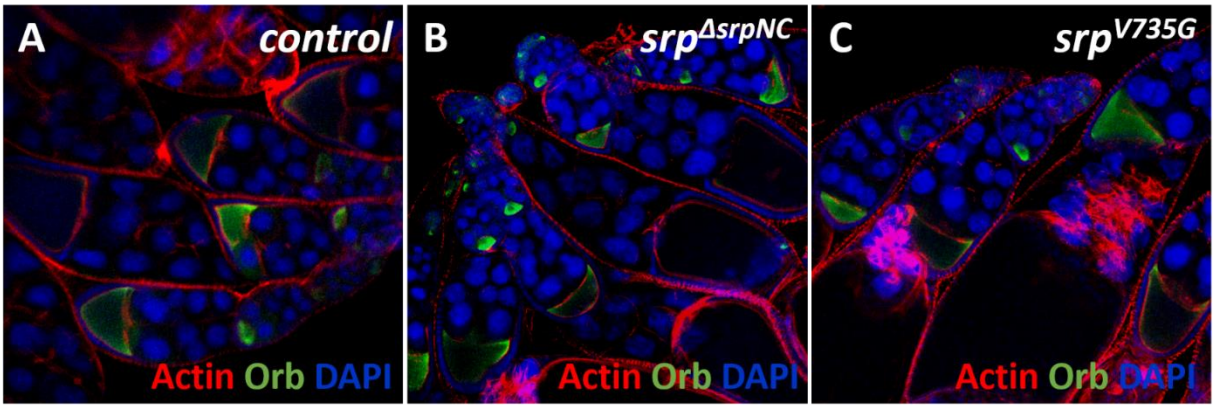


Figure 50. SrpNC is dispensable for ovarian oocyte formation.

Ovaries stained with anti-oo-RNA binding protein (Orb) to mark oocytes and with Phalloidin (Actin) to mark filamentous actin, and with DAPI to mark nuclei of control (A), $srp^{\Delta srpNC}$ (B) and srp^{V735G} (C) flies. No differences of actin or of Orb coloration were found between *srp* mutants and control ovaries.

We next analyzed the formation of different ovarian components in the ovaries of these mutants. The unit in the ovaries is the egg chamber. Each ovarian egg chamber contains 15 nurse cells and one oocyte, surrounded by a monolayer of epithelial follicle cells (E. Fedorova et al., 2019). Developing germline cells that are at the origin of nurse cells and oocyte are spanned by large cytoplasmic structures that can be visualized by different membrane skeletal proteins including α-spectrin (de Cuevas et al., 1996) (Figure 48A). The follicle cells express cell adhesion molecule fasciclin III during early stages of oogenesis, but later Fasciclin III expression becomes restricted to polar follicle cells (St. Johnston, 2001) (Figure 49A). Finally, the developing oocyte is characterized by the expression of the RNA binding protein Orb (Lantz et al., 1994) (Figure 50A). Staining for α-spectrin, Fasciclin III and Orb proteins in $srp^{\Delta srpNC}$ mutants (Figure 48B, 49B, 50B) and in srp^{V735G} ovaries (Figure 48C, 49C, 50C) showed no significant differences of expression in comparison to control ovaries (Figure 48A, 49A, 50A). These results indicate that the germline cells, the follicle cells and the oocyte are formed normally in the absence of SrpNC isoform and of SrpNC/Ush interaction in the early egg chambers, and this suggest that the sterility of $srp^{\Delta srpNC}$ and srp^{V735G} mutants is more likely due to defects during the later stages of oogenesis.

(O) SrpNC/Ush complex, but not SrpC, is required for normal oogenesis

In some cases, in $srp^{\Delta srpNC}$ (Figures 48B, 49B) and srp^{V735G} (Figures 48C, 49C) mutants, we observed pycnotic morphology of the DAPI-stained nuclei of the nurse cells. This phenotype of degeneration of the mid-stage egg chambers in some *Drosophila* egg chambers has been already detected in ovaries of flies with downregulation of both Srp isoforms, and analyzed as the consequence of alteration of the ovarian mid-oogenesis checkpoint function, which detects if an egg chamber has to die or to survive and produce an egg (Lepesant et al., 2020). However,

in $srp^{\Delta srpNC}$ mutants, this phenotype was not very pronounced, indicating that fertility problems in these mutants are not solely due to alteration of the mid-oogenesis checkpoint and that other defects could be responsible.

Recently, it was shown that *srp* is required for *Drosophila* oogenesis and/or egg laying (Lepesant et al., 2020). Similarly, we found that, while wild-type flies lay eggs abundantly, $srp^{\Delta srpNC}$ or srp^{V735G} female flies lay only very few eggs (Figure 51A). This indicates that a SrpNC/Ush interaction is required for normal oocyte release.

In addition, embryos laid by $srp^{\Delta srpNC}$ (Figure 51C, F) and srp^{V735G} (Figure 51D, G) flies show clear physical defects, caused by an absence of the *Drosophila* eggshell, of the egg respiratory appendages (the dorsal appendages), and of the egg sperm entry point (the micropyle) that are normally formed in eggs laid by wild-type females (Figure 51B, E). This indicates that the SrpNC isoform is required to interact with Ush, for normal *Drosophila* egg maturation. In agreement with the role of Ush in oogenesis proposed here, analysis of the Fly atlas expression data shows that Ush is highly expressed in the *Drosophila* ovary (Table 1), which supports the idea of a function for Ush during oogenesis. In order to validate the role of Ush during egg maturation, we induced the expression of Ush interfering RNA (VDRC: 5712) in ovarian follicle cells of normal unmutated flies using the insertion *traffic-jam-Gal4* (*Tj-Gal4*) (Lepesant et al., 2020) that is activated in the ovarian follicle cells, and we found that contrarily to embryos laid by control flies (Figure 52A), embryos laid by flies having Ush downregulation (Figure 52B) have morphological defects identical to those of embryos laid by $srp^{\Delta srpNC}$ and srp^{V735G} mutants including an absence of dorsal appendages, micropyle and eggshell.

In order to determine the role of the SrpC isoform in fly oogenesis, we analyzed the effect on

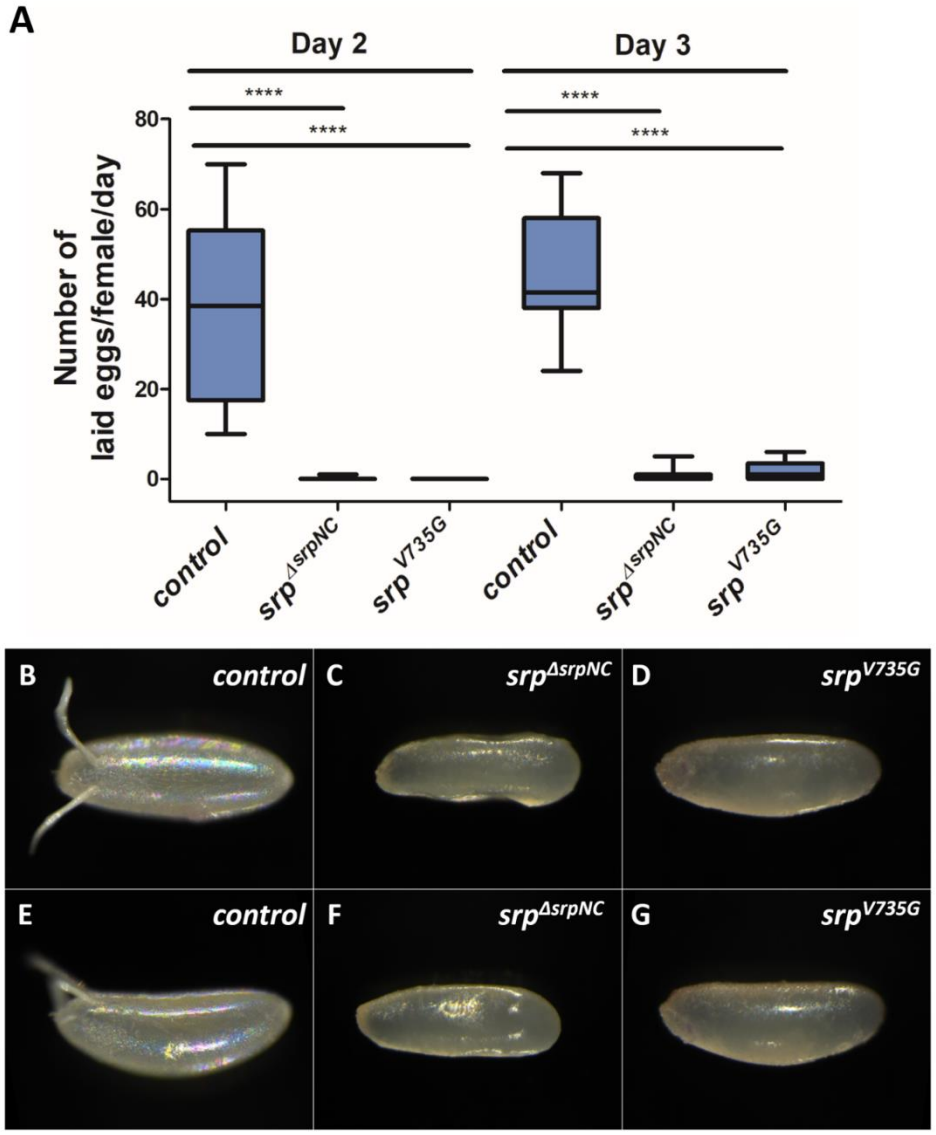


Figure 51. SrpNC/Ush complex is required for normal egg formation and maturation.

(A) Quantification of the number of eggs laid per female in 2 and 3 days after mating, for the genotypes *control* (n=12), $srp^{\Delta srpNC}$ (n=10) and srp^{v735G} (n=13). The Mann-Whitney test was used to compare the phenotype of different genotypes (**** means P-value<0.0001). Eggs laid by flies $srp^{\Delta srpNC}$ (C, F) or srp^{v735G} (D, G) have remarkable physical defects in comparison to normal eggs (B, D). The laid eggs quantification and observation in all genotypes have been repeated at least three times with similar results.

Tissue	mRNA Signal	Present Call	Enrichment	Affy Call
Brain	46 ± 1	4 of 4	0.90	None
Head	41 ± 2	4 of 4	0.80	None
Eye	23 ± 4	3 of 4	0.47	Down
Thoracicoabdominal ganglion	60 ± 4	4 of 4	1.20	None
Salivary gland	41 ± 4	4 of 4	0.84	None
Crop	15 ± 2	4 of 4	0.30	Down
Midgut	6 ± 2	0 of 4	0.10	Down
Tubule	17 ± 1	3 of 4	0.40	Down
Hindgut	20 ± 1	4 of 4	0.40	Down
Heart	158 ± 13	4 of 4	3.23	Up
Fat body	20 ± 6	4 of 4	0.42	Down
Ovary	100 ± 3	4 of 4	2.00	Up
Testis	21 ± 2	4 of 4	0.40	Down
Male accessory glands	115 ± 3	4 of 4	2.40	Up
Virgin spermatheca	78 ± 2	4 of 4	1.61	Up
Mated spermatheca	96 ± 9	4 of 4	1.98	Up
Adult carcass	32 ± 3	4 of 4	0.70	Down
Larval CNS	105 ± 7	4 of 4	2.16	Up
Larval Salivary gland	13 ± 1	0 of 4	0.28	Down
Larval midgut	78 ± 4	4 of 4	1.60	Up
Larval tubule	20 ± 2	4 of 4	0.40	Down
Larval hindgut	27 ± 0	4 of 4	0.56	Down
Larval fat body	82 ± 9	4 of 4	1.70	Up
Larval trachea	54 ± 6	4 of 4	1.12	None

Larval carcass	72 ± 5	4 of 4	1.48	Up
S2 cells (growing)	662 ± 9	4 of 4	13.54	Up
Whole fly	49 ± 2	4 of 4		

Table 1: U-shaped is highly expressed in *Drosophila* ovaries.

Analysis of *Drosophila* gene expression atlas Flyatlas indicates that U-shaped (accession number: CG2762) mRNA is differentially expressed in different *Drosophila* tissues. Values of the mRNA signal (how abundant the mRNA is), of the mRNA enrichment (compared to whole flies), and of the affymetrix present (out of 4 arrays, how many times it was detectably expressed) are represented for each tissue. U-shaped is highly present in adult ovaries (highlited in yellow).

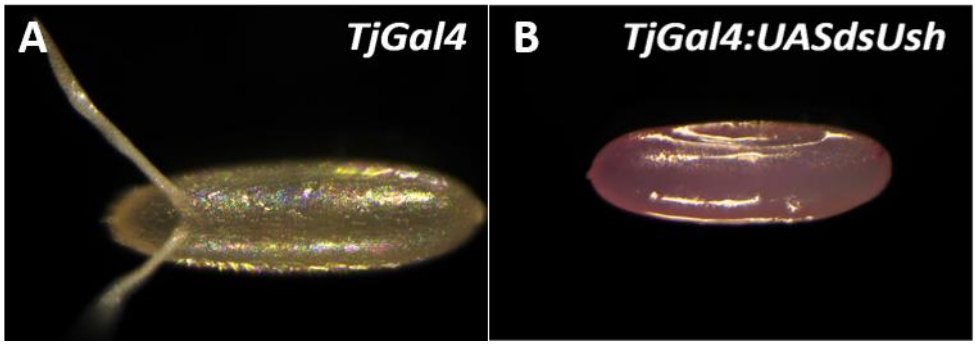
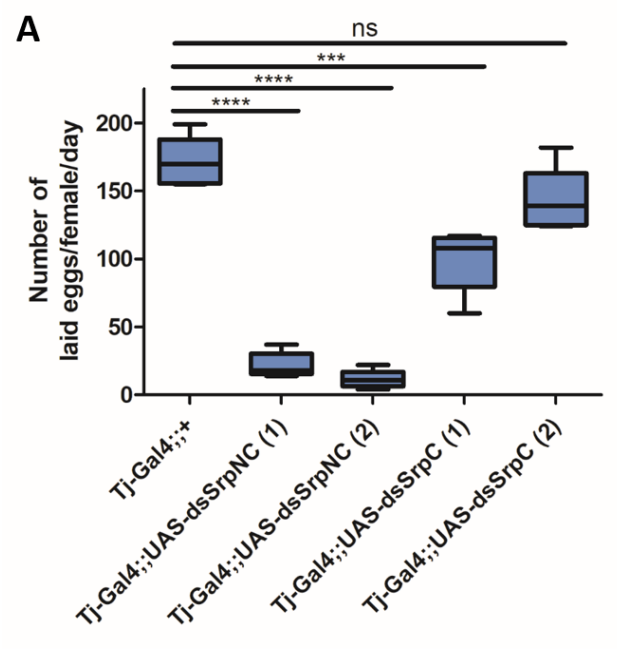


Figure 52. Ush is required for normal egg maturation.

Eggs laid by flies of the genotype *Tj-Gal4;UASdsUsh* (B) have remarkable physical defects in comparison to those of laid by flies of the control genotype *Tj-Gal4* (A)

Drosophila fertility of specific SrpC downregulation in the ovarian follicle cells. As I previously mentioned, the downregulation of SrpC in all tissues (using the *tubulin-Gal4* driver) causes fly lethality (Figure 38). We therefore decided to use the *Tj-Gal4* insertion that downregulates SrpC expression in ovarian follicle cells without interfering with fly viability. Interestingly, flies with downregulation of SrpC under the control of *Tj-Gal4* lay a high number of eggs (Figure 53A) that are able to reach third larval stage (Figure 53B). This contrasts with the results obtained for flies with downregulation of the SrpNC isoform under the control of *Tj-Gal4* driver, as those lay almost no eggs (Figure 53A), which furthermore die during early stages of development (Figure 53B). These results indicate that the isoform SrpC is dispensable for fly female fertility and confirm that SrpNC is required in ovarian follicle cells for adult female fly fertility.

A



B

Genotype	Number of L3 progeny
Tj-Gal4;;+	71
Tj-Gal4;;UAS-dsSrpNC (1)	0
Tj-Gal4;;UAS-dsSrpNC (2)	0
Tj-Gal4;;UAS-dsSrpC (1)	52
Tj-Gal4;;UAS-dsSrpC (2)	130

Figure 53. SrpC is dispensable for fly fertility.

(A) Quantification of the number of eggs laid by flies of the genotypes *Tj-Gal4;+*, *Tj-Gal4;dsSrpNC(1)*, *Tj-Gal4;dsSrpNC(2)*, *Tj-Gal4;dsSrpC(1) and Tj-Gal4;dsSrpC(2)* crossed with fertile males of the genotype *white-eye*. Downregulation of SrpNC, but not SrpC, alters egg formation or release. (B) The number of progeny at third instar larval stage (L1) raised from female flies of the genotypes the genotypes *Tj-Gal4;+*, *Tj-Gal4;dsSrpNC(1)*, *Tj-Gal4;dsSrpNC(2)*, *Tj-Gal4;dsSrpC(1) and Tj-Gal4;dsSrpC(2)*. Flies with downregulation of SrpNC but not SrpC are sterile.

Chapter III. Discussion

Over the years, *Drosophila melanogaster* has emerged as a valuable model to study biological processes implicated in normal development (Gold & Brückner, 2015; Pitsouli & Perrimon, 2008; Saffman & Lasko, 1999; Zaffran, 2003). Genes of the GATA transcription factor family encode zinc finger-containing proteins that are essential for mammalian development (M. Tremblay et al., 2018). Different functions are attributed to each of the two GATA zinc fingers (Martin & Orkin, 1990; Trainor et al., 1996; Yang & Evans, 1992), and mutation in each of them can provoke redundant but also specific developmental defects in many mammalian tissues and at different stages of development (M. H. F. M. Lentjes et al., 2016). During my PhD work, I studied the function of both zinc finger domains of the *Drosophila* GATA factor Serpent (Srp), which is expressed in different organs, where it plays essential functions at specific developmental stages (Rehorn et al., 1996; Reuter, 1994). The particularity of this factor is that it exists as two different isoforms, containing either the canonical N-terminal and C-terminal GATA zinc finger domains (SrpNC), or only the C-zinc finger (SrpC). Both isoforms are simultaneously produced in all *srp*-expressing cells (Waltzer et al., 2002). This characteristic allowed us to investigate *in vivo* the specific roles of each zinc finger domain during fly development.

(A) Using CRISPR/Cas9 to generate flies with Srp isoform loss-of-function

In order to study the functions of SrpNC and SrpC isoforms in *Drosophila*, we used CRISPR/Cas9 technology to generate mutant fly lines devoid of either the SrpNC or SrpC isoform. Using CRISPR/Cas9 has several advantages: it is a simple and efficient method to edit genomic DNA in the fly (Lino et al., 2018), and the modification introduced in the genome can

be easily detected by PCR; CRISPR/Cas9 is the first method that efficiently allows precise control of DNA editing (Lino et al., 2018); it is particularly well adapted in *Drosophila melanogaster*, a model with a profusion of genetic tools, ensuring fast selection and stabilization of the events of interest in the genome; excision of particular genomic region ensures complete elimination of the mRNA of interest; CRISPR/Cas9 has fewer chances to produce off-target effects in comparison to other methods (Boettcher & McManus, 2015). Moreover, there are improved bioinformatics tools (e.g. CRISPR optimal target finder tool) to help identify of the most appropriate guide RNAs targeting the desired genomic region with minimal off-target binding. Coupling this technique with RNA interference is an excellent way to validate the obtained phenotypes, as both methods are unlikely to yield the same off-target effects. Additionally, other approaches can be applied to confirm that the observed phenotypes result from the editing event, and not from off-target modifications, such as increasing the number of independent mutation events and comparing the resulting phenotypes. In our case, screening for mutant flies devoid of SrpNC yielded at least nine different founders positive for the desired mutation. Analysis of these mutants' viability revealed that they are all able to develop until adult stage and that all female flies are sterile, unable to produce offspring. This shows that all the obtained mutants act similarly. In parallel, the screen for mutants with SrpC loss-of-function allowed the recovery of three different founders of the desired mutation. Analysis of their viability showed that they are all unable to reach adulthood. This indicates that mutant lethality is caused by the loss-of-function of SrpC, rather than by another off-target mutation. Finally, for both types of mutation ($srp^{\Delta srpNC}$ and $srp^{\Delta srpC}$ mutants), complementation analyses (performed either with other known *srp* alleles or with a deficiency uncovering the *srp* locus) confirmed that the characterized phenotypes are not due to the genetic background in which our mutants were generated, or to off-target secondary mutations.

(B) Similarities and differences between the mammalian and *Drosophila* GATA zinc finger associated functions

In *Drosophila,* Srp ensures essential functions during several processes and in different organs.

- *In the intestinal system*

In the developing intestine, Srp allows the epithelial-to-mesenchymal transition (EMT) of endodermal cells (K. Campbell et al., 2011; Reuter, 1994). A previous study found that GATA factors in both mammals and *Drosophila* use similar molecular target genes during the induction of the EMT (K. Campbell et al., 2011). Our results show that the GATA C-ZnF domain is sufficient, while the N-ZnF is dispensable for Srp-mediated EMT. Similarly, in vertebrates, it was shown that GATA4 allows EMT of endocardial cells independently of its N-terminal zinc finger-mediated interaction with FOG (Rivera-Feliciano et al., 2006).

- *In the hematopoietic system*

During *Drosophila* hematopoiesis, we found that the C-ZnF is sufficient for proliferation and maintenance of embryonic hemocytes. Similarly, studies on GATA vertebrates showed that GATA1 is required for maintenance of pro-erythroblasts and their differentiation into mature hemocytes during mammalian primitive hematopoiesis (Y. Fujiwara et al., 1996), and that GATA1 C-terminal zinc finger is sufficient to carry out this function (Shimizu et al., 2001). Also, the GATA/FOG interaction is required during mammalian definitive hematopoiesis, in order to regulate specific hematopoietic lineage differentiation. For example, GATA1/FOG1 inhibits differentiation of mast cells progenitors into mast cells, by activating the neutrophil cell fate program in these progenitors (Sugiyama et al., 2008). Likewise, we found that in *Drosophila*, the SrpNC/Ush complex is required for regulation of normal hematopoietic cell

differentiation, mainly by preventing spurious lamellocyte formation in hematopoietic territories.

- *In the ovaries*

During oogenesis, SrpNC/Ush plays essential functions in egg formation and release. This situation can be paralleled in mammals, since GATA4 is required for the expression of essential proteins in ovary follicular cells, which control egg maintenance and release (Kyrönlahti et al., 2011). The role of FOG in the adult mammalian ovary, however, has not been studied yet. Unlike the mammalian GATA4/FOG2 complex, which plays essential functions during development of both ovary and testis (Manuylov et al., 2008; Tevosian et al., 2002), we found that the SrpNC/Ush complex is dispensable for *Drosophila* sex-specific gonad formation.

In summary, our work provides a great deal of new insights concerning the roles and different modes of action of *Drosophila* GATA zinc finger domains in various developmental contexts. Although not all the functions of GATA zinc fingers have been conserved throughout evolution, a striking number of similarities have emerged between mammals and *Drosophila*. Future challenges imply the exploitation of the similarities identified between *Drosophila* and mammalian GATA zinc finger activities, in order to modulate the human GATA zinc finger domains associated with diseases.

(C) Common and different functions for Srp zinc finger domains

Our analysis leads us to divide the functions of Srp into three categories (Figure 53).

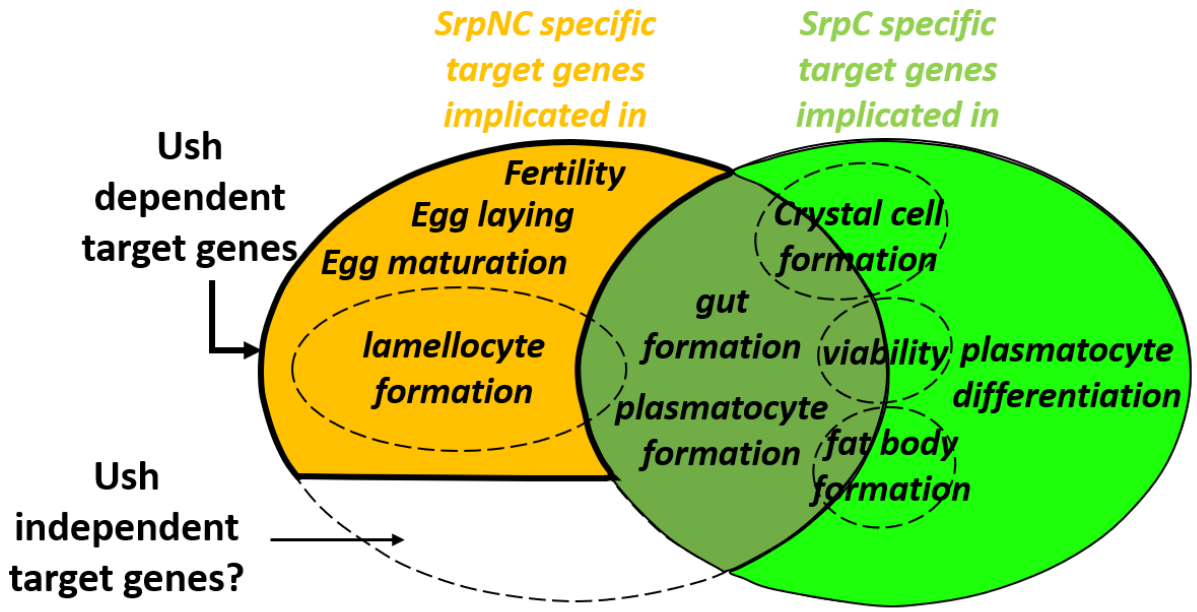


Figure 54. Summary of the processes that depend on Srp function and their division into categories according to the contribution of Srp isoforms, SrpC and SrpNC.

(i) The functions for which the presence of C-terminal zinc finger is sufficient: SrpNC and SrpC are therefore redundant. This is the case for instance during gut and plasmatocyte formation.

(ii) The functions specifically assigned to the SrpC isoform: these are the functions for which SrpNC, despite the fact that it also contains the C-ZnF, cannot compensate for the loss of SrpC. This was observed for fat body and crystal cell formation, as well as for plasmatocyte maturation. These results have several possible explanations. For example, SrpC contains a region encoded by the mutually exclusive exon E4B that can be implicated in the regulation of SrpC specific functions. However, analysis of this sequence by bioinformatics tools did not reveal any known functional or structural motifs. It could be interesting to assess the dispensability of this region by creating a mutant containing both Srp isoforms but without the non-conserved E4B-encoded part of SrpC. 3D structural characterization of GATA zinc finger domains shows that each GATA zinc finger can recognize specific DNA motifs. In fact, both zinc finger domains have a homologous core that interacts with the first three bases of the GATA binding site, while a QTRNRK motif in the C-terminal basic tail of the GATA C-terminal zinc finger allows the recognition of the fourth base pair of the GATA site, adenine/thymine (Bates et al., 2008; Ghirlando & Trainor, 2003). The C-terminal basic tail of the N-terminal zinc finger doesn't have the QTRNRK motif, and it displays more affinity for the GATC, GATG and GATT sites than for the GATA sites (Ghirlando & Trainor, 2003; Pedone et al., 1997). According to this difference in DNA binding site recognition between the two zinc finger domains, the N-ZnF and its adjacent basic region are able to modulate the binding specificity of the C-ZnF, preventing it from recognizing the consensus GATA site (Trainor et al., 2000). Thus, the presence of an isoform devoid of the N-ZnF could be required to recognize specific target genes. A third explanation for the incapability of SrpNC to compensate the loss-of-function of SrpC may depend on the presence of regulatory proteins, such as the FOG cofactor Ush, which interacts with the N-ZnF of SrpNC and interferes with its ability to activate

SrpC target genes. This type of negative regulation has been shown during mast cell progenitor differentiation, where FOG interacts with GATA1 in order to recruit the NuRD repressor complex and inhibit the GATA1-mediated mast cell formation (Cantor et al., 2008; Z. Gao et al., 2010; Gregory et al., 2010). It was also demonstrated by our group that Ush antagonizes SrpNC- but not SrpC-mediated gene transcription in *Drosophila* embryonic mesoderm and in the mammalian COS-7 cell line (Waltzer et al., 2002, 2003).

(iii) Finally, the third category regroups functions that require the presence of both N- and C-terminal zinc fingers, as the downregulation of SrpNC in the follicle cells of *Drosophila* ovaries causes sterility, but the downregulation of SrpC in the presence of SrpNC has no such effect. Additionally, we found that the loss of one copy of SrpNC isoform alters *Drosophila* hematopoietic homeostasis, but the loss of one copy of SrpC in the presence of SrpNC has no effect on this context. However, effects on hematopoietic homeostasis of SrpNC or SrpC downregulation in the lymph gland have to be characterized, in order to consolidate the hypothesis of the existence of SrpNC-dependent specific functions in this context.

To summarize, we envision three modes of action of Srp through the differential use of its two alternatively spliced isoforms: Srp either relies only on its C-terminal zinc finger, or it must be devoid of the N-ZnF, or finally in a third scenario, it would require both zinc finger domains simultaneously in order to properly fulfill its role.

(D) The Srp N-terminal zinc finger domain is dispensable for *Drosophila* development

The most intriguing result from the analysis of *srp* isoform-specific mutants is the ability of the fly to develop until adult stage in the absence of the GATA N-ZnF domain. This domain plays

a role during GATA binding to DNA palindromic sequences and during the interaction of GATA with other proteins (Jason A. Lowry & Mackay, 2006; Trainor et al., 1996; Yang & Evans, 1992). The *in vivo* requirement for this domain has been shown in mammals by analysis of mutant mice harboring in the N-terminal zinc finger the V205G substitution, which abolishes GATA interaction with FOG cofactors. The introduction of this mutation in GATA1 or in both GATA1 and GATA2, as well as in GATA4, causes mice embryonic lethality associated to defects during hematopoiesis or heart formation (J. D. Crispino et al., 2001; Shimizu et al., 2004). Our results show that neither the loss-of-function of SrpNC nor the loss of the interaction of SrpNC with Ush affect the fly development, demonstrating that the essential functions of Srp are not supported by the N-ZnF domain.

(E) Role of Srp N-zinc finger domain during fly development

1) The Srp N-zinc finger is required for regulation of hematopoietic homeostasis

We found that SrpNC and SrpNC/Ush interaction regulates many aspects of larval hematopoiesis. In the lymph gland, mutant analysis showed that crystal cell formation is abolished after loss of SrpNC/Ush interaction, but not in the context of loss-of-function of SrpNC protein, indicating that without a proper interaction with Ush, SrpNC proteins act as negative regulator of crystal cell formation. This suggests that in this specific situation, the cofactor Ush is required to promote crystal cells formation, which contradicts what was already observed in other contexts. That different functions are played by GATA and the GATA/FOG complex has already been seen in the fly, since the induction of larval circulating crystal cells

by overexpression of SrpNC, or of another GATA factor Pnr, is antagonized by Ush (Chatterjee et al., 2019; Nancy Fossett et al., 2003). In addition, we show that during embryonic crystal cell formation, the presence of a SrpNC variant unable to interact with Ush does not block crystal cell formation at the embryonic stage, indicating that SrpNC/Ush interaction is not essential for crystal cell development at this stage. The differences observed between formation of crystal cells in the embryo and in the lymph gland suggest that GATA factors act differently during the two hematopoietic waves.

Interestingly, it was already demonstrated that lymph glands with loss of expression of *ush* expression, display a reduced number of both plasmatocytes and crystal cells (Dragojlovic-Munther & Martinez-Agosto, 2013; H. Gao et al., 2009). Accordingly, we show that the role of Ush during crystal cell formation depends on its interaction with SrpNC. However, plasmatocyte formation is not affected in $srpNC^{V735G}$ mutant, suggesting that the SrpNC/Ush interaction is not required in this process, and that Ush-mediated plasmatocyte formation is likely to depend on its interaction with another GATA factor expressed in the lymph gland.

Apart from crystal cells and plasmatocytes, the fly produces a third type of hemocytes called lamellocytes, in response to various stress situations (Letourneau et al., 2016). Interestingly, we found that both SrpNC and SrpNC/Ush interaction inhibits lamellocyte formation under normal conditions. Analysis of lamellocyte formation in larval hemolymph showed that the phenotype is more severe in larvae with loss of SrpNC/Ush interaction, than in those with SrpNC loss-of-function. In contrast, $srp^{\Delta srpNC}$ mutants display, a lamellocyte production, whose level is more important in the lymph gland than in the larval hemolymph. This discrepancy might reflect the existence of compartment-specific hematopoietic regulatory programs for lamellocyte induction in the lymph gland or in circulation.

2) The Srp N-terminal zinc finger is required for egg maturation

We found that SrpNC and the SrpNC/Ush complex are required for egg dorsal respiratory appendages, micropyle and eggshell formation. The lack of dorsal appendages can be the consequence of an alteration of appendage precursor cell determination, or of subsequent appendage morphogenesis events. To discriminate between these two possibilities, staining of mutant ovaries with Broad-Complex (BR-C) antibody that mark the appendage progenitors would certainly be useful. Additionally, the formation of functional micropyle requires the migration of specialized ovarian follicle cells, called border cells and centripetal cells (Montell et al., 1992; Suzanne et al., 2001; Twombly et al., 1996). Hence, analysis of the border cell and centripetal cell formation in $srp^{\Delta srpNC}$ and $srpNC^{V735G}$ mutant ovaries, would help determine if the loss of micropyle is due to a hampered determination or migration of these specialized cells, or due to deregulation of genes implicated in other steps of micropyle morphogenesis. Finally, synthesis of eggshell components requires the expression of different vitelline membrane and chorion genes, which have to be finely regulated both spatially and temporally (Cavaliere et al., 2008). Identifying SrpNC and SrpNC/Ush transcriptional targets participating in eggshell secretion would therefore be of great interest. As we didn't find any defects in $srp^{\Delta srpNC}$ and $srpNC^{V735G}$ mutant ovaries morphogenesis, it would be imaginable to carry on transcriptomic analysis by RNA sequencing and compare gene expression in wild type, $srp^{\Delta srpNC}$ mutant and $srpNC^{V735G}$ mutant ovaries, in order to identify the cause of sterility at the transcriptional level.

(F) The functions of the Srp N-terminal zinc finger depend mostly on its interaction with the FOG factor U-shaped

Although dispensable for *Drosophila* development, we found that the Srp N-ZnF is required for maintenance of hematopoietic homeostasis and for female fertility. The N-ZnF plays a role during GATA binding to DNA palindromic sequences, as well as during the interaction of GATA with other proteins, such as FOG proteins, the Krüppel-like family factor KLF13, and the LIM domain containing protein 2 (LMO2) that bridges GATA to transcription factors of the bHLH family (Lavallée et al., 2006; Jason A. Lowry & Mackay, 2006; Trainor et al., 1996; Yang & Evans, 1992). The most studied regulator of GATA function is FOG (Chlon & Crispino, 2012). During GATA/FOG interaction, the only valine present in the N-ZnF domain is required for the interaction with FOG, and its substitution into glycine alters GATA/FOG association without affecting GATA binding to DNA, or to KLF13 and LMO2 (J. D. Crispino et al., 1999; Lavallée et al., 2006; Wilkinson-White et al., 2011). Accordingly, only SrpNC has the ability to interact with Ush (Waltzer et al., 2002). Importantly, the N-ZnF amino acid sequence of SrpNC is highly conserved, with the equivalent mammalian sequence including the valine residue (Figure 7) (Waltzer et al., 2002), and the valine to glycine substitution of SrpNC alters the SrpNC/Ush interaction (Nancy Fossett et al., 2003). Interestingly, during this work we showed that mutants with the V735G substitution have the same defects as those observed in mutants with specific SrpNC loss-of-function, indicating that the interaction of SrpNC with Ush is required for *Drosophila* hematopoietic homeostasis and female fertility (Figure 54). Thus, the functions played here by the N-ZnF domain depend on its ability to interact with its FOG cofactor, rather than on its binding to specific DNA sequences, or its interaction with the other aforementioned regulators.

The role of Ush in hematopoietic homeostasis regulation has been already described (Avet-Rochex et al., 2010; Dragojlovic-Munther & Martinez-Agosto, 2013; H. Gao et al., 2009; R. P. Sorrentino et al., 2007), and we determine here that most of its role in this context is due to its

interaction with SrpNC. Also, during this work, a role for Ush during the *Drosophila* oogenesis was detected for the first time. The role of SrpNC/Ush interaction in oogenesis could be confirmed by restoring fly fertility in the $srpNC^{V735G}$ mutant, by expressing in srp^{V735G} ovaries a mutant form of Ush protein harboring the compensatory substitution serine to arginine, identified in the mammalian FOG zinc finger domain as responsible for restoration of the interaction of FOG with vertebrate GATA proteins that contain the substitution valine to glycine mimicked in $SrpNC^{V735G}$ (J. D. Crispino et al., 1999). Of note, among the five *Drosophila* GATA factors, only Srp and GATAd are expressed in *Drosophila* ovaries (Lepesant et al., 2020). GATAd being devoid of the zinc finger domains that allows the interaction GATA/FOG (Okumura et al., 2005), the expression of a form of Ush with this compensatory mutation in the ovaries should not interfere with the binding of Ush to GATA proteins other than SrpNC in this organ.

However, although all the identified SrpNC mediated functions are dependent of the interaction with Ush, we cannot exclude the possibility of presence of some no identified functions for SrpNC that are independent of Ush (Figure 54).

(G) Why has nature selected a Srp isoform with only one zinc finger domain?

The ancestor of GATA transcription factors of triploblastic animals is the diploblastic cnidarians that have a single GATA factor, which like SrpNC, contains two conserved zinc finger domains. During evolution, some GATA genes in the fly have lost DNA sequence coding for the N-zinc finger, giving rise to factors with only one zinc finger, such as SrpC (W. J. Gillis et al., 2007; W. Q. Gillis et al., 2008). This type of GATA factor with only one zinc finger domain is also found in other triploblastic protostomes, such as *C.elegans*, which have eleven

GATA factors, yet only one of them (elt-1) containing double zinc finger domains (all the others contain only the C-ZnF) (W. Q. Gillis et al., 2008). In the fly, we found that the N-ZnF domain is dispensable for development. Similarly, in *C.elegans*, a high number of GATA-mediated functions are endorsed by the single zinc finger GATA factors (Block & Shapira, 2015), and analysis of elt-1 mode of action revealed that deletion of elt-1 N-ZnF domain has no effect on the elt-1-dependant gene activation (Shim et al., 1995). It thus appears that dispensability of GATA N-ZnF during development is a common feature of invertebrates. This striking difference with mammals may reflect the different biological complexity of these organisms, as it was shown that in Human the estimated number of protein/protein interactions is significantly higher than in organisms of lesser complexity, such as *Drosophila* and *Caenorhabditis* (M. P. H. Stumpf et al., 2008). Thus, mammals might have a bigger need for protein/protein interactions mediated by the GATA N-ZnF, in order to control their development.

In addition, we found that SrpC has essential functions during fly development. Accordingly, under normal conditions, flies produce about 5 times more SrpC than SrpNC isoform (Waltzer et al., 2002). Single zing finger containing versions exhibit a smaller surface available for interactions with other factors than a protein with two zinc finger domains. The ability of SrpC to only mildly interact with other proteins could make it more rapid and efficient in DNA-binding, prior to transcription activation of target genes. Conversely, SrpNC could interact with other proteins that interfere with, or reduce the speed of, its activity during transcriptional regulation.

Finally, although dispensable for fly development, we found that the N-ZnF domain plays essential functions during the hematopoietic homeostasis and fly oogenesis (discussed above). This is hypothesized to be why the fly gene has kept production of this isoform with two zinc

finger domains in *srp* expressing cells (Waltzer et al., 2002). Processes regulated by the N-zinc finger domain are highly dynamic and renewable, which suggests that these processes require more complex regulation programs during their realization than in other stable developmental processes, such as gut and fat body formation, for which the presence of the C-zinc finger domain is sufficient.

(H) Subfunctionnalization in Serpent isoforms

Gene duplication and alternative splicing are dominant factors in the evolution of eukaryote complexity and diversity. After gene duplication, two major fates can be adopted by the generated gene copies, namely neofunctionalisation (acquisition of novel functional properties in comparison to the ancestral gene) and subfunctionalisation (performing of complementary functions that jointly match that of the ancestral gene) (Sandve et al., 2017). We found that the majority of the functions played by Srp are portioned between the isoforms SrpC and SrpNC, suggesting that a subfunctionalization process occurs with the *srp* gene. However, we also found that both isoforms still have redundant functions during embryonic gut and plasmatocyte formation, suggesting that degeneration and complementation occuring before subfunctionalisation (Force et al., 1999) are not completely achieved.

Individual gene duplication is common in all organisms (Taylor & Raes, 2004). In the fly, *srp* gene was obtained after duplication of its ancestral gene, which gave rise to *srp* and to other gene copies, including the GATA factor Pnr-encoding gene (W. Q. Gillis et al., 2008), which has a distinct expression pattern and produces only canonic GATA factor isoforms with two zinc finger domains (Fromental-Ramain et al., 2008; Heitzler et al., 1996; Ramain et al., 1993).

Intriguingly, production of SrpC factor, relies on alternative splicing of the *srp* transcript, rather than on gene duplication and subsequent degenaration of the N-ZnF coding sequence in one of the two duplicates. Production of both Srp isoforms by alternative splicing keeps the expression of both *srp* isoforms under the control of the same regulatory region. It is unknown, however, why exactly the fly has chosen the subfunctionnalization process by alternative splicing instead of gene duplication. One possibility is that in some territories, during fly development, the simultaneous presence of both Srp isoforms is splicing required, therefore exerting an evolutionary pressure on this alternative splicing event. Thus, in the end, this work illustrates that, like genome duplication in vertebrates, alternative splicing provides an efficient strategy to promote subfunctionalization and generate GATA functional diversity in invertebrates. Identification and analysis of more genes sharing this particularity would be of high interest for a better understanding of the complex functional diversity in the fly and other organisms.

Chapter IV. Materials and Methods

1- Generation of $srp^{\Delta srpNC}$ and $srp^{\Delta srpC}$ mutant fly lines

The $srp^{\Delta srpNC}$ and $srp^{\Delta srpC}$ mutant fly lines were created by generating deletions in the *srp* locus using the CRISPR/Cas9 genome editing system, with tools and reagents published in (Port et al., 2014). For each fly line, two different single guide RNAs (sgRNAs) were used, which determine the limits of the genomic DNA region to be removed. The sgRNAs were designed using the "CRISPR optimal target finder tool", and sgRNAs that recognize zero off-target sites were selected. sgRNAs sense sequences were as follows:

- $srp^{\Delta srpNC}$: 5'-ATAAGTCATAAGGTTTTGCT and AATTAAACAGCCTAGAAGAT
- $srp^{\Delta srpC}$: 5'-CACTTTTCGATTAAACTAGT and CGTAGTAAGGCTAACACGAG

For each sgRNA, sense and anti-sense oligonucleotides were annealed and inserted into the pCFD3 plasmid (addgene #49410) following the specific protocol in the CRISPR fly design website (https://www.crisprflydesign.org/). After validation of cloning efficiency by sequencing, recombinant plasmids were purified by the Qiagen plasmid maxiprep purification kit, checked by sequencing and injected into embryos expressing the nuclease Cas9 in the germline cells under the control of the vasa promoter (Bloomington stock 51323). The injection mix contains a specific pair of sgRNA recombinant plasmids at a concentration of 250 ng/µl each. For injections, embryos at early stages of embryogenesis were covered by Voltalef 10S oil (VWR Chemicals, # 24627.188) in order to make the embryo clearly visible under the chorion. The mix was injected at the embryonic posterior pole, using the Leica mechanical micromanipulator and a needle pulled from a capillary glass (Harvard Apparatus, #30-0019) by a Sutter instruments brand horizontal puller. Injected embryos were allowed to develop until adulthood. These F0 adult flies were crossed in order to yield F1 progeny. F1 males were subsequently crossed for 4-5 days, then removed from the cross and their DNA extracted as described in (Gloor et al., 1993).

The screening PCR was performed using the following primers:

- $srp^{\Delta srpNC}$F: 5'-AATTCAAACCAACAAAAAGACACCT

- $srp^{\Delta srpNC}$R: 5'- GCAGATGCAGATTGATCAGATTTTC

- $srp^{\Delta srpC}$F: 5'- CCCCTGAAAAGCTCCAATGC

- $srp^{\Delta srpC}$R: 5'- CTCAGTGGCCAAGGAGGTTT

These primers bind regions located outside of the expected deletions. Flies positive for the deletion will give an amplicon of size smaller than that obtained from the control chromosome. Upon detection, presence of the deletion was validated by Sanger sequencing.

2- Generation of *srp^{V735G}* mutant fly line

The *srp^{V735G}* mutant fly line was created by introducing a missense mutation (a GGA codon replacing the wild-type GTC codon) in *srp*, using the CRISPR/Cas9 genome editing system. A sgRNA was used to target the nuclease Cas9 to a 5'-ATGCCGATTATTTCACTGAG sequence in the vicinity of the nucleotides intended to be modified. A single-strand DNA donor (ssODN) of 200 base pairs harboring the mutated nucleotides was used as a repair template. The recombinant sgRNA-containing pCFD3 plasmid was prepared as described above. The ssODN was manufactured by Integrated DNA Technologies. The injection mix constituted of 250 ng/µl of the recombinant plasmid and 100 ng/µl of the ssDNA. ssODN sequence: CGTTCATTTT CATGTACAAG CCGCAGGCAT TGCACAAATA GTGTCCCGTG TTATCGCGTC GCCATAATGG GGTTGAAATC GCACCACAGT TTCCACACTC ACGTCCTTCT GTAAAATAAT CGGCATCGAA TAATGCTTCA AAGATTGACA TAGAAATAT CGGAATAAGT CATAAGGTTT TGCTAGGTGT TTGTTTGATT.

The screening PCR of F1 males' genomic DNA involved the following primers:

- srp^{V735G}F: 5'-ACACAATACGCAAATGC,

- srp^{V735G}R: 5'-ACTGAGAGAGGATGTTGC

- $srp^{V735G}R_m$: 5'-CAGTTTCCACACTCACG.

F and R primers bind outside of the modified DNA region, while R_m is a reverse primer that specifically recognizes the modified nucleotides incorporated from the ssODN-directed repair event. Note that some silent mutations have been inserted in the ssODN (Figure 5), in order to allow the design of a R_m primer efficient at discriminating the modified DNA from the unmodified DNA. Upon detection, events of interest were confirmed by sequencing.

3- Generating transgenic RNAi fly lines specific to *srp* isoforms

RNAi constructs were designed using the E-RNAi web service (https://www.dkfz.de/signaling/e-rnai3//). The 21-nucleotide sequences correspond to specific regions of *srp* gene exon 4A or 4B, in order to specifically interfere with expression of *srpNC* and *srpC*, respectively. Sequences with the lowest off-target score were selected and blasted using the National Center for Biotechnology and Information (NCBI) website against the *D. melanogaster* RNA-sequences available at the NCBI Reference RNA Sequences (Refseq_rna) database, in order to validate the absence of matches with off-target sites. Short-hairpin RNA (shRNA) have been designed as described as described in (Jian-Quan Ni, et al, 2011) and synthetized by the Integrated DNA Technologies (IDT) platform.

- Exon 4A specific shRNA (1) sense strand: CTAGCAGTCGATAACACGGGACACTATTT TAGTTATATTCAAGCATAAAATAGTGTCCCGTGTTATCGGCG

- Exon 4A specific shRNA (1) anti-sense strand: AATTCGCCGATAACACGGGACACTATTT TATGCTTGAATATAACTAAAATAGTGTCCCGTGTTATCGACTG

- Exon 4A specific shRNA (2) sense strand: CTAGCAGTGCATGAATCGACCCCTAATTA TAGTTATATTCAAGCATATAATTAGGGGTCGATTCATGCGCG

- Exon 4A specific shRNA (2) anti-sense strand: AATTCGCGCATGAATCGACCCCTAATTA TATGCTTGAATATAACTATAATTAGGGGTCGATTCATGCACTG

- Exon 4B specific shRNA (1) sense strand: CTAGCAGTCGCTGAATCAGGCGGGGATTT TAGTTATATTCAAGCATAAAATCCCCGCCTGATTCAGCGGCG

- Exon 4B specific shRNA (1) anti-sense strand: AATTCGCCGCTGAATCAGGCGGGGATTT TATGCTTGAATATAACTAAAATCCCCGCCTGATTCAGCGACTG

- Exon 4B specific shRNA (2) sense strand: CTAGCAGTCTATAAACCCAACTCATTTAA TAGTTATATTCAAGCATATTAAATGAGTTGGGTTTATAGGCG

- Exon 4B specific shRNA (2) anti-sense strand: AATTCGCCTATAAACCCAACTCATTTAA TATGCTTGAATATAACTATTAAATGAGTTGGGTTTATAGACTG

For each shRNA construct, the synthetized sense and anti-sense strands were annealed and cloned into pWalium20 plasmid (DGRC: 1472) following a protocol adapted from the cloning in pCFD3 vector protocol (https://www.crisprflydesign.org/). The recombinant plasmids were individually injected into flies containing AttP2 sites that allow insertion of a plasmid into the AttP2 platform in genomic DNA.

4- Fly strains

The fly strains *srp*^{3}, *srp*^{As}, *Df(3R)BSC728* were obtained from the Bloomington Drosophila Stock Center. The fly strains *lz-Gal4* and *srp*^{6G} were already available in our group (Waltzer et al., 2002, 2003). The fly strains *ush*^{VX22}, *ush*^{Rev24} were kindly provided by P.Heitzler's team. The fly strain *msnF9-mCherry* was kindly provided by R. Schulz's group.

5- Embryonic RNA extraction and RT-PCR

Embryos of the genotypes w^{1118}, $srp^{\Delta srpNC}$, $srp^{\Delta srpC}$ were allowed to develop on agar plates, at 25°C until stages 14 to 16 of embryogenesis. Total embryonic RNA was extracted using the RNeasy Plus Mini kit (Qiagen). Reverse transcription was performed using random primers (Invitrogen, *P/N* 58875), and PCR was performed using the following primer pairs: SrpNC-F: 5'-GATACCTGGTTCGATCC / SrpNC-R: 5'-TGGTGTCCTTTTTCATG, and SrpC-F: 5'-CTCGGCATCGTTGTC / SrpC-R: 5'-TCCGGCTCGCTTTGAGG.

6- Survival analysis

Throughout the survival analysis period, flies were raised at 25°C. For each analyzed genotype, embryos were collected on agar plates at stages 14 to 16 of embryogenesis, and their ability to hatch was recorded. 48 hours later, the ability of the hatched first instar larvae to reach the third instar larval stage was quantified, and the third instar larvae L3 were transferred to tubes containing standard media, where analysis of their ability to develop into pupal and adult flies was recorded 48 hours and 5 days after transfer.

For the table in figure 28, crosses were performed in tubes containing standard media, and the progeny was allowed to develop at 25°C until the third instar larval or adult stage.

7- Bioinformatic tools

Multiple alignment of the SrpNC amino acid sequence with the mammalian GATA protein sequences was done by using the bioinformatics software Geneious. SrpNC (NP_732098.1),

GATA1 (AAH09797.1), GATA2 (AAH51272.1), GATA3 (NP_001002295.1), GATA4 (AAI43480.1), GATA5 (AAH47790.1) and GATA6 (NP_005248.2).

Multiple alignment of the SrpC amino acid sequence with GATA protein sequences of other species was done by using the combine popular aligners M-Coffee of the multiple sequence alignment server T-coffee. Dmel-SrpC (NP_732100.2), Lcup-GATA (XP_023301929), Amel-GATA (XP_016769881.2), Tcas-GATA (XP_008200496.1), Mdom- GATA (XP_005187504.3), Hsap-GATA2 (AAH51272.1), Hsap-GATA3 (NP_001002295.1) and Hsap-GATA4 (AAI43480.1).

Analysis of the presence of known functional domains in the amino acid sequence encoded by the exon E4B was done by the bioinformatics tools InterPro, ExPASY and NCBI Conserved Domain Search.

8- RNA in situ hybridization

Anti-sense probes specific for *alcohol dehydrogenase*, *croquemort*, *GATAe*, *glutactin*, *grain*, *lozenge*, *Prophenoloxidase 2*, *Peroxidasin*, *tiggrin* and *viking* transcripts were synthesized after linearization of plasmids containing the cDNA sequences of each of the mentioned target (Drosophila Golden Collection website). Probe transcription was performed for 2 hours at 37°C in the presence of the appropriate T7, T3 or SP6 RNA polymerase (Promega), RNasin ribonuclease inhibitors (Promega) and a mixture of dATP, dCTP, dGTP and dioxigenin (Dig)-coupled dUTP (Roche). The synthetized probe was precipitated in the presence of EDTA PH8, LiCl, torula RNA and ethanol and conserved at -20°C in Hybridization Buffer (HB; 50ml of HB: 25 ml of deionized formamide, 10 ml of 10% Roche blocking reagent, 5 ml of saline-sodium citrate buffer 20x (x), 1 ml of torula RNA (50 mg/ml), 0,5 ml CHAPS (10%), 0,5 ml EDTA (0,5M), 50 µl heparin (0,05 g/ml) and 50 µl of 100% Tween^{TM} 20).

Embryos were collected on agar plate at 25°C. They were dechorionated by soaking the eggs in a 50% bleach solution for 5 minutes at room temperature, rinsed under tap-water and then fixed in a fixation solution (1.8 ml of 37% formaldehyde solution, 200 µl of EGTA 0.5M and 2 ml of heptane). Fixation was performed by shaking embryos on a rotomax at room temperature for 20 minutes. Fixed embryos were washed several times with 100% methanol, then with 1% phosphate buffered saline containing 0.1% of Tween™ 20 (PBSTw), and incubated for 1h at 65°C in HB. Probe hybridization was performed overnight at 65°C in HB containing probe at the appropriate dilution. The following day, embryos were incubated for one hour at 65°C in HB solution and then for 20 minutes at the same temperature, after adding equivalent volumes of HB solution and PBSTw to the embryos. Samples were then washed at room temperature in PBSTw several times and in PBSTw containing 1% of bovine serum albumin (PBSTw-BSA) for 30 minutes. Embryos were then incubated overnight at 4°C in PBSTw-BSA containing α-DIG antibodies conjugated with alkaline phosphatase (Roche). On the third day, embryos were washed 5 times with PBSTw for 10 minutes, and later once for 10 minutes with the Staining Buffer solution (SB) containing MgCl2, NaCl, Tris-HCl PH 9.5 and Tween ™ 20. Finally, the coloration solution formed by the NBT/BCIP substrate diluted in the SB solution was added to the embryos in order to reveal the anti-DIG binding sites on the embryonic tissues. Images were acquired by the Nikon Eclipse 80i microscope (NIS-Element software).

9- Lymph gland and circulating hemocytes immunostaining

Larvae were grown on standard medium at 25°C. For lymph gland analysis, we selected mid-third larval instar larvae, and for circulating hemocytes analysis we selected wandering larvae prior to spiracle eversion. All the analyzed larvae having the X-linked transgenes *msnF9-mCherry* are heterozygous females for the transgene. Samples were fixed with 4%

paraformaldehyde for 20 minutes at room temperature. They were washed several times with 1X phosphate buffered saline (PBS) and then with 1X PBS solution containing 0.3% Triton X-100 (PBSTr). All samples were then blocked for 30 minutes in PBSTr containing 1% of bovine serum albumin (PBSTr-BSA). Incubation lasted one night (α-P1, α-L1, Phalloidin) or 36 hours (α-Collier, α-PPO1), in PBSTr-BSA. Antibodies used were: Mouse α-P1 (kindly provided by I. Ando, 1/30), Mouse α-Col (kindly provided by M. Crozatier, 1/50) and Rabbit α- PPO1 ((Li et al., 2012), 1/10 000). Samples were washed several times with PBSTr-BSA and incubated with Alexa Fluor-labeled secondary antibodies (Molecular Probes, 1/400). DAPI (Sigma-Aldrich) was adjusted to 1 mg/L in blocking reagent and added either to the primary antibodies solution for lymph gland staining, or incubated alone for 20 minutes at room temperature for circulating hemocyte staining. Samples were mounted in Vectashield medium (Vector laboratories). Slides were then analyzed using the Leica SP8 inverted-based confocal microscope. The optimized system option was chosen to collect Z-series of images by the Leica LAS-X Life Science software. Maximum intensity projection of Z-stack of each image was created by the image processing software ImageJ. For each experiment, images of different genotypes were taken with the same signal intensity.

10- Quantification of lamellocyte phenotype penetrance

Larvae were raised on standard medium at 25°C. Live larvae were observed on a Leica fluorescence detecting macroscope and were separated into 4 categories, depending on the number of *msnF9-mCherry* expressing circulating cells detected in each larva. Activation of reporter expression in the hemolymph is restricted to lamellocytes (T. Tokusumi et al., 2009). Larvae with no positive cells in their hemolymph were placed in the “0 lamellocytes” category. Larvae with 1-5 cells expressing the reporter in their hemolymph were placed in the “<5

lamellocytes category". Larvae with more than 5 cherry-positive circulating cells were placed in the ">5 lamellocytes category". Finally, larvae with large clusters of mCherry fluorescent cells were placed in the category of larvae with "lamellocyte aggregates". Only the percentage of larvae with extreme phenotypes (0 lamellocytes and aggregates of lamellocytes) were taken into consideration during the analysis of the ability of different genotypes to produce lamellocytes.

11- Analysis of the fertility phenotype

Virgin female w^{1118}, $srp^{\Delta srpNC}$, srp^{V735G} flies were mated with w^{1118} males and allowed to lay eggs on agar plates, at 25°C. The number of laid eggs was quantified 2 and 3 days after the beginning of the crosses. Photos of the eggs laid were aquired with a Leica macroscope.

Chapter V. References

Abel, T., Michelson, A. M., & Maniatis, T. (1993). A Drosophila GATA family member that binds to Adh regulatory sequences is expressed in the developing fat body. *Development*, *119*(3), 623-633.

Allen, H. L., Flanagan, S. E., Shaw-Smith, C., De Franco, E., Akerman, I., Caswell, R., International Pancreatic Agenesis Consortium, Ferrer, J., Hattersley, A. T., & Ellard, S. (2011). GATA6 haploinsufficiency causes pancreatic agenesis in humans. *Nature Genetics*, *44*(1), 20-22. https://doi.org/10.1038/ng.1035

Arceci, R. J., King, A. A., Simon, M. C., Orkin, S. H., & Wilson, D. B. (1993). Mouse GATA-4 : A retinoic acid-inducible GATA-binding transcription factor expressed in endodermally derived tissues and heart. *Molecular and Cellular Biology*, *13*(4), 2235-2246. https://doi.org/10.1128/mcb.13.4.2235

Aronson, B. E., Stapleton, K. A., & Krasinski, S. D. (2014). Role of GATA factors in development, differentiation, and homeostasis of the small intestinal epithelium. *American Journal of Physiology - Gastrointestinal and Liver Physiology*, *306*(6), G474-G490. https://doi.org/10.1152/ajpgi.00119.2013

Asselin-Labat, M.-L., Sutherland, K. D., Barker, H., Thomas, R., Shackleton, M., Forrest, N. C., Hartley, L., Robb, L., Grosveld, F. G., van der Wees, J., Lindeman, G. J., & Visvader, J. E. (2007). Gata-3 is an essential regulator of mammary-gland morphogenesis and luminal-cell differentiation. *Nature Cell Biology*, *9*(2), 201-209. https://doi.org/10.1038/ncb1530

Avet-Rochex, A., Boyer, K., Polesello, C., Gobert, V., Osman, D., Roch, F., Augé, B., Zanet, J., Haenlin, M., & Waltzer, L. (2010). An in vivo RNA interference screen identifies gene networks controlling Drosophila melanogaster blood cell homeostasis. *BMC Developmental Biology*, *10*, 65. https://doi.org/10.1186/1471-213X-10-65

Balduini, C. L., Pecci, A., Loffredo, G., Izzo, P., Noris, P., Grosso, M., Bergamaschi, G., Rosti, V., Magrini, U., Ceresa, I. F., Conti, V., Poggi, V., & Savoia, A. (2004). Effects of the R216Q mutation of GATA-1 on erythropoiesis and megakaryocytopoiesis. *Thrombosis and Haemostasis*, *91*(1), 129-140. https://doi.org/10.1160/TH03-05-0290

Banerjee, U., Girard, J. R., Goins, L. M., & Spratford, C. M. (2019). Drosophila as a Genetic Model for Hematopoiesis. *Genetics*, *211*(2), 367-417. https://doi.org/10.1534/genetics.118.300223

Barnett, T., Pachl, C., Gergen, J. P., & Wensink, P. C. (1980). The isolation and characterization of Drosophila yolk protein genes. *Cell*, *21*(3), 729-738. https://doi.org/10.1016/0092-8674(80)90436-5

Bates, D. L., Chen, Y., Kim, G., Guo, L., & Chen, L. (2008). Crystal structures of multiple GATA zinc fingers bound to DNA reveal new insights into DNA recognition and self-association by GATA. *Journal of molecular biology*, *381*(5), 1292-1306. https://doi.org/10.1016/j.jmb.2008.06.072

Battle, M. A., Bondow, B. J., Iverson, M. A., Adams, S. J., Jandacek, R. J., Tso, P., & Duncan, S. A. (2008). GATA4 is essential for jejunal function in mice. *Gastroenterology*, *135*(5), 1676-1686.e1. https://doi.org/10.1053/j.gastro.2008.07.074

Belaguli, N. S., Sepulveda, J. L., Nigam, V., Charron, F., Nemer, M., & Schwartz, R. J. (2000). Cardiac Tissue Enriched Factors Serum Response Factor and GATA-4 Are Mutual Coregulators. *Molecular and Cellular Biology*, *20*(20), 7550-7558.

Benmimoun, B., Haenlin, M., & Waltzer, L. (2015). Blood cell progenitor maintenance : Collier barks out of the niche. *Fly*, *9*(4), 160-164. https://doi.org/10.1080/19336934.2016.1151130

Bennett, J., Wu, Y.-G., Gossen, J., Zhou, P., & Stocco, C. (2012). Loss of GATA-6 and GATA-4 in Granulosa Cells Blocks Folliculogenesis, Ovulation, and Follicle Stimulating Hormone Receptor Expression Leading to Female Infertility. *Endocrinology*, *153*(5), 2474-2485. https://doi.org/10.1210/en.2011-1969

Beuling, E., Aronson, B. E., Tran, L. M. D., Stapleton, K. A., ter Horst, E. N., Vissers, L. A. T. M., Verzi, M. P., & Krasinski, S. D. (2012). GATA6 Is Required for Proliferation, Migration, Secretory Cell Maturation, and Gene Expression in the Mature Mouse Colon. *Molecular and Cellular Biology*, *32*(17), 3392-3402. https://doi.org/10.1128/MCB.00070-12

Beuling, E., Baffour-Awuah, N. Y. A., Stapleton, K. A., Aronson, B. E., Noah, T. K., Shroyer, N. F., Duncan, S. A., Fleet, J. C., & Krasinski, S. D. (2011). GATA factors regulate proliferation, differentiation, and gene expression in small intestine of mature mice. *Gastroenterology*, *140*(4), 1219-1229.e1-2. https://doi.org/10.1053/j.gastro.2011.01.033

Bier, E. (2005). Drosophila , the golden bug, emerges as a tool for human genetics. *Nature Reviews Genetics*, *6*(1), 9-23. https://doi.org/10.1038/nrg1503

Blobel, G. A., Nakajima, T., Eckner, R., Montminy, M., & Orkin, S. H. (1998). CREB-binding protein cooperates with transcription factor GATA-1 and is required for erythroid differentiation. *Proceedings of the National Academy of Sciences of the United States of America*, *95*(5), 2061-2066. https://doi.org/10.1073/pnas.95.5.2061

Block, D. H., & Shapira, M. (2015). GATA transcription factors as tissue-specific master regulators for induced responses. *Worm*, *4*(4). https://doi.org/10.1080/21624054.2015.1118607

Boettcher, M., & McManus, M. T. (2015). Choosing the Right Tool for the Job : RNAi, TALEN or CRISPR. *Molecular cell*, *58*(4), 575-585. https://doi.org/10.1016/j.molcel.2015.04.028

Bonachea, E. M., Chang, S.-W., Zender, G., LaHaye, S., Fitzgerald-Butt, S., McBride, K. L., & Garg, V. (2014). Rare GATA5 sequence variants identified in individuals with bicuspid aortic valve. *Pediatric Research*, *76*(2), 211-216. https://doi.org/10.1038/pr.2014.67

Bossard, P., & Zaret, K. S. (1998). GATA transcription factors as potentiators of gut endoderm differentiation. *Development (Cambridge, England)*, *125*(24), 4909-4917.

Bosse, T., Piaseckyj, C. M., Burghard, E., Fialkovich, J. J., Rajagopal, S., Pu, W. T., & Krasinski, S. D. (2006). Gata4 Is Essential for the Maintenance of Jejunal-Ileal Identities in the Adult Mouse Small Intestine. *Molecular and Cellular Biology*, *26*(23), 9060-9070. https://doi.org/10.1128/MCB.00124-06

Brade, T., Pane, L. S., Moretti, A., Chien, K. R., & Laugwitz, K.-L. (2013). Embryonic Heart Progenitors and Cardiogenesis. *Cold Spring Harbor Perspectives in Medicine*, *3*(10). https://doi.org/10.1101/cshperspect.a013847

Bresnick, E. H., Lee, H.-Y., Fujiwara, T., Johnson, K. D., & Keles, S. (2010). GATA switches as developmental drivers. *The Journal of Biological Chemistry*, *285*(41), 31087-31093. https://doi.org/10.1074/jbc.R110.159079

Brookman, J. J., Toosy, A. T., Shashidhara, L. S., & White, R. A. (1992). The 412 retrotransposon and the development of gonadal mesoderm in Drosophila. *Development (Cambridge, England)*, *116*(4), 1185-1192.

Brown, S., & Castelli-Gair Hombría, J. (2000). Drosophila grain encodes a GATA transcription factor required for cell rearrangement during morphogenesis. *Development (Cambridge, England)*, *127*(22), 4867-4876.

Bruneau, B. G. (2013). Signaling and transcriptional networks in heart development and regeneration. *Cold Spring Harbor Perspectives in Biology*, *5*(3), a008292. https://doi.org/10.1101/cshperspect.a008292

Buchon, N., Osman, D., David, F. P. A., Fang, H. Y., Boquete, J.-P., Deplancke, B., & Lemaitre, B. (2013). Morphological and molecular characterization of adult midgut compartmentalization in Drosophila. *Cell Reports*, *3*(5), 1725-1738. https://doi.org/10.1016/j.celrep.2013.04.001

Campbell, A. E., Wilkinson-White, L., Mackay, J. P., Matthews, J. M., & Blobel, G. A. (2013). Analysis of disease-causing GATA1 mutations in murine gene complementation systems. *Blood*, *121*(26), 5218-5227. https://doi.org/10.1182/blood-2013-03-488080

Campbell, K., Whissell, G., Franch-Marro, X., Batlle, E., & Casanova, J. (2011). Specific GATA Factors Act as Conserved Inducers of an Endodermal-EMT. *Developmental Cell*, *21*(6), 1051-1061. https://doi.org/10.1016/j.devcel.2011.10.005

Cantor, A. B., Iwasaki, H., Arinobu, Y., Moran, T. B., Shigematsu, H., Sullivan, M. R., Akashi, K., & Orkin, S. H. (2008). Antagonism of FOG-1 and GATA factors in fate choice for the mast cell lineage. *The Journal of Experimental Medicine*, *205*(3), 611-624. https://doi.org/10.1084/jem.20070544

Capo, F., Wilson, A., & Di Cara, F. (2019). The Intestine of Drosophila melanogaster : An Emerging Versatile Model System to Study Intestinal Epithelial Homeostasis and Host-Microbial Interactions in Humans. *Microorganisms*, *7*(9). https://doi.org/10.3390/microorganisms7090336

Cavaliere, V., Bernardi, F., Romani, P., Duchi, S., & Gargiulo, G. (2008). Building up the Drosophila eggshell : First of all the eggshell genes must be transcribed. *Developmental Dynamics: An Official Publication of the American Association of Anatomists*, *237*(8), 2061-2072. https://doi.org/10.1002/dvdy.21625

Chang, A. N., Cantor, A. B., Fujiwara, Y., Lodish, M. B., Droho, S., Crispino, J. D., & Orkin, S. H. (2002). GATA-factor dependence of the multitype zinc-finger protein FOG-1 for its essential role in megakaryopoiesis. *Proceedings of the National Academy of Sciences*

of the United States of America, *99*(14), 9237-9242. https://doi.org/10.1073/pnas.142302099

Chatterjee, A., Aavula, K., & Nongthomba, U. (2019). Beadex, a homologue of the vertebrate LIM domain only protein, is a novel regulator of crystal cell development in Drosophila melanogaster. *Journal of Genetics*, *98*.

Cheng, Y., Wu, W., Ashok Kumar, S., Yu, D., Deng, W., Tripic, T., King, D. C., Chen, K.-B., Zhang, Y., Drautz, D., Giardine, B., Schuster, S. C., Miller, W., Chiaromonte, F., Zhang, Y., Blobel, G. A., Weiss, M. J., & Hardison, R. C. (2009). Erythroid GATA1 function revealed by genome-wide analysis of transcription factor occupancy, histone modifications, and mRNA expression. *Genome Research*, *19*(12), 2172-2184. https://doi.org/10.1101/gr.098921.109

Chlon, T. M., & Crispino, J. D. (2012). Combinatorial regulation of tissue specification by GATA and FOG factors. *Development (Cambridge, England)*, *139*(21), 3905-3916. https://doi.org/10.1242/dev.080440

Cirillo, L. A., Lin, F. R., Cuesta, I., Friedman, D., Jarnik, M., & Zaret, K. S. (2002). Opening of compacted chromatin by early developmental transcription factors HNF3 (FoxA) and GATA-4. *Molecular Cell*, *9*(2), 279-289. https://doi.org/10.1016/s1097-2765(02)00459-8

Clowes, C., Boylan, M. G., Ridge, L. A., Barnes, E., Wright, J. A., & Hentges, K. E. (2014). The functional diversity of essential genes required for mammalian cardiac development. *Genesis (New York, N.y. : 2000)*, *52*(8), 713-737. https://doi.org/10.1002/dvg.22794

Crispino, J. D., Lodish, M. B., MacKay, J. P., & Orkin, S. H. (1999). Use of altered specificity mutants to probe a specific protein-protein interaction in differentiation : The GATA-1:FOG complex. *Molecular Cell*, *3*(2), 219-228. https://doi.org/10.1016/s1097-2765(00)80312-3

Crispino, J. D., Lodish, M. B., Thurberg, B. L., Litovsky, S. H., Collins, T., Molkentin, J. D., & Orkin, S. H. (2001). Proper coronary vascular development and heart morphogenesis depend on interaction of GATA-4 with FOG cofactors. *Genes & Development*, *15*(7), 839-844. https://doi.org/10.1101/gad.875201

Crispino, John D., & Horwitz, M. S. (2017). GATA factor mutations in hematologic disease. *Blood*, *129*(15), 2103. https://doi.org/10.1182/blood-2016-09-687889

Crozatier, M., & Meister, M. (2007). Drosophila haematopoiesis. *Cellular Microbiology*, *9*(5), 1117-1126. https://doi.org/10.1111/j.1462-5822.2007.00930.x

Crozatier, M., Ubeda, J.-M., Vincent, A., & Meister, M. (2004). Cellular immune response to parasitization in Drosophila requires the EBF orthologue collier. *PLoS Biology*, *2*(8), E196. https://doi.org/10.1371/journal.pbio.0020196

Cubadda, Y., Heitzler, P., Ray, R. P., Bourouis, M., Ramain, P., Gelbart, W., Simpson, P., & Haenlin, M. (1997). U-shaped encodes a zinc finger protein that regulates the proneural genes achaete and scute during the formation of bristles in Drosophila. *Genes & Development*, *11*(22), 3083-3095. https://doi.org/10.1101/gad.11.22.3083

Dai, Y.-S., & Markham, B. E. (2001). P300 Functions as a Coactivator of Transcription Factor GATA-4. *Journal of Biological Chemistry*, *276*(40), 37178-37185. https://doi.org/10.1074/jbc.M103731200

de Cuevas, M., Lee, J. K., & Spradling, A. C. (1996). Alpha-spectrin is required for germline cell division and differentiation in the Drosophila ovary. *Development (Cambridge, England)*, *122*(12), 3959-3968.

De Franco, E., Shaw-Smith, C., Flanagan, S. E., Shepherd, M. H., Hattersley, A. T., & Ellard, S. (2013). GATA6 Mutations Cause a Broad Phenotypic Spectrum of Diabetes From Pancreatic Agenesis to Adult-Onset Diabetes Without Exocrine Insufficiency. *Diabetes*, *62*(3), 993-997. https://doi.org/10.2337/db12-0885

Debacker, C., Catala, M., & Labastie, M. C. (1999). Embryonic expression of the human GATA-3 gene. *Mechanisms of Development*, *85*(1-2), 183-187. https://doi.org/10.1016/s0925-4773(99)00088-x

DeFalco, T., & Capel, B. (2009). GONAD MORPHOGENESIS IN VERTEBRATES : DIVERGENT MEANS TO A CONVERGENT END. *Annual review of cell and developmental biology*, *25*, 457-482. https://doi.org/10.1146/annurev.cellbio.042308.13350

DiMario, P. J., & Mahowald, A. P. (1987). Female sterile (1) yolkless : A recessive female sterile mutation in Drosophila melanogaster with depressed numbers of coated pits and

coated vesicles within the developing oocytes. *The Journal of Cell Biology*, *105*(1), 199-206. https://doi.org/10.1083/jcb.105.1.199

Dong, X., Tsuda, L., Zavitz, K. H., Lin, M., Li, S., Carthew, R. W., & Zipursky, S. L. (1999). Ebi regulates epidermal growth factor receptor signaling pathways in Drosophila. *Genes & Development*, *13*(8), 954-965.

Doré, L. C., Chlon, T. M., Brown, C. D., White, K. P., & Crispino, J. D. (2012). Chromatin occupancy analysis reveals genome-wide GATA factor switching during hematopoiesis. *Blood*, *119*(16), 3724-3733. https://doi.org/10.1182/blood-2011-09-380634

Dragojlovic-Munther, M., & Martinez-Agosto, J. A. (2013). Extracellular matrix-modulated Heartless signaling in Drosophila blood progenitors regulates their differentiation via a Ras/ETS/FOG pathway and target of rapamycin function. *Developmental Biology*, *384*(2), 313-330. https://doi.org/10.1016/j.ydbio.2013.04.004

Du, J., Stankiewicz, M. J., Liu, Y., Xi, Q., Schmitz, J. E., Lekstrom-Himes, J. A., & Ackerman, S. J. (2002). Novel Combinatorial Interactions of GATA-1, PU.1, and C/EBPε Isoforms Regulate Transcription of the Gene Encoding Eosinophil Granule Major Basic Protein. *Journal of Biological Chemistry*, *277*(45), 43481-43494. https://doi.org/10.1074/jbc.M204777200

Durocher, D., Charron, F., Warren, R., Schwartz, R. J., & Nemer, M. (1997). The cardiac transcription factors Nkx2-5 and GATA-4 are mutual cofactors. *The EMBO Journal*, *16*(18), 5687-5696. https://doi.org/10.1093/emboj/16.18.5687

Dutta, D., Dobson, A. J., Houtz, P. L., Gläßer, C., Revah, J., Korzelius, J., Patel, P. H., Edgar, B. A., & Buchon, N. (2015). Regional Cell-Specific Transcriptome Mapping Reveals Regulatory Complexity in the Adult Drosophila Midgut. *Cell Reports*, *12*(2), 346-358. https://doi.org/10.1016/j.celrep.2015.06.009

Efimenko, E., Padua, M. B., Manuylov, N. L., Fox, S. C., Morse, D. A., & Tevosian, S. G. (2013). The transcription factor GATA4 is required for follicular development and normal ovarian function. *Developmental Biology*, *381*(1), 144-158. https://doi.org/10.1016/j.ydbio.2013.06.004

Elagib, K. E., Racke, F. K., Mogass, M., Khetawat, R., Delehanty, L. L., & Goldfarb, A. N. (2003). RUNX1 and GATA-1 coexpression and cooperation in megakaryocytic

differentiation. *Blood*, *101*(11), 4333-4341. https://doi.org/10.1182/blood-2002-09-2708

Ellis, M. J., Ding, L., Shen, D., Luo, J., Suman, V. J., Wallis, J. W., Van Tine, B. A., Hoog, J., Goiffon, R. J., Goldstein, T. C., Ng, S., Lin, L., Crowder, R., Snider, J., Ballman, K., Weber, J., Chen, K., Koboldt, D. C., Kandoth, C., … Mardis, E. R. (2012). Whole Genome Analysis Informs Breast Cancer Response to Aromatase Inhibition. *Nature*, *486*(7403), 353-360. https://doi.org/10.1038/nature11143

Evans, T., Reitman, M., & Felsenfeld, G. (1988). An erythrocyte-specific DNA-binding factor recognizes a regulatory sequence common to all chicken globin genes. *Proceedings of the National Academy of Sciences of the United States of America*, *85*(16), 5976-5980. https://doi.org/10.1073/pnas.85.16.5976

Fang, R., Olds, L. C., & Sibley, E. (2006). Spatio-temporal patterns of intestine-specific transcription factor expression during postnatal mouse gut development. *Gene Expression Patterns: GEP*, *6*(4), 426-432. https://doi.org/10.1016/j.modgep.2005.09.003

Fedorova, E., Dorogova, N., Bolobolova, E., Fedorova, S., Karagodin, D., Ogienko, A., Khruscheva, A., & Baricheva, E. (2018). GAGA protein is required for multiple aspects of Drosophila oogenesis and female fertility. *genesis*, *57*, e23269. https://doi.org/10.1002/dvg.23269

Force, A., Lynch, M., Pickett, F. B., Amores, A., Yan, Y. L., & Postlethwait, J. (1999). Preservation of duplicate genes by complementary, degenerative mutations. *Genetics*, *151*(4), 1531-1545.

Fossett, N., Tevosian, S. G., Gajewski, K., Zhang, Q., Orkin, S. H., & Schulz, R. A. (2001). The Friend of GATA proteins U-shaped, FOG-1, and FOG-2 function as negative regulators of blood, heart, and eye development in Drosophila. *Proceedings of the National Academy of Sciences of the United States of America*, *98*(13), 7342-7347. https://doi.org/10.1073/pnas.131215798

Fossett, N., Zhang, Q., Gajewski, K., Choi, C. Y., Kim, Y., & Schulz, R. A. (2000). The multitype zinc-finger protein U-shaped functions in heart cell specification in the Drosophila embryo. *Proceedings of the National Academy of Sciences of the United States of America*, *97*(13), 7348-7353. https://doi.org/10.1073/pnas.97.13.7348

Fossett, Nancy, Hyman, K., Gajewski, K., Orkin, S. H., & Schulz, R. A. (2003). Combinatorial interactions of Serpent, Lozenge, and U-shaped regulate crystal cell lineage commitment during Drosophila hematopoiesis. *Proceedings of the National Academy of Sciences*, *100*(20), 11451-11456. https://doi.org/10.1073/pnas.1635050100

Fox, A. H., Liew, C., Holmes, M., Kowalski, K., Mackay, J., & Crossley, M. (1999). Transcriptional cofactors of the FOG family interact with GATA proteins by means of multiple zinc fingers. *The EMBO Journal*, *18*(10), 2812-2822. https://doi.org/10.1093/emboj/18.10.2812

Fox, A., Kowalski, K., King, G., Mackay, J., & Crossley, M. (1999). Key residues characteristic of GATA N-fingers are recognized by FOG. *The Journal of biological chemistry*, *273*, 33595-33603. https://doi.org/10.1074/jbc.273.50.33595

Frandsen, J. L., Gunn, B., Muratoglu, S., Fossett, N., & Newfeld, S. J. (2008). Salmonella pathogenesis reveals that BMP signaling regulates blood cell homeostasis and immune responses in Drosophila. *Proceedings of the National Academy of Sciences of the United States of America*, *105*(39), 14952-14957. https://doi.org/10.1073/pnas.0808208105

Frelin, C., Herrington, R., Janmohamed, S., Barbara, M., Tran, G., Paige, C. J., Benveniste, P., Zuñiga-Pflücker, J.-C., Souabni, A., Busslinger, M., & Iscove, N. N. (2013). GATA-3 REGULATES THE SELF-RENEWAL OF LONG-TERM HEMATOPOIETIC STEM CELLS. *Nature immunology*, *14*(10), 1037-1044. https://doi.org/10.1038/ni.2692

Freson, K., Devriendt, K., Matthijs, G., Van Hoof, A., De Vos, R., Thys, C., Minner, K., Hoylaerts, M. F., Vermylen, J., & Van Geet, C. (2001). Platelet characteristics in patients with X-linked macrothrombocytopenia because of a novel GATA1 mutation. *Blood*, *98*(1), 85-92. https://doi.org/10.1182/blood.v98.1.85

Freson, Kathleen, Matthijs, G., Thys, C., Mariën, P., Hoylaerts, M. F., Vermylen, J., & Van Geet, C. (2002). Different substitutions at residue D218 of the X-linked transcription factor GATA1 lead to altered clinical severity of macrothrombocytopenia and anemia and are associated with variable skewed X inactivation. *Human Molecular Genetics*, *11*(2), 147-152. https://doi.org/10.1093/hmg/11.2.147

Fromental-Ramain, C., Vanolst, L., Delaporte, C., & Ramain, P. (2008). Pannier encodes two structurally related isoforms that are differentially expressed during Drosophila development and display distinct functions during thorax patterning. *Mechanisms of Development*, *125*(1-2), 43-57. https://doi.org/10.1016/j.mod.2007.10.008

Fujiwara, T. (2017). GATA Transcription Factors : Basic Principles and Related Human Disorders. *The Tohoku Journal of Experimental Medicine*, *242*(2), 83-91. https://doi.org/10.1620/tjem.242.83

Fujiwara, Y., Browne, C. P., Cunniff, K., Goff, S. C., & Orkin, S. H. (1996). Arrested development of embryonic red cell precursors in mouse embryos lacking transcription factor GATA-1. *Proceedings of the National Academy of Sciences of the United States of America*, *93*(22), 12355-12358.

Gajewski, K., Fossett, N., Molkentin, J. D., & Schulz, R. A. (1999). The zinc finger proteins Pannier and GATA4 function as cardiogenic factors in Drosophila. *Development (Cambridge, England)*, *126*(24), 5679-5688.

Gajewski, K., Zhang, Q., Choi, C. Y., Fossett, N., Dang, A., Kim, Y. H., Kim, Y., & Schulz, R. A. (2001). Pannier is a transcriptional target and partner of Tinman during Drosophila cardiogenesis. *Developmental Biology*, *233*(2), 425-436. https://doi.org/10.1006/dbio.2001.0220

Gao, H., Baldeosingh, R., Wu, X., & Fossett, N. (2016). The Friend of GATA Transcriptional Co-Regulator, U-Shaped, Is a Downstream Antagonist of Dorsal-Driven Prohemocyte Differentiation in Drosophila. *PloS One*, *11*(5), e0155372. https://doi.org/10.1371/journal.pone.0155372

Gao, H., Wu, X., & Fossett, N. (2009). Upregulation of the Drosophila Friend of GATA gene U-shaped by JAK/STAT signaling maintains lymph gland prohemocyte potency. *Molecular and Cellular Biology*, *29*(22), 6086-6096. https://doi.org/10.1128/MCB.00244-09

Gao, Z., Huang, Z., Olivey, H. E., Gurbuxani, S., Crispino, J. D., & Svensson, E. C. (2010). FOG-1-mediated recruitment of NuRD is required for cell lineage re-enforcement during haematopoiesis. *The EMBO Journal*, *29*(2), 457-468. https://doi.org/10.1038/emboj.2009.368

Garcia-Garcia, M. J., Ramain, P., Simpson, P., & Modolell, J. (1999). Different contributions of pannier and wingless to the patterning of the dorsal mesothorax of Drosophila. *Development*, *126*(16), 3523-3532.

Gavaghan, M. (1998). Cardiac anatomy and physiology : A review. *AORN Journal*, *67*(4), 802-822; quiz 824-828. https://doi.org/10.1016/s0001-2092(06)62644-6

Ghirlando, R., & Trainor, C. D. (2003). Determinants of GATA-1 Binding to DNA THE ROLE OF NON-FINGER RESIDUES. *Journal of Biological Chemistry*, *278*(46), 45620-45628. https://doi.org/10.1074/jbc.M306410200

Gillis, W. J., Bowerman, B., & Schneider, S. Q. (2007). Ectoderm- and endomesoderm-specific GATA transcription factors in the marine annelid Platynereis dumerilli. *Evolution & Development*, *9*(1), 39-50. https://doi.org/10.1111/j.1525-142X.2006.00136.x

Gillis, W. Q., Bowerman, B. A., & Schneider, S. Q. (2008). The evolution of protostome GATA factors : Molecular phylogenetics, synteny, and intron/exon structure reveal orthologous relationships. *BMC Evolutionary Biology*, *8*, 112. https://doi.org/10.1186/1471-2148-8-112

Gloor, G. B., Preston, C. R., Johnson-Schlitz, D. M., Nassif, N. A., Phillis, R. W., Benz, W. K., Robertson, H. M., & Engels, W. R. (1993). Type I Repressors of P Element Mobility. *Genetics*, *135*(1), 81-95.

Gold, K. S., & Brückner, K. (2015). Macrophages and cellular immunity in Drosophila melanogaster. *Seminars in Immunology*, *27*(6), 357-368. https://doi.org/10.1016/j.smim.2016.03.010

Grass, J. A., Boyer, M. E., Pal, S., Wu, J., Weiss, M. J., & Bresnick, E. H. (2003). GATA-1-dependent transcriptional repression of GATA-2 via disruption of positive autoregulation and domain-wide chromatin remodeling. *Proceedings of the National Academy of Sciences of the United States of America*, *100*(15), 8811-8816. https://doi.org/10.1073/pnas.1432147100

Grass, J. A., Jing, H., Kim, S.-I., Martowicz, M. L., Pal, S., Blobel, G. A., & Bresnick, E. H. (2006). Distinct functions of dispersed GATA factor complexes at an endogenous gene locus. *Molecular and Cellular Biology*, *26*(19), 7056-7067. https://doi.org/10.1128/MCB.01033-06

Gratz, S. J., Rubinstein, C. D., Harrison, M. M., Wildonger, J., & O'Connor-Giles, K. M. (2015). CRISPR-Cas9 Genome Editing in Drosophila. *Current Protocols in Molecular Biology*, *111*, 31.2.1-31.2.20. https://doi.org/10.1002/0471142727.mb3102s111

Greenwood-Van Meerveld, B., Johnson, A. C., & Grundy, D. (2017). Gastrointestinal Physiology and Function. *Handbook of Experimental Pharmacology*, *239*, 1-16. https://doi.org/10.1007/164_2016_118

Gregory, G. D., Miccio, A., Bersenev, A., Wang, Y., Hong, W., Zhang, Z., Poncz, M., Tong, W., & Blobel, G. A. (2010). FOG1 requires NuRD to promote hematopoiesis and maintain lineage fidelity within the megakaryocytic-erythroid compartment. *Blood*, *115*(11), 2156-2166. https://doi.org/10.1182/blood-2009-10-251280

Gutiérrez, L., Nikolic, T., van Dijk, T. B., Hammad, H., Vos, N., Willart, M., Grosveld, F., Philipsen, S., & Lambrecht, B. N. (2007). Gata1 regulates dendritic-cell development and survival. *Blood*, *110*(6), 1933-1941. https://doi.org/10.1182/blood-2006-09-048322

Haenlin, M., Cubadda, Y., Blondeau, F., Heitzler, P., Lutz, Y., Simpson, P., & Ramain, P. (1997). Transcriptional activity of pannier is regulated negatively by heterodimerization of the GATA DNA-binding domain with a cofactor encoded by the u-shaped gene of Drosophila. *Genes & Development*, *11*(22), 3096-3108. https://doi.org/10.1101/gad.11.22.3096

Hames, B. D., & Bownes, M. (1978). Synthesis of yolk proteins in Drosophila melanogaster. *Insect Biochemistry*, *8*(5), 319-328. https://doi.org/10.1016/0020-1790(78)90016-1

Han, Z., & Olson, E. N. (2005). Hand is a direct target of Tinman and GATA factors during Drosophila cardiogenesis and hematopoiesis. *Development (Cambridge, England)*, *132*(15), 3525-3536. https://doi.org/10.1242/dev.01899

Harigae, H., Takahashi, S., Suwabe, N., Ohtsu, H., Gu, L., Yang, Z., Tsai, F. Y., Kitamura, Y., Engel, J. D., & Yamamoto, M. (1998). Differential roles of GATA-1 and GATA-2 in growth and differentiation of mast cells. *Genes to Cells: Devoted to Molecular & Cellular Mechanisms*, *3*(1), 39-50. https://doi.org/10.1046/j.1365-2443.1998.00166.x

Hartenstein, V. (1993). *Atlas of Drosophila Development*. Cold Spring Harbor Laboratory Press.

Hartenstein, V. (2006). Blood cells and blood cell development in the animal kingdom. *Annual Review of Cell and Developmental Biology*, *22*, 677-712. https://doi.org/10.1146/annurev.cellbio.22.010605.093317

Heitzler, P., Haenlin, M., Ramain, P., Calleja, M., & Simpson, P. (1996). A genetic analysis of pannier, a gene necessary for viability of dorsal tissues and bristle positioning in Drosophila. *Genetics*, *143*(3), 1271-1286.

Hernández de Madrid, B., & Casanova, J. (2018). GATA factor genes in the Drosophila midgut embryo. *PLoS ONE*, *13*(3). https://doi.org/10.1371/journal.pone.0193612

Hirasawa, R., Shimizu, R., Takahashi, S., Osawa, M., Takayanagi, S., Kato, Y., Onodera, M., Minegishi, N., Yamamoto, M., Fukao, K., Taniguchi, H., Nakauchi, H., & Iwama, A. (2002). Essential and instructive roles of GATA factors in eosinophil development. *The Journal of Experimental Medicine*, *195*(11), 1379-1386. https://doi.org/10.1084/jem.20020170

Ho, I.-C., Tai, T.-S., & Pai, S.-Y. (2009). GATA3 and the T-cell lineage : Essential functions before and after T-helper-2-cell differentiation. *Nature Reviews Immunology*, *9*(2), 125-135. https://doi.org/10.1038/nri2476

Holz, A., Bossinger, B., Strasser, T., Janning, W., & Klapper, R. (2003). The two origins of hemocytes in Drosophila. *Development (Cambridge, England)*, *130*(20), 4955-4962. https://doi.org/10.1242/dev.00702

Hong, W., Nakazawa, M., Chen, Y.-Y., Kori, R., Vakoc, C. R., Rakowski, C., & Blobel, G. A. (2005). FOG-1 recruits the NuRD repressor complex to mediate transcriptional repression by GATA-1. *The EMBO Journal*, *24*(13), 2367-2378. https://doi.org/10.1038/sj.emboj.7600703

Hoppe, P. S., Schwarzfischer, M., Loeffler, D., Kokkaliaris, K. D., Hilsenbeck, O., Moritz, N., Endele, M., Filipczyk, A., Gambardella, A., Ahmed, N., Etzrodt, M., Coutu, D. L., Rieger, M. A., Marr, C., Strasser, M. K., Schauberger, B., Burtscher, I., Ermakova, O., Bürger, A., … Schroeder, T. (2016). Early myeloid lineage choice is not initiated by random PU.1 to GATA1 protein ratios. *Nature*, *535*(7611), 299-302. https://doi.org/10.1038/nature18320

Hu, Y.-C., Okumura, L. M., & Page, D. C. (2013). Gata4 is required for formation of the genital ridge in mice. *PLoS Genetics*, *9*(7), e1003629. https://doi.org/10.1371/journal.pgen.1003629

Huang, J., Liu, X., Li, D., Shao, Z., Cao, H., Zhang, Y., Trompouki, E., Bowman, T. V., Zon, L. I., Yuan, G.-C., Orkin, S. H., & Xu, J. (2016). Dynamic Control of Enhancer Repertoires Drives Lineage and Stage-Specific Transcription during Hematopoiesis. *Developmental Cell*, *36*(1), 9-23. https://doi.org/10.1016/j.devcel.2015.12.014

Ikonomi, P., Rivera, C. E., Riordan, M., Washington, G., Schechter, A. N., & Noguchi, C. T. (2000). Overexpression of GATA-2 inhibits erythroid and promotes megakaryocyte differentiation. *Experimental Hematology*, *28*(12), 1423-1431. https://doi.org/10.1016/S0301-472X(00)00553-1

Immarigeon, C., Bernat-Fabre, S., Augé, B., Faucher, C., Gobert, V., Haenlin, M., Waltzer, L., Payet, A., Cribbs, D. L., Bourbon, H.-M. G., & Boube, M. (2019). Drosophila Mediator Subunit Med1 Is Required for GATA-Dependent Developmental Processes : Divergent Binding Interfaces for Conserved Coactivator Functions. *Molecular and Cellular Biology*, *39*(7). https://doi.org/10.1128/MCB.00477-18

Ivanovs, A., Rybtsov, S., Ng, E. S., Stanley, E. G., Elefanty, A. G., & Medvinsky, A. (2017). Human haematopoietic stem cell development : From the embryo to the dish. *Development (Cambridge, England)*, *144*(13), 2323-2337. https://doi.org/10.1242/dev.134866

Jiang, H., Grenley, M. O., Bravo, M.-J., Blumhagen, R. Z., & Edgar, B. A. (2011). EGFR/Ras/MAPK Signaling Mediates Adult Midgut Epithelial Homeostasis and Regeneration in Drosophila. *Cell Stem Cell*, *8*(1), 84-95. https://doi.org/10.1016/j.stem.2010.11.026

Jing, H., Vakoc, C. R., Ying, L., Mandat, S., Wang, H., Zheng, X., & Blobel, G. A. (2008). Exchange of GATA factors mediates transitions in looped chromatin organization at a developmentally regulated gene locus. *Molecular Cell*, *29*(2), 232-242. https://doi.org/10.1016/j.molcel.2007.11.020

Jung, S.-H., Evans, C. J., Uemura, C., & Banerjee, U. (2005). The Drosophila lymph gland as a developmental model of hematopoiesis. *Development (Cambridge, England)*, *132*(11), 2521-2533. https://doi.org/10.1242/dev.01837

Kadam, S., McAlpine, G. S., Phelan, M. L., Kingston, R. E., Jones, K. A., & Emerson, B. M. (2000). Functional selectivity of recombinant mammalian SWI/SNF subunits. *Genes & Development*, *14*(19), 2441-2451. https://doi.org/10.1101/gad.828000

Kakita, T., Hasegawa, K., Morimoto, T., Kaburagi, S., Wada, H., & Sasayama, S. (1999). P300 Protein as a Coactivator of GATA-5 in the Transcription of Cardiac-restricted Atrial Natriuretic Factor Gene. *Journal of Biological Chemistry*, *274*(48), 34096-34102. https://doi.org/10.1074/jbc.274.48.34096

Kaneko, H., Kobayashi, E., Yamamoto, M., & Shimizu, R. (2012). N- and C-terminal Transactivation Domains of GATA1 Protein Coordinate Hematopoietic Program. *The Journal of Biological Chemistry*, *287*(25), 21439-21449. https://doi.org/10.1074/jbc.M112.370437

Katsumura, K. R., & Bresnick, E. H. (2017). The GATA factor revolution in hematology. *Blood*, *129*(15), 2092-2102. https://doi.org/10.1182/blood-2016-09-687871

Kim, S.-I., Bultman, S. J., Kiefer, C. M., Dean, A., & Bresnick, E. H. (2009). BRG1 requirement for long-range interaction of a locus control region with a downstream promoter. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(7), 2259-2264. https://doi.org/10.1073/pnas.0806420106

Klinedinst, S. L., & Bodmer, R. (2003). Gata factor Pannier is required to establish competence for heart progenitor formation. *Development (Cambridge, England)*, *130*(13), 3027-3038. https://doi.org/10.1242/dev.00517

Kouros-Mehr, H., Slorach, E. M., Sternlicht, M. D., & Werb, Z. (2006). GATA-3 maintains the differentiation of the luminal cell fate in the mammary gland. *Cell*, *127*(5), 1041-1055. https://doi.org/10.1016/j.cell.2006.09.048

Kowalski, K., Czolij, R., King, G. F., Crossley, M., & Mackay, J. P. (1999). The solution structure of the N-terminal zinc finger of GATA-1 reveals a specific binding face for the transcriptional co-factor FOG. *Journal of Biomolecular NMR*, *13*(3), 249-262. https://doi.org/10.1023/a:1008309602929

Krishnan, A., Samtani, R., Dhanantwari, P., Lee, E., Yamada, S., Shiota, K., Donofrio, M. T., Leatherbury, L., & Lo, C. W. (2014). A Detailed Comparison of Mouse and Human Cardiac Development. *Pediatric research*, *76*(6), 500-507. https://doi.org/10.1038/pr.2014.128

Ku, C.-J., Hosoya, T., Maillard, I., & Engel, J. D. (2012). GATA-3 regulates hematopoietic stem cell maintenance and cell-cycle entry. *Blood*, *119*(10), 2242-2251. https://doi.org/10.1182/blood-2011-07-366070

Kuo, C. T., Morrisey, E. E., Anandappa, R., Sigrist, K., Lu, M. M., Parmacek, M. S., Soudais, C., & Leiden, J. M. (1997). GATA4 transcription factor is required for ventral morphogenesis and heart tube formation. *Genes & Development*, *11*(8), 1048-1060. https://doi.org/10.1101/gad.11.8.1048

Kurek, D., Garinis, G. A., van Doorninck, J. H., van der Wees, J., & Grosveld, F. G. (2007). Transcriptome and phenotypic analysis reveals Gata3-dependent signalling pathways in murine hair follicles. *Development (Cambridge, England)*, *134*(2), 261-272. https://doi.org/10.1242/dev.02721

Kuuluvainen, E., Hakala, H., Havula, E., Sahal Estimé, M., Rämet, M., Hietakangas, V., & Mäkelä, T. P. (2014). Cyclin-dependent kinase 8 module expression profiling reveals requirement of mediator subunits 12 and 13 for transcription of Serpent-dependent innate immunity genes in Drosophila. *The Journal of Biological Chemistry*, *289*(23), 16252-16261. https://doi.org/10.1074/jbc.M113.541904

Kyrönlahti, A., Vetter, M., Euler, R., Bielinska, M., Jay, P. Y., Anttonen, M., Heikinheimo, M., & Wilson, D. B. (2011). GATA4 deficiency impairs ovarian function in adult mice. *Biology of Reproduction*, *84*(5), 1033-1044. https://doi.org/10.1095/biolreprod.110.086850

Labastie, M.-C., Catala, M., Gregoire, J.-M., & Peault, B. (1995). The GATA-3 gene is expressed during human kidney embryogenesis. *Kidney International*, *47*(6), 1597-1603. https://doi.org/10.1038/ki.1995.223

Lamers Wouter H., & Moorman Antoon F.M. (2002). Cardiac Septation. *Circulation Research*, *91*(2), 93-103. https://doi.org/10.1161/01.RES.0000027135.63141.89

Lanot, R., Zachary, D., Holder, F., & Meister, M. (2001). Postembryonic hematopoiesis in Drosophila. *Developmental Biology*, *230*(2), 243-257. https://doi.org/10.1006/dbio.2000.0123

Lantz, V., Chang, J. S., Horabin, J. I., Bopp, D., & Schedl, P. (1994). The Drosophila orb RNA-binding protein is required for the formation of the egg chamber and establishment of polarity. *Genes & Development*, *8*(5), 598-613. https://doi.org/10.1101/gad.8.5.598

Lavallée, G., Andelfinger, G., Nadeau, M., Lefebvre, C., Nemer, G., Horb, M. E., & Nemer, M. (2006). The Kruppel-like transcription factor KLF13 is a novel regulator of heart development. *The EMBO Journal*, *25*(21), 5201-5213. https://doi.org/10.1038/sj.emboj.7601379

Laverriere, A. C., MacNeill, C., Mueller, C., Poelmann, R. E., Burch, J. B., & Evans, T. (1994). GATA-4/5/6, a subfamily of three transcription factors transcribed in developing heart and gut. *The Journal of Biological Chemistry*, *269*(37), 23177-23184.

Lebestky, T., Chang, T., Hartenstein, V., & Banerjee, U. (2000). Specification of Drosophila hematopoietic lineage by conserved transcription factors. *Science (New York, N.Y.)*, *288*(5463), 146-149. https://doi.org/10.1126/science.288.5463.146

Lebestky, Tim, Jung, S.-H., & Banerjee, U. (2003). A Serrate-expressing signaling center controls Drosophila hematopoiesis. *Genes & Development*, *17*(3), 348-353. https://doi.org/10.1101/gad.1052803

Lemaitre, B., & Miguel-Aliaga, I. (2013). The digestive tract of Drosophila melanogaster. *Annual Review of Genetics*, *47*, 377-404. https://doi.org/10.1146/annurev-genet-111212-133343

Lentjes, M. H. F. M., Niessen, H. E. C., Akiyama, Y., de Bruïne, A. P., Melotte, V., & van Engeland, M. (2016). The emerging role of GATA transcription factors in development and disease. *Expert Reviews in Molecular Medicine*, *18*, e3. https://doi.org/10.1017/erm.2016.2

Lepesant, J. M. J., Iampietro, C., Galeota, E., Augé, B., Aguirrenbengoa, M., Mercé, C., Chaubet, C., Rocher, V., Haenlin, M., Waltzer, L., Pelizzola, M., & Di Stefano, L. (2020). A dual role of dLsd1 in oogenesis : Regulating developmental genes and repressing transposons. *Nucleic Acids Research*, *48*(3), 1206-1224. https://doi.org/10.1093/nar/gkz1142

Leptin, M. (1995). Drosophila gastrulation : From pattern formation to morphogenesis. *Annual Review of Cell and Developmental Biology*, *11*, 189-212. https://doi.org/10.1146/annurev.cb.11.110195.001201

Letourneau, M., Lapraz, F., Sharma, A., Vanzo, N., Waltzer, L., & Crozatier, M. (2016). Drosophila hematopoiesis under normal conditions and in response to immune stress. *FEBS Letters*, *590*(22), 4034-4051. https://doi.org/10.1002/1873-3468.12327

Li, X., Ma, M., Liu, F., Chen, Y., Lu, A., Ling, Q.-Z., Li, J., Beerntsen, B. T., Yu, X.-Q., Liu, C., & Ling, E. (2012). Properties of Drosophila melanogaster prophenoloxidases expressed in Escherichia coli. *Developmental & Comparative Immunology*, *36*(4), 648-656. https://doi.org/10.1016/j.dci.2011.11.005

Lin, W.-H., Huang, L.-H., Yeh, J.-Y., Hoheisel, J., Lehrach, H., Sun, Y. H., & Tsai, S.-F. (1995). Expression of a Drosophila GATA Transcription Factor in Multiple Tissues in the Developing Embryos IDENTIFICATION OF HOMOZYGOUS LETHAL MUTANTS WITH P-ELEMENT INSERTION AT THE PROMOTER REGION. *Journal of Biological Chemistry*, *270*(42), 25150-25158. https://doi.org/10.1074/jbc.270.42.25150

Lino, C. A., Harper, J. C., Carney, J. P., & Timlin, J. A. (2018). Delivering CRISPR : A review of the challenges and approaches. *Drug Delivery*, *25*(1), 1234-1257. https://doi.org/10.1080/10717544.2018.1474964

Liu, C., Glasser, S. W., Wan, H., & Whitsett, J. A. (2002). GATA-6 and thyroid transcription factor-1 directly interact and regulate surfactant protein-C gene expression. *The Journal of Biological Chemistry*, *277*(6), 4519-4525. https://doi.org/10.1074/jbc.M107585200

Lossky, M., & Wensink, P. C. (1995). Regulation of Drosophila yolk protein genes by an ovary-specific GATA factor. *Molecular and Cellular Biology*, *15*(12), 6943-6952. https://doi.org/10.1128/MCB.15.12.6943

Lourenço, D., Brauner, R., Rybczyńska, M., Nihoul-Fékété, C., McElreavey, K., & Bashamboo, A. (2011). Loss-of-function mutation in GATA4 causes anomalies of human testicular development. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(4), 1597-1602. https://doi.org/10.1073/pnas.1010257108

Lowry, J. A., & Atchley, W. R. (2000). Molecular evolution of the GATA family of transcription factors : Conservation within the DNA-binding domain. *Journal of Molecular Evolution*, *50*(2), 103-115. https://doi.org/10.1007/s002399910012

Lowry, Jason A., & Mackay, J. P. (2006). GATA-1 : One protein, many partners. *The International Journal of Biochemistry & Cell Biology*, *38*(1), 6-11. https://doi.org/10.1016/j.biocel.2005.06.017

Lu, J., McKinsey, T. A., Xu, H., Wang, D., Richardson, J. A., & Olson, E. N. (1999). FOG-2, a Heart- and Brain-Enriched Cofactor for GATA Transcription Factors. *Molecular and Cellular Biology*, *19*(6), 4495-4502.

Mahowald, A. P. (1962). Fine structure of pole cells and polar granules in Drosophila melanogaster. *Journal of Experimental Zoology*, *151*(3), 201-215. https://doi.org/10.1002/jez.1401510302

Mandal, L., Banerjee, U., & Hartenstein, V. (2004). Evidence for a fruit fly hemangioblast and similarities between lymph-gland hematopoiesis in fruit fly and mammal aorta-gonadal-mesonephros mesoderm. *Nature Genetics*, *36*(9), 1019-1023. https://doi.org/10.1038/ng1404

Manuylov, N. L., Smagulova, F. O., Leach, L., & Tevosian, S. G. (2008). Ovarian development in mice requires the GATA4-FOG2 transcription complex. *Development*, *135*(22), 3731-3743. https://doi.org/10.1242/dev.024653

Martin, D. I., & Orkin, S. H. (1990). Transcriptional activation and DNA binding by the erythroid factor GF-1/NF-E1/Eryf 1. *Genes & Development*, *4*(11), 1886-1898. https://doi.org/10.1101/gad.4.11.1886

Martinez de LaPiscina, I., de Mingo, C., Riedl, S., Rodriguez, A., Pandey, A. V., Fernández-Cancio, M., Camats, N., Sinclair, A., Castaño, L., Audi, L., & Flück, C. E. (2018). GATA4 Variants in Individuals With a 46,XY Disorder of Sex Development (DSD) May or May Not Be Associated With Cardiac Defects Depending on Second Hits in Other DSD Genes. *Frontiers in Endocrinology*, *9*. https://doi.org/10.3389/fendo.2018.00142

Martínez-Corrales, G., Cabrero, P., Dow, J. A. T., Terhzaz, S., & Davies, S.-A. (2019). Novel roles for GATAe in growth, maintenance and proliferation of cell populations in the Drosophila renal tubule. *Development (Cambridge, England)*, *146*(9). https://doi.org/10.1242/dev.178087

Martowicz, M. L., Grass, J. A., Boyer, M. E., Guend, H., & Bresnick, E. H. (2005). Dynamic GATA factor interplay at a multicomponent regulatory region of the GATA-2 locus. *The Journal of Biological Chemistry*, *280*(3), 1724-1732. https://doi.org/10.1074/jbc.M406038200

Martynova, E., Zhao, Y., Xie, Q., Zheng, D., & Cvekl, A. (2019). Transcriptomic analysis and novel insights into lens fibre cell differentiation regulated by Gata3. *Open Biology*, *9*, 190220. https://doi.org/10.1098/rsob.190220

McDevitt, M. A., Shivdasani, R. A., Fujiwara, Y., Yang, H., & Orkin, S. H. (1997). A « knockdown » mutation created by cis-element gene targeting reveals the dependence of erythroid cell maturation on the level of transcription factor GATA-1. *Proceedings of the National Academy of Sciences of the United States of America*, *94*(13), 6781-6785. https://doi.org/10.1073/pnas.94.13.6781

Mehaffey, M. G., Newton, A. L., Gandhi, M. J., Crossley, M., & Drachman, J. G. (2001). X-linked thrombocytopenia caused by a novel mutation of GATA-1. *Blood*, *98*(9), 2681-2688. https://doi.org/10.1182/blood.v98.9.2681

Merika, M., & Orkin, S. H. (1995). Functional synergy and physical interactions of the erythroid transcription factor GATA-1 with the Krüppel family proteins Sp1 and EKLF. *Molecular and Cellular Biology*, *15*(5), 2437-2447. https://doi.org/10.1128/mcb.15.5.2437

Miguel-Aliaga, I., Jasper, H., & Lemaitre, B. (2018). Anatomy and Physiology of the Digestive Tract of Drosophila melanogaster. *Genetics*, *210*(2), 357-396. https://doi.org/10.1534/genetics.118.300224

Minakhina, S., Tan, W., & Steward, R. (2011). JAK/STAT and the GATA factor Pannier control hemocyte maturation and differentiation in Drosophila. *Developmental Biology*, *352*(2), 308-316. https://doi.org/10.1016/j.ydbio.2011.01.035

Mirzoyan, Z., Sollazzo, M., Allocca, M., Valenza, A. M., Grifoni, D., & Bellosta, P. (2019). Drosophila melanogaster : A Model Organism to Study Cancer. *Frontiers in Genetics*, *10*. https://doi.org/10.3389/fgene.2019.00051

Molkentin, J. D., Lin, Q., Duncan, S. A., & Olson, E. N. (1997). Requirement of the transcription factor GATA4 for heart tube formation and ventral morphogenesis. *Genes & Development*, *11*(8), 1061-1072. https://doi.org/10.1101/gad.11.8.1061

Molkentin, J. D., Tymitz, K. M., Richardson, J. A., & Olson, E. N. (2000). Abnormalities of the genitourinary tract in female mice lacking GATA5. *Molecular and Cellular Biology*, *20*(14), 5256-5260. https://doi.org/10.1128/mcb.20.14.5256-5260.2000

Montell, D. J., Rorth, P., & Spradling, A. C. (1992). Slow border cells, a locus required for a developmentally regulated cell migration during oogenesis, encodes Drosophila C/EBP. *Cell*, *71*(1), 51-62. https://doi.org/10.1016/0092-8674(92)90265-e

Morimoto, T., Hasegawa, K., Kaburagi, S., Kakita, T., Masutani, H., Kitsis, R. N., Matsumori, A., & Sasayama, S. (1999). GATA-5 is involved in leukemia inhibitory factor-responsive transcription of the beta-myosin heavy chain gene in cardiac myocytes. *The Journal of Biological Chemistry*, *274*(18), 12811-12818. https://doi.org/10.1074/jbc.274.18.12811

Morin, S., Charron, F., Robitaille, L., & Nemer, M. (2000). GATA-dependent recruitment of MEF2 proteins to target promoters. *The EMBO Journal*, *19*(9), 2046-2055. https://doi.org/10.1093/emboj/19.9.2046

Morin, Steves, Paradis, P., Aries, A., & Nemer, M. (2001). Serum Response Factor-GATA Ternary Complex Required for Nuclear Signaling by a G-Protein-Coupled Receptor. *Molecular and Cellular Biology*, *21*(4), 1036-1044. https://doi.org/10.1128/MCB.21.4.1036-1044.2001

Morrisey, E. E., Ip, H. S., Tang, Z., Lu, M. M., & Parmacek, M. S. (1997). GATA-5 : A transcriptional activator expressed in a novel temporally and spatially-restricted pattern during embryonic development. *Developmental Biology*, *183*(1), 21-36. https://doi.org/10.1006/dbio.1996.8485

Morrisey, E. E., Ip, H. S., Tang, Z., & Parmacek, M. S. (1997). GATA-4 activates transcription via two novel domains that are conserved within the GATA-4/5/6 subfamily. *The Journal of Biological Chemistry*, *272*(13), 8515-8524. https://doi.org/10.1074/jbc.272.13.8515

Morrisey, Edward E., Ip, H. S., Lu, M. M., & Parmacek, M. S. (1996). GATA-6 : A Zinc Finger Transcription Factor That Is Expressed in Multiple Cell Lineages Derived from Lateral Mesoderm. *Developmental Biology*, *177*(1), 309-322. https://doi.org/10.1006/dbio.1996.0165

Murakami, R., Takashima, S., & Hamaguchi, T. (1999). Developmental genetics of the Drosophila gut : Specification of primordia, subdivision and overt-differentiation. *Cellular and Molecular Biology (Noisy-Le-Grand, France)*, *45*(5), 661-676.

Murakami, Ryutaro, Shigenaga, A., Matsumoto, A., Yamaoka, I., & Tanimura, T. (1994). Novel tissue units of regional differentiation in the gut epithelium of Drosopbila, as revealed by P-element-mediated detection of enhancer. *Roux's Archives of Developmental Biology: The Official Organ of the EDBO*, *203*(5), 243-249. https://doi.org/10.1007/BF00360519

Muratoglu, S., Hough, B., Mon, S. T., & Fossett, N. (2007). The GATA factor Serpent cross-regulates lozenge and u-shaped expression during Drosophila blood cell development. *Developmental biology*, *311*(2), 636-649. https://doi.org/10.1016/j.ydbio.2007.08.015

Muroya, K., Hasegawa, T., Ito, Y., Nagai, T., Isotani, H., Iwata, Y., Yamamoto, K., Fujimoto, S., Seishu, S., Fukushima, Y., Hasegawa, Y., & Ogata, T. (2001). GATA3 abnormalities and the phenotypic spectrum of HDR syndrome. *Journal of Medical Genetics*, *38*(6), 374-380. https://doi.org/10.1136/jmg.38.6.374

Nelson, R. E., Fessler, L. I., Takagi, Y., Blumberg, B., Keene, D. R., Olson, P. F., Parker, C. G., & Fessler, J. H. (1994). Peroxidasin : A novel enzyme-matrix protein of Drosophila development. *The EMBO Journal*, *13*(15), 3438-3447.

Nerlov, C., Querfurth, E., Kulessa, H., & Graf, T. (2000). GATA-1 interacts with the myeloid PU.1 transcription factor and represses PU.1-dependent transcription. *Blood*, *95*(8), 2543-2551.

Nichols, K. E., Crispino, J. D., Poncz, M., White, J. G., Orkin, S. H., Maris, J. M., & Weiss, M. J. (2000). Familial dyserythropoietic anaemia and thrombocytopenia due to an inherited mutation in GATA1. *Nature Genetics*, *24*(3), 266-270. https://doi.org/10.1038/73480

Noah, T. K., Donahue, B., & Shroyer, N. F. (2011). Intestinal development and differentiation. *Experimental cell research*, *317*(19), 2702-2710. https://doi.org/10.1016/j.yexcr.2011.09.006

Okawa, T., Yoshida, M., Usui, T., Kudou, T., Iwasaki, Y., Fukuoka, K., Takahashi, N., Uehara, Y., & Oiso, Y. (2015). A novel loss-of-function mutation of GATA3 (p.R299Q) in a Japanese family with Hypoparathyroidism, Deafness, and Renal Dysplasia (HDR) syndrome. *BMC Endocrine Disorders*, *15*. https://doi.org/10.1186/s12902-015-0065-7

Okumura, T., Matsumoto, A., Tanimura, T., & Murakami, R. (2005). An endoderm-specific GATA factor gene, dGATAe, is required for the terminal differentiation of the Drosophila endoderm. *Developmental Biology*, *278*(2), 576-586. https://doi.org/10.1016/j.ydbio.2004.11.021

Okumura, T., Tajiri, R., Kojima, T., Saigo, K., & Murakami, R. (2007). GATAe-dependent and -independent expressions of genes in the differentiated endodermal midgut of Drosophila. *Gene Expression Patterns*, *7*(1), 178-186. https://doi.org/10.1016/j.modgep.2006.07.001

Okumura, T., Takeda, K., Kuchiki, M., Akaishi, M., Taniguchi, K., & Adachi-Yamada, T. (2016). GATAe regulates intestinal stem cell maintenance and differentiation in Drosophila adult midgut. *Developmental Biology*, *410*(1), 24-35. https://doi.org/10.1016/j.ydbio.2015.12.017

Omichinski, J. G., Clore, G. M., Schaad, O., Felsenfeld, G., Trainor, C., Appella, E., Stahl, S. J., & Gronenborn, A. M. (1993). NMR structure of a specific DNA complex of Zn-

containing DNA binding domain of GATA-1. *Science (New York, N.Y.)*, *261*(5120), 438-446. https://doi.org/10.1126/science.8332909

Organs. (2015, juin 7). *Droso4schools*. https://droso4schools.wordpress.com/organs/

Orkin, S. H. (1990). Globin gene regulation and switching : Circa 1990. *Cell*, *63*(4), 665-672. https://doi.org/10.1016/0092-8674(90)90133-y

Osada, H., Grutz, G., Axelson, H., Forster, A., & Rabbitts, T. H. (1995). Association of erythroid transcription factors : Complexes involving the LIM protein RBTN2 and the zinc-finger protein GATA1. *Proceedings of the National Academy of Sciences of the United States of America*, *92*(21), 9585-9589.

Oyallon, J., Vanzo, N., Krzemień, J., Morin-Poulard, I., Vincent, A., & Crozatier, M. (2016). Two Independent Functions of Collier/Early B Cell Factor in the Control of Drosophila Blood Cell Homeostasis. *PloS One*, *11*(2), e0148978. https://doi.org/10.1371/journal.pone.0148978

Padua, M. B., Fox, S. C., Jiang, T., Morse, D. A., & Tevosian, S. G. (2014). Simultaneous gene deletion of gata4 and gata6 leads to early disruption of follicular development and germ cell loss in the murine ovary. *Biology of Reproduction*, *91*(1), 24. https://doi.org/10.1095/biolreprod.113.117002

Patient, R. K., & McGhee, J. D. (2002). The GATA family (vertebrates and invertebrates). *Current Opinion in Genetics & Development*, *12*(4), 416-422. https://doi.org/10.1016/s0959-437x(02)00319-2

Pedone, P. V., Omichinski, J. G., Nony, P., Trainor, C., Gronenborn, A. M., Clore, G. M., & Felsenfeld, G. (1997). The N-terminal fingers of chicken GATA-2 and GATA-3 are independent sequence-specific DNA binding domains. *The EMBO Journal*, *16*(10), 2874-2882. https://doi.org/10.1093/emboj/16.10.2874

Peterkin, T., Gibson, A., Loose, M., & Patient, R. (2005). The roles of GATA-4, -5 and -6 in vertebrate heart development. *Seminars in Cell & Developmental Biology*, *16*(1), 83-94. https://doi.org/10.1016/j.semcdb.2004.10.003

Pevny, L., Lin, C. S., D'Agati, V., Simon, M. C., Orkin, S. H., & Costantini, F. (1995). Development of hematopoietic cells lacking transcription factor GATA-1. *Development (Cambridge, England)*, *121*(1), 163-172.

Pevny, L., Simon, M. C., Robertson, E., Klein, W. H., Tsai, S. F., D'Agati, V., Orkin, S. H., & Costantini, F. (1991). Erythroid differentiation in chimaeric mice blocked by a targeted mutation in the gene for transcription factor GATA-1. *Nature*, *349*(6306), 257-260. https://doi.org/10.1038/349257a0

Phillips, J. D., Steensma, D. P., Pulsipher, M. A., Spangrude, G. J., & Kushner, J. P. (2007). Congenital erythropoietic porphyria due to a mutation in GATA1 : The first trans-acting mutation causative for a human porphyria. *Blood*, *109*(6), 2618-2621. https://doi.org/10.1182/blood-2006-06-022848

Pikkarainen, S., Tokola, H., Kerkelä, R., & Ruskoaho, H. (2004). GATA transcription factors in the developing and adult heart. *Cardiovascular Research*, *63*(2), 196-207. https://doi.org/10.1016/j.cardiores.2004.03.025

Ping, N., Sun, A., Song, Y., Wang, Q., Yin, J., Cheng, W., Xu, Y., Wen, L., Yao, H., Ma, L., Qiu, H., Ruan, C., Wu, D., & Chen, S. (2017). Exome sequencing identifies highly recurrent somatic GATA2 and CEBPA mutations in acute erythroid leukemia. *Leukemia*, *31*(1), 195-202. https://doi.org/10.1038/leu.2016.162

Pitsouli, C., & Perrimon, N. (2008). Developmental biology : Our fly cousins' gut. *Nature*, *454*(7204), 592-593. https://doi.org/10.1038/454592a

Port, F., Chen, H.-M., Lee, T., & Bullock, S. L. (2014). Optimized CRISPR/Cas tools for efficient germline and somatic genome engineering in Drosophila. *Proceedings of the National Academy of Sciences of the United States of America*, *111*(29), E2967-E2976. https://doi.org/10.1073/pnas.1405500111

Querfurth, E., Schuster, M., Kulessa, H., Crispino, J. D., Döderlein, G., Orkin, S. H., Graf, T., & Nerlov, C. (2000). Antagonism between C/EBPβ and FOG in eosinophil lineage commitment of multipotent hematopoietic progenitors. *Genes & Development*, *14*(19), 2515-2525. https://doi.org/10.1101/gad.177200

Ramain, P., Heitzler, P., Haenlin, M., & Simpson, P. (1993). Pannier, a negative regulator of achaete and scute in Drosophila, encodes a zinc finger protein with homology to the vertebrate transcription factor GATA-1. *Development (Cambridge, England)*, *119*(4), 1277-1291.

Reamon-Buettner, S. M. (2005). GATA4 zinc finger mutations as a molecular rationale for septation defects of the human heart. *Journal of Medical Genetics*, *42*(5), e32-e32. https://doi.org/10.1136/jmg.2004.025395

Rehorn, K. P., Thelen, H., Michelson, A. M., & Reuter, R. (1996). A molecular aspect of hematopoiesis and endoderm development common to vertebrates and Drosophila. *Development (Cambridge, England)*, *122*(12), 4023-4031.

Reiter, L. T., Potocki, L., Chien, S., Gribskov, M., & Bier, E. (2001). A Systematic Analysis of Human Disease-Associated Gene Sequences In Drosophila melanogaster. *Genome Research*, *11*(6), 1114-1125. https://doi.org/10.1101/gr.169101

Rekhtman, N., Radparvar, F., Evans, T., & Skoultchi, A. I. (1999). Direct interaction of hematopoietic transcription factors PU.1 and GATA-1 : Functional antagonism in erythroid cells. *Genes & Development*, *13*(11), 1398-1411. https://doi.org/10.1101/gad.13.11.1398

Reuter, R. (1994). The gene serpent has homeotic properties and specifies endoderm versus ectoderm within the Drosophila gut. *Development (Cambridge, England)*, *120*(5), 1123-1135.

Riechmann, V., Rehorn, K. P., Reuter, R., & Leptin, M. (1998). The genetic control of the distinction between fat body and gonadal mesoderm in Drosophila. *Development (Cambridge, England)*, *125*(4), 713-723.

Rivera-Feliciano, J., Lee, K.-H., Kong, S. W., Rajagopal, S., Ma, Q., Springer, Z., Izumo, S., Tabin, C. J., & Pu, W. T. (2006). Development of heart valves requires Gata4 expression in endothelial-derived cells. *Development (Cambridge, England)*, *133*(18), 3607-3618. https://doi.org/10.1242/dev.02519

Rivolta, M. N., & Holley, M. C. (1998). GATA3 is downregulated during hair cell differentiation in the mouse cochlea. *Journal of Neurocytology*, *27*(9), 637-647. https://doi.org/10.1023/A:1006951813063

Rizki, M. T. M., & Rizki, R. M. (1959). Functional Significance of the Crystal Cells in the Larva of Drosophila melanogaster. *The Journal of biophysical and biochemical cytology*. https://doi.org/10.1083/jcb.5.2.235

Romano, O., & Miccio, A. (2020). GATA factor transcriptional activity : Insights from genome-wide binding profiles. *IUBMB Life*, *72*(1), 10-26. https://doi.org/10.1002/iub.2169

Rotstein, B., & Paululat, A. (2016). On the Morphology of the Drosophila Heart. *Journal of Cardiovascular Development and Disease*, *3*(2). https://doi.org/10.3390/jcdd3020015

Rylski, M., Welch, J. J., Chen, Y.-Y., Letting, D. L., Diehl, J. A., Chodosh, L. A., Blobel, G. A., & Weiss, M. J. (2003). GATA-1-Mediated Proliferation Arrest during Erythroid Maturation. *Molecular and Cellular Biology*, *23*(14), 5031-5042. https://doi.org/10.1128/MCB.23.14.5031-5042.2003

Saffman, E. E., & Lasko, P. (1999). Germline development in vertebrates and invertebrates. *Cellular and Molecular Life Sciences: CMLS*, *55*(8-9), 1141-1163. https://doi.org/10.1007/s000180050363

Sam, S., Leise, W., & Keiko Hoshizaki, D. (1996). The serpent gene is necessary for progression through the early stages of fat-body development. *Mechanisms of Development*, *60*(2), 197-205. https://doi.org/10.1016/S0925-4773(96)00615-6

Sandve, S. R., Rohlfs, R. V., & Hvidsten, T. R. (2017). Subfunctionalization versus neofunctionalization after whole-genome duplication. *BioRxiv*, 190090. https://doi.org/10.1101/190090

Santos, A. C., & Lehmann, R. (2004). Germ Cell Specification and Migration in Drosophila and beyond. *Current Biology*, *14*(14), R578-R589. https://doi.org/10.1016/j.cub.2004.07.018

Schleich, J.-M., Abdulla, T., Summers, R., & Houyel, L. (2013). An overview of cardiac morphogenesis. *Archives of Cardiovascular Diseases*, *106*(11), 612-623. https://doi.org/10.1016/j.acvd.2013.07.001

Senger, K., Harris, K., & Levine, M. (2006). GATA factors participate in tissue-specific immune responses in Drosphila larvae. *Proceedings of the National Academy of Sciences of the United States of America*, *103*, 15957-15962. https://doi.org/10.1073/pnas.0607608103

Shim, Y. H., Bonner, J. J., & Blumenthal, T. (1995). Activity of a C. elegans GATA transcription factor, ELT-1, expressed in yeast. *Journal of Molecular Biology*, *253*(5), 665-676. https://doi.org/10.1006/jmbi.1995.0581

Shimizu, R., Ohneda, K., Engel, J. D., Trainor, C. D., & Yamamoto, M. (2004). Transgenic rescue of GATA-1-deficient mice with GATA-1 lacking a FOG-1 association site phenocopies patients with X-linked thrombocytopenia. *Blood*, *103*(7), 2560-2567. https://doi.org/10.1182/blood-2003-07-2514

Shimizu, R., Takahashi, S., Ohneda, K., Engel, J. D., & Yamamoto, M. (2001). In vivo requirements for GATA-1 functional domains during primitive and definitive erythropoiesis. *The EMBO Journal*, *20*(18), 5250-5260. https://doi.org/10.1093/emboj/20.18.5250

Shivdasani, R. A., Fujiwara, Y., McDevitt, M. A., & Orkin, S. H. (1997). A lineage-selective knockout establishes the critical role of transcription factor GATA-1 in megakaryocyte growth and platelet development. *The EMBO Journal*, *16*(13), 3965-3973. https://doi.org/10.1093/emboj/16.13.3965

Si, W., Huang, W., Zheng, Y., Yang, Y., Liu, X., Shan, L., Zhou, X., Wang, Y., Su, D., Gao, J., Yan, R., Han, X., Li, W., He, L., Shi, L., Xuan, C., Liang, J., Sun, L., Wang, Y., & Shang, Y. (2015). Dysfunction of the Reciprocal Feedback Loop between GATA3- and ZEB2-Nucleated Repression Programs Contributes to Breast Cancer Metastasis. *Cancer Cell*, *27*(6), 822-836. https://doi.org/10.1016/j.ccell.2015.04.011

Smith, C. (2003). Hematopoietic stem cells and hematopoiesis. *Cancer Control: Journal of the Moffitt Cancer Center*, *10*(1), 9-16. https://doi.org/10.1177/107327480301000103

Smith, V. M., Lee, P. P., Szychowski, S., & Winoto, A. (1995). GATA-3 Dominant Negative Mutant FUNCTIONAL REDUNDANCY OF THE T CELL RECEPTOR α AND β ENHANCERS. *Journal of Biological Chemistry*, *270*(4), 1515-1520. https://doi.org/10.1074/jbc.270.4.1515

Song, S.-H., Hou, C., & Dean, A. (2007). A positive role for NLI/Ldb1 in long-range beta-globin locus control region function. *Molecular Cell*, *28*(5), 810-822. https://doi.org/10.1016/j.molcel.2007.09.025

Sorrentino, R., Carton, Y., & Govind, S. (2002). Cellular Immune Response to Parasite Infection in the Drosophila Lymph Gland Is Developmentally Regulated. *Developmental biology*, *243*, 65-80. https://doi.org/10.1006/dbio.2001.0542

Sorrentino, R. P., Tokusumi, T., & Schulz, R. A. (2007). The Friend of GATA protein U-shaped functions as a hematopoietic tumor suppressor in Drosophila. *Developmental Biology*, *311*(2), 311-323. https://doi.org/10.1016/j.ydbio.2007.08.011

Spence, J. R., Lauf, R., & Shroyer, N. F. (2011). Vertebrate Intestinal Endoderm Development. *Developmental dynamics : an official publication of the American Association of Anatomists*, *240*(3), 501-520. https://doi.org/10.1002/dvdy.22540

Spilianakis, C. G., & Flavell, R. A. (2004). Long-range intrachromosomal interactions in the T helper type 2 cytokine locus. *Nature Immunology*, *5*(10), 1017-1027. https://doi.org/10.1038/ni1115

Spinner, M. A., Sanchez, L. A., Hsu, A. P., Shaw, P. A., Zerbe, C. S., Calvo, K. R., Arthur, D. C., Gu, W., Gould, C. M., Brewer, C. C., Cowen, E. W., Freeman, A. F., Olivier, K. N., Uzel, G., Zelazny, A. M., Daub, J. R., Spalding, C. D., Claypool, R. J., Giri, N. K., … Holland, S. M. (2014). GATA2 deficiency : A protean disorder of hematopoiesis, lymphatics, and immunity. *Blood*, *123*(6), 809-821. https://doi.org/10.1182/blood-2013-07-515528

St. Johnston, H. L.-S. and D. (2001). Delta signaling from the germ line controls the proliferation and differentiation of the somatic follicle cells during Drosophila oogenesis. *Genes & Development*, *15*(11), 1393-1405. https://doi.org/10.1101/gad.200901

Steger, D. J., Lefterova, M. I., Ying, L., Stonestrom, A. J., Schupp, M., Zhuo, D., Vakoc, A. L., Kim, J.-E., Chen, J., Lazar, M. A., Blobel, G. A., & Vakoc, C. R. (2008). DOT1L/KMT4 recruitment and H3K79 methylation are ubiquitously coupled with gene transcription in mammalian cells. *Molecular and Cellular Biology*, *28*(8), 2825-2839. https://doi.org/10.1128/MCB.02076-07

Stumpf, M. P. H., Thorne, T., de Silva, E., Stewart, R., An, H. J., Lappe, M., & Wiuf, C. (2008). Estimating the size of the human interactome. *Proceedings of the National Academy of Sciences of the United States of America*, *105*(19), 6959-6964. https://doi.org/10.1073/pnas.0708078105

Stumpf, M., Waskow, C., Krötschel, M., van Essen, D., Rodriguez, P., Zhang, X., Guyot, B., Roeder, R. G., & Borggrefe, T. (2006). The mediator complex functions as a coactivator for GATA-1 in erythropoiesis via subunit Med1/TRAP220. *Proceedings of the National*

Academy of Sciences of the United States of America, *103*(49), 18504-18509. https://doi.org/10.1073/pnas.0604494103

Sugiyama, D., Tanaka, M., Kitajima, K., Zheng, J., Yen, H., Murotani, T., Yamatodani, A., & Nakano, T. (2008). Differential context-dependent effects of friend of GATA-1 (FOG-1) on mast-cell development and differentiation. *Blood*, *111*(4), 1924-1932. https://doi.org/10.1182/blood-2007-08-104489

Suzanne, M., Perrimon, N., & Noselli, S. (2001). The Drosophila JNK pathway controls the morphogenesis of the egg dorsal appendages and micropyle. *Developmental Biology*, *237*(2), 282-294. https://doi.org/10.1006/dbio.2001.0384

Svensson, E. C., Tufts, R. L., Polk, C. E., & Leiden, J. M. (1999). Molecular cloning of FOG-2 : A modulator of transcription factor GATA-4 in cardiomyocytes. *Proceedings of the National Academy of Sciences of the United States of America*, *96*(3), 956-961.

Takahashi, S., Onodera, K., Motohashi, H., Suwabe, N., Hayashi, N., Yanai, N., Nabesima, Y., & Yamamoto, M. (1997). Arrest in primitive erythroid cell development caused by promoter-specific disruption of the GATA-1 gene. *The Journal of Biological Chemistry*, *272*(19), 12611-12615. https://doi.org/10.1074/jbc.272.19.12611

Tang, H. (2009). Regulation and function of the melanization reaction in Drosophila. *Fly*, *3*(1), 105-111. https://doi.org/10.4161/fly.3.1.7747

Taylor, J. S., & Raes, J. (2004). Duplication and divergence : The evolution of new genes and old ideas. *Annual Review of Genetics*, *38*, 615-643. https://doi.org/10.1146/annurev.genet.38.072902.092831

Tepass, U., Fessler, L. I., Aziz, A., & Hartenstein, V. (1994). Embryonic origin of hemocytes and their relationship to cell death in Drosophila. *Development (Cambridge, England)*, *120*(7), 1829-1837.

Tevosian, S. G., Albrecht, K. H., Crispino, J. D., Fujiwara, Y., Eicher, E. M., & Orkin, S. H. (2002). Gonadal differentiation, sex determination and normal Sry expression in mice require direct interaction between transcription partners GATA4 and FOG2. *Development (Cambridge, England)*, *129*(19), 4627-4634.

Tevosian, S. G., Deconinck, A. E., Cantor, A. B., Rieff, H. I., Fujiwara, Y., Corfas, G., & Orkin, S. H. (1999). FOG-2 : A novel GATA-family cofactor related to multitype zinc-finger

proteins Friend of GATA-1 and U-shaped. *Proceedings of the National Academy of Sciences, 96*(3), 950-955. https://doi.org/10.1073/pnas.96.3.950

Tokusumi, T., Sorrentino, R. P., Russell, M., Ferrarese, R., Govind, S., & Schulz, R. A. (2009). Characterization of a Lamellocyte Transcriptional Enhancer Located within the misshapen Gene of Drosophila melanogaster. *PLoS ONE, 4*(7). https://doi.org/10.1371/journal.pone.0006429

Tokusumi, Y., Tokusumi, T., & Schulz, R. A. (2018). Mechanical stress to Drosophila larvae stimulates a cellular immune response through the JAK/STAT signaling pathway. *Biochemical and Biophysical Research Communications, 502*(3), 415-421. https://doi.org/10.1016/j.bbrc.2018.05.192

Tokusumi, Y., Tokusumi, T., Stoller-Conrad, J., & Schulz, R. A. (2010). Serpent, Suppressor of Hairless and U-shaped are crucial regulators of hedgehog niche expression and prohemocyte maintenance during Drosophila larval hematopoiesis. *Development (Cambridge, England), 137*(21), 3561-3568. https://doi.org/10.1242/dev.053728

Trainor, C. D., Ghirlando, R., & Simpson, M. A. (2000). GATA zinc finger interactions modulate DNA binding and transactivation. *The Journal of Biological Chemistry, 275*(36), 28157-28166. https://doi.org/10.1074/jbc.M000020200

Trainor, C D, Omichinski, J. G., Vandergon, T. L., Gronenborn, A. M., Clore, G. M., & Felsenfeld, G. (1996). A palindromic regulatory site within vertebrate GATA-1 promoters requires both zinc fingers of the GATA-1 DNA-binding domain for high-affinity interaction. *Molecular and Cellular Biology, 16*(5), 2238-2247.

Tremblay, K. D., & Zaret, K. S. (2005). Distinct populations of endoderm cells converge to generate the embryonic liver bud and ventral foregut tissues. *Developmental Biology, 280*(1), 87-99. https://doi.org/10.1016/j.ydbio.2005.01.003

Tremblay, M., Sanchez-Ferras, O., & Bouchard, M. (2018). GATA transcription factors in development and disease. *Development, 145*(20), dev164384. https://doi.org/10.1242/dev.164384

Tsai, F. Y., & Orkin, S. H. (1997). Transcription factor GATA-2 is required for proliferation/survival of early hematopoietic cells and mast cell formation, but not for erythroid and myeloid terminal differentiation. *Blood, 89*(10), 3636-3643.

Tsai, F.-Y., Keller, G., Kuo, F. C., Weiss, M., Chen, J., Rosenblatt, M., Alt, F. W., & Orkin, S. H. (1994). An early haematopoietic defect in mice lacking the transcription factor GATA-2. *Nature*, *371*(6494), 221-226. https://doi.org/10.1038/371221a0

Tsang, A. P., Visvader, J. E., Turner, C. A., Fujiwara, Y., Yu, C., Weiss, M. J., Crossley, M., & Orkin, S. H. (1997). FOG, a multitype zinc finger protein, acts as a cofactor for transcription factor GATA-1 in erythroid and megakaryocytic differentiation. *Cell*, *90*(1), 109-119. https://doi.org/10.1016/s0092-8674(00)80318-9

Tubman, V. N., Levine, J. E., Campagna, D. R., Monahan-Earley, R., Dvorak, A. M., Neufeld, E. J., & Fleming, M. D. (2007). X-linked gray platelet syndrome due to a GATA1 Arg216Gln mutation. *Blood*, *109*(8), 3297-3299. https://doi.org/10.1182/blood-2006-02-004101

Twombly, V., Blackman, R. K., Jin, H., Graff, J. M., Padgett, R. W., & Gelbart, W. M. (1996). The TGF-beta signaling pathway is essential for Drosophila oogenesis. *Development (Cambridge, England)*, *122*(5), 1555-1565.

Usary, J., Llaca, V., Karaca, G., Presswala, S., Karaca, M., He, X., Langerød, A., Kåresen, R., Oh, D. S., Dressler, L. G., Lønning, P. E., Strausberg, R. L., Chanock, S., Børresen-Dale, A.-L., & Perou, C. M. (2004). Mutation of GATA3 in human breast tumors. *Oncogene*, *23*(46), 7669-7678. https://doi.org/10.1038/sj.onc.1207966

Vakoc, C. R., Letting, D. L., Gheldof, N., Sawado, T., Bender, M. A., Groudine, M., Weiss, M. J., Dekker, J., & Blobel, G. A. (2005). Proximity among distant regulatory elements at the beta-globin locus requires GATA-1 and FOG-1. *Molecular Cell*, *17*(3), 453-462. https://doi.org/10.1016/j.molcel.2004.12.028

Vecchio, G. C. D., Giordani, L., Santis, A. D., & Mattia, D. D. (2005). Dyserythropoietic Anemia and Thrombocytopenia due to a Novel Mutation in GATA-1. *Acta Haematologica*, *114*(2), 113-116. https://doi.org/10.1159/000086586

Viger, R. S., Mertineit, C., Trasler, J. M., & Nemer, M. (1998). Transcription factor GATA-4 is expressed in a sexually dimorphic pattern during mouse gonadal development and is a potent activator of the Müllerian inhibiting substance promoter. *Development (Cambridge, England)*, *125*(14), 2665-2675.

Vogler, G., & Bodmer, R. (2015). Cellular Mechanisms of Drosophila Heart Morphogenesis. *Journal of Cardiovascular Development and Disease*, *2*(1), 2-16. https://doi.org/10.3390/jcdd2010002

Vyas, P., Ault, K., Jackson, C. W., Orkin, S. H., & Shivdasani, R. A. (1999). Consequences of GATA-1 Deficiency in Megakaryocytes and Platelets. *Blood*, *93*(9), 2867-2875. https://doi.org/10.1182/blood.V93.9.2867

Wada, H., Hasegawa, K., Morimoto, T., Kakita, T., Yanazume, T., & Sasayama, S. (2000). A p300 protein as a coactivator of GATA-6 in the transcription of the smooth muscle-myosin heavy chain gene. *The Journal of Biological Chemistry*, *275*(33), 25330-25335. https://doi.org/10.1074/jbc.M000828200

Wadman, I. A., Osada, H., Grütz, G. G., Agulnick, A. D., Westphal, H., Forster, A., & Rabbitts, T. H. (1997). The LIM-only protein Lmo2 is a bridging molecule assembling an erythroid, DNA-binding complex which includes the TAL1, E47, GATA-1 and Ldb1/NLI proteins. *The EMBO Journal*, *16*(11), 3145-3157. https://doi.org/10.1093/emboj/16.11.3145

Walker, E. M., Thompson, C. A., & Battle, M. A. (2014). GATA4 and GATA6 regulate intestinal epithelial cytodifferentiation during development. *Developmental Biology*, *392*(2), 283-294. https://doi.org/10.1016/j.ydbio.2014.05.017

Waltzer, L., Bataillé, L., Peyrefitte, S., & Haenlin, M. (2002). Two isoforms of Serpent containing either one or two GATA zinc fingers have different roles in Drosophila haematopoiesis. *The EMBO Journal*, *21*(20), 5477-5486. https://doi.org/10.1093/emboj/cdf545

Waltzer, L., Ferjoux, G., Bataillé, L., & Haenlin, M. (2003). Cooperation between the GATA and RUNX factors Serpent and Lozenge during Drosophila hematopoiesis. *The EMBO Journal*, *22*(24), 6516-6525. https://doi.org/10.1093/emboj/cdg622

Watt, A. J., Zhao, R., Li, J., & Duncan, S. A. (2007). Development of the mammalian liver and ventral pancreas is dependent on GATA4. *BMC Developmental Biology*, *7*, 37. https://doi.org/10.1186/1471-213X-7-37

Weiss, M. J., & Orkin, S. H. (1995). Transcription factor GATA-1 permits survival and maturation of erythroid precursors by preventing apoptosis. *Proceedings of the National Academy of Sciences of the United States of America*, *92*(21), 9623-9627.

Welch, J. J., Watts, J. A., Vakoc, C. R., Yao, Y., Wang, H., Hardison, R. C., Blobel, G. A., Chodosh, L. A., & Weiss, M. J. (2004). Global regulation of erythroid gene expression by transcription factor GATA-1. *Blood*, *104*(10), 3136-3147. https://doi.org/10.1182/blood-2004-04-1603

Wilkinson-White, L., Gamsjaeger, R., Dastmalchi, S., Wienert, B., Stokes, P. H., Crossley, M., Mackay, J. P., & Matthews, J. M. (2011). Structural basis of simultaneous recruitment of the transcriptional regulators LMO2 and FOG1/ZFPM1 by the transcription factor GATA1. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(35), 14443-14448. https://doi.org/10.1073/pnas.1105898108

Winick, J., Abel, T., Leonard, M. W., Michelson, A. M., Chardon-Loriaux, I., Holmgren, R. A., Maniatis, T., & Engel, J. D. (1993). A GATA family transcription factor is expressed along the embryonic dorsoventral axis in Drosophila melanogaster. *Development (Cambridge, England)*, *119*(4), 1055-1065.

Wolf, M. J. (2012). Modeling Dilated Cardiomyopathies in Drosophila. *Trends in cardiovascular medicine*, *22*(3), 55-61. https://doi.org/10.1016/j.tcm.2012.06.012

Wood, W., & Jacinto, A. (2007). Drosophila melanogaster embryonic haemocytes : Masters of multitasking. *Nature Reviews Molecular Cell Biology*, *8*(7), 542-551. https://doi.org/10.1038/nrm2202

Wu, J., Ding, X., & Wang, J. (2016). Stem Cells in Mammalian Gonads. In R. P. Piprek (Éd.), *Molecular Mechanisms of Cell Differentiation in Gonad Development* (p. 289-307). Springer International Publishing. https://doi.org/10.1007/978-3-319-31973-5_11

Wu, X., Tanwar, P. S., & Raftery, L. A. (2008). Drosophila follicle cells : Morphogenesis in an eggshell. *Seminars in cell & developmental biology*, *19*(3), 271-282. https://doi.org/10.1016/j.semcdb.2008.01.004

Yamamoto, M., Ko, L. J., Leonard, M. W., Beug, H., Orkin, S. H., & Engel, J. D. (1990). Activity and tissue-specific expression of the transcription factor NF-E1 multigene family. *Genes & Development*, *4*(10), 1650-1662. https://doi.org/10.1101/gad.4.10.1650

Yang, H. Y., & Evans, T. (1992). Distinct roles for the two cGATA-1 finger domains. *Molecular and Cellular Biology*, *12*(10), 4562-4570. https://doi.org/10.1128/mcb.12.10.4562

Yu, C., Niakan, K. K., Matsushita, M., Stamatoyannopoulos, G., Orkin, S. H., & Raskind, W. H. (2002). X-linked thrombocytopenia with thalassemia from a mutation in the amino finger of GATA-1 affecting DNA binding rather than FOG-1 interaction. *Blood*, *100*(6), 2040-2045. https://doi.org/10.1182/blood-2002-02-0387

Yu, L., Wynn, J., Cheung, Y. H., Shen, Y., Mychaliska, G. B., Crombleholme, T. M., Azarow, K. S., Lim, F. Y., Chung, D. H., Potoka, D., Warner, B. W., Bucher, B., Stolar, C., Aspelund, G., Arkovitz, M. S., & Chung, W. K. (2013). Variants in GATA4 Are a Rare Cause of Familial and Sporadic Congenital Diaphragmatic Hernia. *Human genetics*, *132*(3), 285-292. https://doi.org/10.1007/s00439-012-1249-0

Yu, Min, Wang, U., & Wang, Z. (2016). E2F and GATA switches turn off WD repeat domain 77 expression in differentiating cells. *The Biochemical Journal*, *473*(15), 2331-2343. https://doi.org/10.1042/BCJ20160130

Yu, Ming, Riva, L., Xie, H., Schindler, Y., Moran, T. B., Cheng, Y., Yu, D., Hardison, R., Weiss, M. J., Orkin, S. H., Bernstein, B. E., Fraenkel, E., & Cantor, A. B. (2009). Insights into GATA-1-mediated gene activation versus repression via genome-wide chromatin occupancy analysis. *Molecular Cell*, *36*(4), 682-695. https://doi.org/10.1016/j.molcel.2009.11.002

Zaffran, S. (2003). Early Cardiac Development in Vertebrates and Invertebrates : Transcription Factors and Signaling Processes. *Proceedings of Indian National Science Academy*, *5*, 703-718.

Zaret, K. S. (2002). Regulatory phases of early liver development : Paradigms of organogenesis. *Nature Reviews Genetics*, *3*(7), 499-512. https://doi.org/10.1038/nrg837

Zaret, K. S., & Carroll, J. S. (2011). Pioneer transcription factors : Establishing competence for gene expression. *Genes & Development*, *25*(21), 2227-2241. https://doi.org/10.1101/gad.176826.111

Zhang, J., Ding, L., Holmfeldt, L., Wu, G., Heatley, S. L., Payne-Turner, D., Easton, J., Chen, X., Wang, J., Rusch, M., Lu, C., Chen, S.-C., Wei, L., Collins-Underwood, J. R., Ma, J., Roberts, K. G., Pounds, S. B., Ulyanov, A., Becksfort, J., … Mullighan, C. G. (2012). The genetic basis of early T-cell precursor acute lymphoblastic leukaemia. *Nature*, *481*(7380), 157-163. https://doi.org/10.1038/nature10725

Zhang, P., Behre, G., Pan, J., Iwama, A., Wara-Aswapati, N., Radomska, H. S., Auron, P. E., Tenen, D. G., & Sun, Z. (1999). Negative cross-talk between hematopoietic regulators : GATA proteins repress PU.1. *Proceedings of the National Academy of Sciences of the United States of America*, *96*(15), 8705-8710. https://doi.org/10.1073/pnas.96.15.8705

Zhang, S.-J., Ma, L.-Y., Huang, Q.-H., Li, G., Gu, B.-W., Gao, X.-D., Shi, J.-Y., Wang, Y.-Y., Gao, L., Cai, X., Ren, R.-B., Zhu, J., Chen, Z., & Chen, S.-J. (2008). Gain-of-function mutation of GATA-2 in acute myeloid transformation of chronic myeloid leukemia. *Proceedings of the National Academy of Sciences of the United States of America*, *105*(6), 2076-2081. https://doi.org/10.1073/pnas.0711824105

Zhang, Y., & Xi, Y. (2015). Fat Body Development and its Function in Energy Storage and Nutrient Sensing in Drosophila melanogaster. *Journal of Tissue Science and Engineering*, *6*(1), 1-8. https://doi.org/10.4172/2157-7552.1000141

Zhao, R., Watt, A. J., Battle, M. A., Li, J., Bondow, B. J., & Duncan, S. A. (2008). Loss of both GATA4 and GATA6 blocks cardiac myocyte differentiation and results in acardia in mice. *Developmental Biology*, *317*(2), 614-619. https://doi.org/10.1016/j.ydbio.2008.03.013

Zhao, R., Watt, A. J., Li, J., Luebke-Wheeler, J., Morrisey, E. E., & Duncan, S. A. (2005). GATA6 is essential for embryonic development of the liver but dispensable for early heart formation. *Molecular and Cellular Biology*, *25*(7), 2622-2631. https://doi.org/10.1128/MCB.25.7.2622-2631.2005

Zhu, W., Shiojima, I., Hiroi, Y., Zou, Y., Akazawa, H., Mizukami, M., Toko, H., Yazaki, Y., Nagai, R., & Komuro, I. (2000). Functional analyses of three Csx/Nkx-2.5 mutations that cause human congenital heart disease. *The Journal of Biological Chemistry*, *275*(45), 35291-35296. https://doi.org/10.1074/jbc.M000525200