



**HAL**  
open science

# Spatio-temporal control of muscle diversity in *Drosophila* leg

Camille Guillermin

► **To cite this version:**

Camille Guillermin. Spatio-temporal control of muscle diversity in *Drosophila* leg. *Development Biology*. Ecole normale supérieure de lyon - ENS LYON, 2023. English. NNT : 2023ENSL0103 . tel-04692677

**HAL Id: tel-04692677**

**<https://theses.hal.science/tel-04692677>**

Submitted on 10 Sep 2024

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

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



## THESE

en vue de l'obtention du grade de Docteur, délivré par  
l'ECOLE NORMALE SUPERIEURE DE LYON

**Ecole Doctorale N° 340**  
Biologie Moléculaire Intégrative et Cellulaire

### **Discipline :**

Sciences de la vie et de la santé

Soutenue publiquement le 11/12/2023, par :

**Camille GUILLERMIN**

---

# **Spatio-temporal control of muscle diversity in *Drosophila* leg**

---

Devant le jury composé de :

**SOLER, Cédric** / CR1 MCU / Laboratoire de Génétique Reproduction et Développement de Clermont-Ferrand / Rapporteur

**BOUKHATMI, Hadi** / CR / Institut de Génétique et Développement de Rennes / Rapporteur

**RUGGIERO, Florence** / DR1 / Institut de Génomique Fonctionnelle de Lyon / Examinatrice

**DAVY, Alice** / DR / Centre de Biologie Intégrative de Toulouse / Examinatrice

**LEGRAND, Fabien** / DR2 / Institut NeuroMyoGène de Lyon / Examineur

**ENRIQUEZ, Jonathan** / CR1 / Institut de Génomique Fonctionnelle de Lyon / Directeur de thèse



## TABLE OF CONTENTS

<b>TABLE OF CONTENTS</b> .....	<b>2</b>
<b>REMERCIEMENTS</b> .....	<b>6</b>
<b>RESUMÉ</b> .....	<b>10</b>
<b>ABSTRACT</b> .....	<b>12</b>
<b>ABBREVIATIONS</b> .....	<b>14</b>
<b>INTRODUCTION</b> .....	<b>20</b>
A. Vertebrate muscles.....	20
A.1. General organization of vertebrate muscles .....	20
A.2. Muscle development.....	22
A.2.a. Overview of muscle development .....	22
A.2.b. Somitogenesis.....	24
A.2.c. Compartmentalization of the dermomyotome.....	26
A.2.d. Development of limb muscles .....	28
A.2.d.1. Cellular description of limb muscle development .....	28
A.2.d.1.1. Delamination and migration .....	30
A.2.d.1.2. Proliferation .....	30
A.2.d.1.3. Determination and differentiation.....	32
A.2.d.1.4. Fusion.....	32
A.2.d.1.5. Second wave of myogenesis .....	34
A.2.d.2. Molecular description of limb muscle development .....	34
A.2.d.2.1. Specification of muscles precursors.....	34
A.2.d.2.2. Delamination of muscle precursors .....	36
A.2.d.2.3. Migration of muscle precursors .....	38
A.2.d.2.4. Proliferation of muscle precursors.....	40
A.2.d.2.5. Determination and differentiation.....	40
A.3. Generation of muscles morphological diversity .....	42
B. <i>Drosophila melanogaster</i> .....	46
B.1. The larval body wall muscles .....	48





B.1.a. Organization of the larval body wall muscles .....	48
B.1.b. Larvae body wall muscles development .....	50
B.1.b.1. Cellular basis of larvae myogenesis .....	50
B.1.b.2. Molecular basis of larvae myogenesis .....	52
B.1.b.2.1. Segmentation of the mesoderm .....	52
B.1.b.2.2. Molecular regulation of larvae myogenesis .....	54
B.1.b.2.2.1. Specification of the promuscular clusters and selection of progenitors.....	56
B.1.b.2.2.2. Formation of founder cells .....	56
B.1.b.2.2.3. Fusion of founder cells with fusion competent myoblasts .....	56
B.2. The adult leg muscles .....	58
B.2.b. Generality about adult leg myogenesis .....	58
B.2.c. The immature leg .....	60
B.2.d. Patterning of the leg disc .....	62
B.2.d.1. The antero-posterior axis of the leg disc .....	62
B.2.d.2. The Dorso-ventral axis of the leg disc .....	62
B.2.d.3. The proximo-distal axis of the leg disc.....	64
B.3. Generation of muscle identity .....	66
B.3.a. The intra-segmental identity.....	66
B.3.b. The inter-segmental identity .....	68
<b>THESIS PROJECT .....</b>	<b>72</b>
<b>PAPER IN PREPARATION .....</b>	<b>78</b>
<b>DISCUSSION .....</b>	<b>190</b>
A. Control of muscle morphologies development .....	190
A.1. Early specification of myoblasts .....	190
A.2. Control of the spatio-temporal organization of myoblasts .....	194
B. Development of the muscle innervation .....	196
C. Maintenance of the adult leg architecture.....	202
D. CONCLUSION .....	204
<b>BIBLIOGRAPHY.....</b>	<b>206</b>



## REMERCIEMENTS

Au cours de ces dernières années, de mon stage de Master 1 à la soutenance de cette thèse j'ai eu la chance d'être entourée et soutenue par de nombreuses personnes que je ne remerciais jamais assez.

Je tiens dans un premier temps à remercier Dr Cédric Soler et Dr Hadi Boukhatmi qui ont accepté d'être les rapporteurs de ma thèse. Je les remercie pour leur expertise et leurs retours sur mon manuscrit. Je remercie aussi Dr Florence Ruggiero, Dr Alice Davy et Dr Fabien Legrand d'avoir acceptés de faire partie de mon jury de thèse et de leur expertise scientifique sur mes travaux. Je tiens aussi à remercier Dr Stéphane Vincent et Dr Alain Vincent d'avoir fait partie de mon Comité de Suivi de Thèse, de leurs retours et leurs conseils avisés.

Jon, je crois que je ne trouverais jamais les bons mots pour la chance que tu m'as donné dès mon master et de me permettre de réaliser ma thèse dans ton équipe. Je ne te remerciais jamais de toute l'aide que tu m'as apporté, professionnellement et personnellement. Être ta première thésarde est un honneur et je suis désolé pour les cheveux blancs... Tu m'as aidé à prendre confiance en moi, à avoir un regard critique et de devenir une meilleure scientifique chaque jour ! Merci de ton soutien, ta persévérance envers moi et surtout de ta grande patience... ! Tu es quelqu'un d'authentique, de sincère et j'ai eu une grande chance d'avoir pu réaliser ma thèse avec toi.

Anne, tu as toujours été là pour m'écouter, me soutenir, me conseiller ! Je ne saurais te remercier assez avec du chocolat et des cookies à foisons pour ce que tu as fait pour moi. Tu m'as aidé à « respirer », à méditer et prendre du recul. Merci pour toute ton énergie et ta présence au quotidien.

Mathilde, merci pour les vacances, les repos, les chocolats, les aides, les ... je n'aurai pas les mots une fois de plus pour tout ce que tu as fait ! Tu es devenue une deuxième maman ! Ne change jamais, tu es parfaitement parfaite dans ton soutien infaillible, ton écoute, ton aide pro oui mais surtout personnelle ! Merci et encore merci !

Violaine, tu es arrivée pour ma dernière année mais c'est comme si tu avais toujours été là ! Je ne sais même pas comment j'aurai pu finir cette année sans toi. Oui tu m'as aidé pour les expériences mais tu as surtout été là pour me soutenir, m'écouter, tu es une troisième maman. Merci !

Juliette, je n'ai qu'une chose à dire : Samedi 13h ! Merci pour tout, ton travail, ton soutien, ton aide, ton écoute, ton expertise en pioupiou ! Tu complètes le groupe du guet-apens qui permet la survie de tous !

Manuela, Ziyang, courage pour la fin de vos thèses respectives ! Essayer de ne pas trop abimer Jon et Anne, je crois que je leur ai déjà donné beaucoup top de fil à retordre...

Imane, Stéphanie, Wenyue, vous étiez là les premières années et sans vous tout aurait été différent ! Merci pour votre aide scientifique et votre soutien infaillible. Imane et nos discussions sur oui ou non le « screen » a apporté de nouveaux candidats. Stéphanie pour toute ton aide informatique..., pour ton soutien et ton amitié. Wenyue merci d'avoir été ma



première encadrante lors de mon master, de m'avoir appris tout ce qu'il fallait savoir sur les drosophiles !

Je tiens aussi à remercier tous les membres de l'IGFL. Le collège doc-post doc qui apporte un soutien réel dans l'épreuve d'une thèse. Béryl, Amélie merci pour vos nombreux conseils et votre écoute, je n'oublierais jamais cela ! Sébastien, Benjamin, Sandrine merci pour votre aide et votre expertise. Sergio, Laurent merci de m'avoir sauvé avec ces nombreuses data, sans vous la thèse aurait un sens tout différent ! Merci à toute l'équipe administrative pour leur aide et leur patience, merci Fabienne, merci Martine. Enfin merci à toutes les équipes, chercheurs, ingénieurs pour leurs aides scientifiques, leurs conseils et leurs encouragements.

Dans un contexte plus personnel je remercie amis et famille qui m'ont apporté leur amour, soutien, bulles d'oxygène. Merci pour les sorties cinés, les randonnées, les vacances, les weekends, votre écoute, votre soutien, et tout simplement votre présence.

Auriane, je ne te le dirais jamais assez tu es comme ma sœur, merci pour tout, je sais que je pourrais toujours compter sur toi et je n'aurai jamais assez de mots pour te remercier. Julie, merci pour ton soutien, ton aide, ton amitié, ta présence. Adrien, merci pour ton écoute, les brunchs, les cinés, les ralleries, les vacances, toi. Marine, je pense que je ne pouvais pas rêver d'une meilleure colocation pour ces dernières années de thèse. Merci de ton soutien, de ces fous rires, de cette connexion, de ton aide au quotidien pour me permettre d'avoir réalisé cette thèse dans les meilleures conditions.

Enfin, je n'aurai jamais les mots assez forts pour eux, ma famille, mes parents, mon frère, ma belle-sœur, mes amours de nièce et neveu. Je vous aime. Je suis aujourd'hui la personne que je suis grâce à vous. Votre amour inconditionnel est ce qui me permet d'avancer, de tenir tout le long, de persévérer, d'être qui je suis. Cette thèse n'aurait jamais été possible sans votre aide, soutien, amour !

Et oui... Je ne pourrais écrire ces remerciements sans mentionner ce qui a changé ma vie : le Crossfit, ma deuxième (troisième ?) maison, Crossfit Gerland ! Une passion qui m'a donné confiance en moi, persévérance, force mentale et surtout une communauté qui m'a apporté son soutien et plus encore. Merci les copains, les APA ! Merci pour tout, les apéros, les PR, les compète, l'eau de bourgogne ! Merci à tous les amis de CFG qui ont été là quand il le fallait, qui ont cru en moi, qui m'ont encouragé et soutenu tout du long. Cette thèse fut le plus long WOD d'une vie je pense.

A mes parents



## RESUMÉ

La locomotion se traduit par un ensemble de mouvements indispensables pour la survie de l'espèce. La précision de ces mouvements est intrinsèquement liée à la morphologie unique des muscles. *Drosophila melanogaster* se distingue comme un organisme modèle remarquable pour disséquer les complexités du développement des morphologies musculaires en offrant une analogie structurale et développementale aux membres supérieurs et inférieurs des vertébrés. Alors que le programme moléculaire général, orchestrant le développement musculaire est bien élucidé, le mystère demeure quant au programme spécifique qui gouverne la diversité de ces morphologies. C'est autour de cette problématique que s'articule ma thèse : existe-t-il, en parallèle du programme général de la myogenèse, un programme spécifique qui dicte la mise en place de la morphologie unique de chaque muscle ? Au cours de ma thèse, j'ai réalisé une découverte majeure concernant le programme génétique sous-tendant la diversité morphologique des muscles. J'ai utilisé diverses approches, combinant génétique, microscopie confocale, et séquençage de cellules uniques. Mes résultats ont révélé que les myoblastes possèdent initialement un état naïf lorsqu'ils rejoignent les cellules épithéliales au sein de la patte immature (disque imaginal). Au début du développement, ils se spécifient en deux sous-populations : l'une, localisée au centre du disque, à l'origine des muscles distaux, tandis que l'autre, en périphérie, donnera les muscles proximaux. Ultérieurement, les myoblastes voués à former le même muscle se spécialisent et se regroupent spatialement dans la région préfigurant la position finale du muscle. Les analyses transcriptomiques à différents stades ont révélé des gènes candidats semblant jouer un rôle clé dans ce processus de spécification. J'ai pu assigner des codes de facteurs de transcriptions spécifiques aux différentes populations de myoblastes évoluant au cours du temps et commencé à analyser leur rôle fonctionnel dans la formation du système musculaire. Ce projet mènera à de nouveaux concepts biologiques qui accroîtront nos connaissances fondamentales sur la biologie du développement.





## ABSTRACT

Locomotion is a fundamental behavior in animals, crucial for finding food, mates, and avoiding predators. In appendages, precise movement relies on the unique muscle morphologies. *Drosophila melanogaster* stands as a remarkable model organism for untangling the complexities of muscle morphologies development. It offers a structural and developmental analogous model to vertebrates upper and lower limbs. The general molecular program of muscle development is well characterized in flies and vertebrates, but the specific program controlling muscle morphology diversity in appendages remains unknown. My thesis is organized around one problematic: Is there, in parallel with the general myogenesis program, a specific muscle program that governs unique muscle morphologies? During my thesis, I made a significant discovery regarding the genetic program that governs the morphology diversity of adult muscles. I used multiple approaches integrating genetics, confocal microscopy, single-cell sequencing, image post-processing and bioinformatic. My research findings reveal that myoblasts initially possess a naive state upon joining the epithelial cells of the leg disc (immature leg). During early stages, myoblasts are specified into two subpopulations. One population, located in the center of the disc, will give rise to distal muscles, while the other one, at the periphery, will generate proximal muscles. Then at later stages, myoblasts giving rise to the same muscle are specified and spatially cluster together in the region prefiguring the position of the muscle they will form. The analysis of the single cell sequencing perform at different stages highlights the potential genes controlling this specification. I was able to assign specific transcription factor codes, evolving over time, to the different myoblasts' populations, and begin to analyze their functional role in shaping muscle architecture. This project has the potential to contribute novel knowledge and understanding of the intricate cellular and molecular mechanisms controlling muscle diversity.



## **ABBREVIATIONS**

A – Abdominal segments

Abd – Abdominal

AEL – After Egg Laying

AMP - Adult Muscle Precursor

Antp – Antennapedia

APF – After Pupal Formation

ap – apterous

bi – bifid

BMP – Bone Morphogenetic Protein

brk – brinker

cDM – central dermomyotome

Dac – Dachshund

DIP-alpha – Dpr-interacting protein alpha

Dll – Distal-less

DML – Dorsal Medial Lip

Doc1 – Dorsocross 1

Doc2 – Dorsocross 2

Dpp – Decapentaplegic

Dpr10 – Defective proboscis extension response 10

ECM – Extracellular Matrix

EMT – Epithelial-to-Mesenchymal Transition

en – engrailed

eve – even-skipped



FC – Founder Cells

FCM – Fusion Competent Myoblasts

FGF – Fibroblast Growth Factor

HGF – Hepatocyte Growth Factor

Hh – Hedgehog

hth – homothorax

IGFs – Insulin-like Growth Factors

ko – knockout

lbe – ladybird early

Lbx1 – Ladybird Homeobox 1

lim1 – LIM homeobox 1

MAPK – Mitogen Activated Protein Kinase

MCK – Muscle Creatinine Kinase

MCT – Muscle Connective Tissue

Mef2 – Myocyte Enhancer Factor 2

MERFISH – Multiplexed Error-Robust Fluorescence In Situ Hybridization

Mhc – Myosin heavy chain

mid – midline

MRF – Myogenic Regulatory Factor

Msh – Muscle segment homeobox

Msx1 – Muscle segment homeobox 1

mTF – morphological Transcription Factor

Myf-5 – Myogenic factor 5

MyoD – Myogenic Differentiation 1

PAGA – Partition-based Graph Abstraction



PI3K/Akt – phosphoinositide 3-kinase/protein kinase B

Pitx2 – Pituitary homeobox 2

Scr – Sex combs reduced

SF – Scatter Factor

Shh – Sonic Hedgehog

slp – sloppy paired

T – Thoracic segment

TF – Transcription Factor

TGF- $\beta$  – transforming Growth Factor-beta

tilm – tibia levator muscle

tnc – tenectin

twi – twist

Ubx – Ultrabithorax

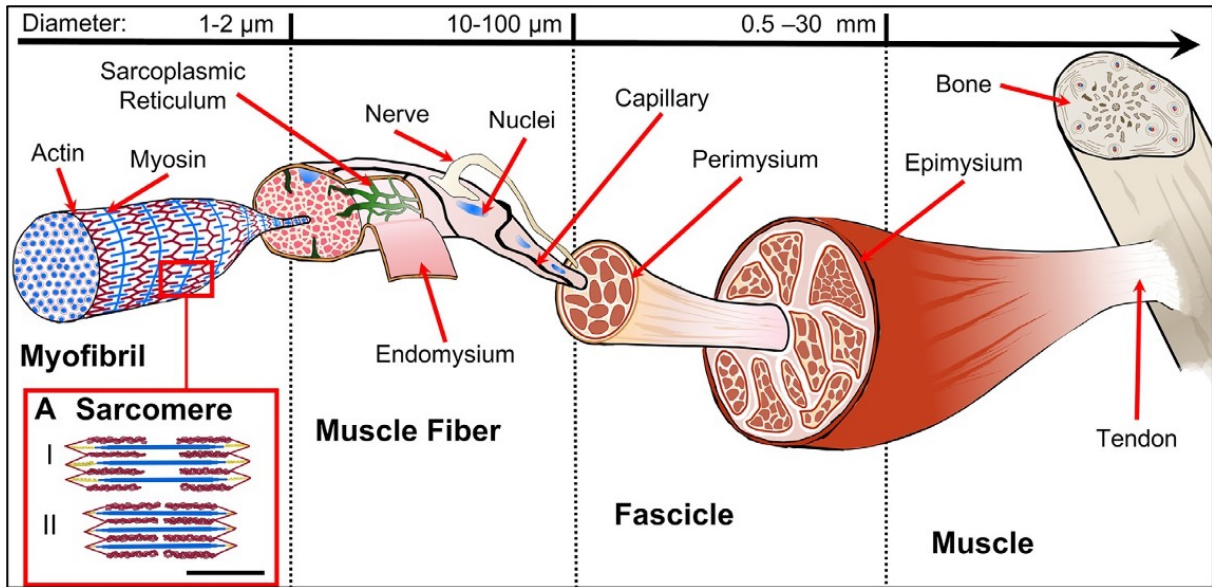
unpg – unplugged

VLL – Ventral Lateral Lip

wg – wingless

zfh1 – Zn finger homeodomain 1





**Figure 1 - Skeletal muscle structure**

Schematic representation of the skeletal muscle tissue components. Each muscle has three layers of connective tissue; epimysium, perimysium and endomysium. Myofibers are composed of multiple myofibrils and are bundled together forming muscle fascicle.

(A) Sarcomere morphology and sliding: Actin (red), Myosin (blue) and Titin (yellow) filaments are shown in the relaxed state (I) and during the contraction (II) (Scale bar 0.5 nm).

Figure from: Gotti, et al., 2020

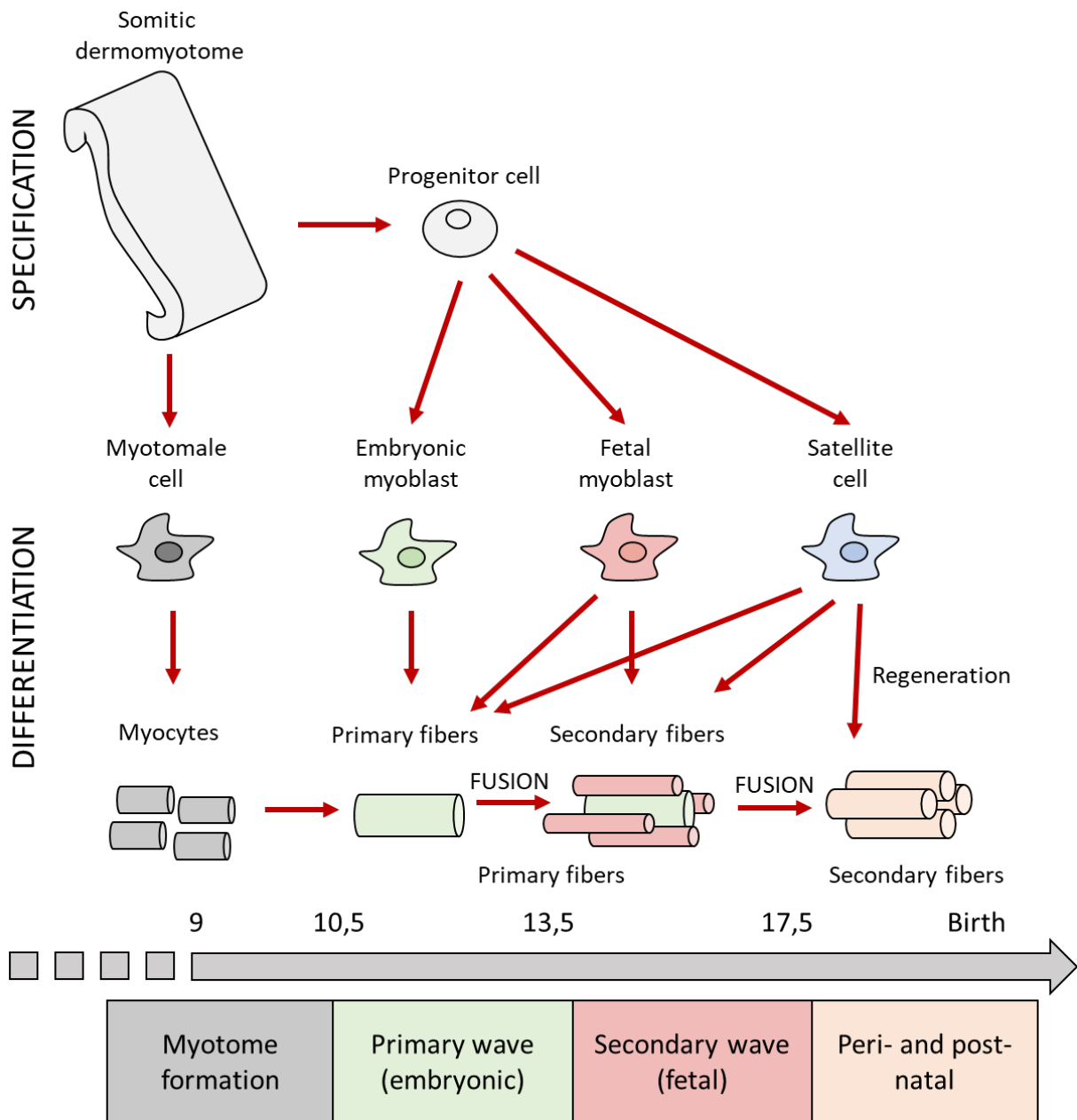
## INTRODUCTION

### A. Vertebrate muscles

#### A.1. General organization of vertebrate muscles

The general organization of a muscle is a complex and hierarchical structure that facilitates movement and contraction. Muscles are composed of muscle fibers, connective tissue, blood vessels, and nerves, and they are linked to the skeleton through tendons. This organization, from the whole muscle down to individual muscle fibers, ensures efficient force generation, coordination, and control of movement throughout the body. Moreover, different types of muscle fibers exist, classified based on their contractile properties and metabolic characteristics. Slow-twitch fibers (Type I) contract slowly, have high endurance, and rely on aerobic metabolism (oxygen-dependent) for energy production. They are suitable for sustained, low-intensity activities like endurance exercises. In contrast, fast-twitch fibers (Type II) contract rapidly, produce high force, and rely on anaerobic metabolism (oxygen-independent) for energy production. Fast-twitch fibers can be further categorized into Type IIa (fast oxidative) and Type IIb (fast glycolytic) fibers, each with unique characteristics (Schiaffino and Reggiani, 2011).

At the macroscopic level, the entire muscle is surrounded by a layer of connective tissue called the epimysium, providing structural support and protection (**Figure 1**, Gotti et al., 2020). Within the muscle, there are smaller bundles of muscle fibers known as fascicles, each surrounded by a layer of connective tissue called the perimysium. The perimysium contains blood vessels and nerves that supply the fascicles with oxygen, nutrients, and nerve signals. At the microscopic level, each fascicle contains individual muscle fibers. Vertebrate skeletal muscles are tubular and striated multi-fiber muscles. Muscle fibers are elongated, striated and multinucleated cells running in parallel to the muscle's length (Spletter and Schnorrer, 2014). They are surrounded by a layer of connective tissue called the endomysium, which provides structural support to the individual muscle fibers. Inside each muscle fiber, myofibrils are responsible for the contractile properties of the muscle (Scime et al., 2009). Myofibrils consist



**Figure 2 - Vertebrate myogenesis**

Schematic representation of muscle formation in vertebrates (*Timeline: mouse developmental time (days post conception)*). There is two waves of myogenesis: the primary wave, embryonic myogenesis, serves as the foundation for early muscle development with the primary fibers, while the secondary wave, fetal myogenesis, adds to muscle size, complexity, and adaptability, shaping the final architecture of the musculature with the secondary fibers.

*Adapted from Rossi and Messina, 2014*

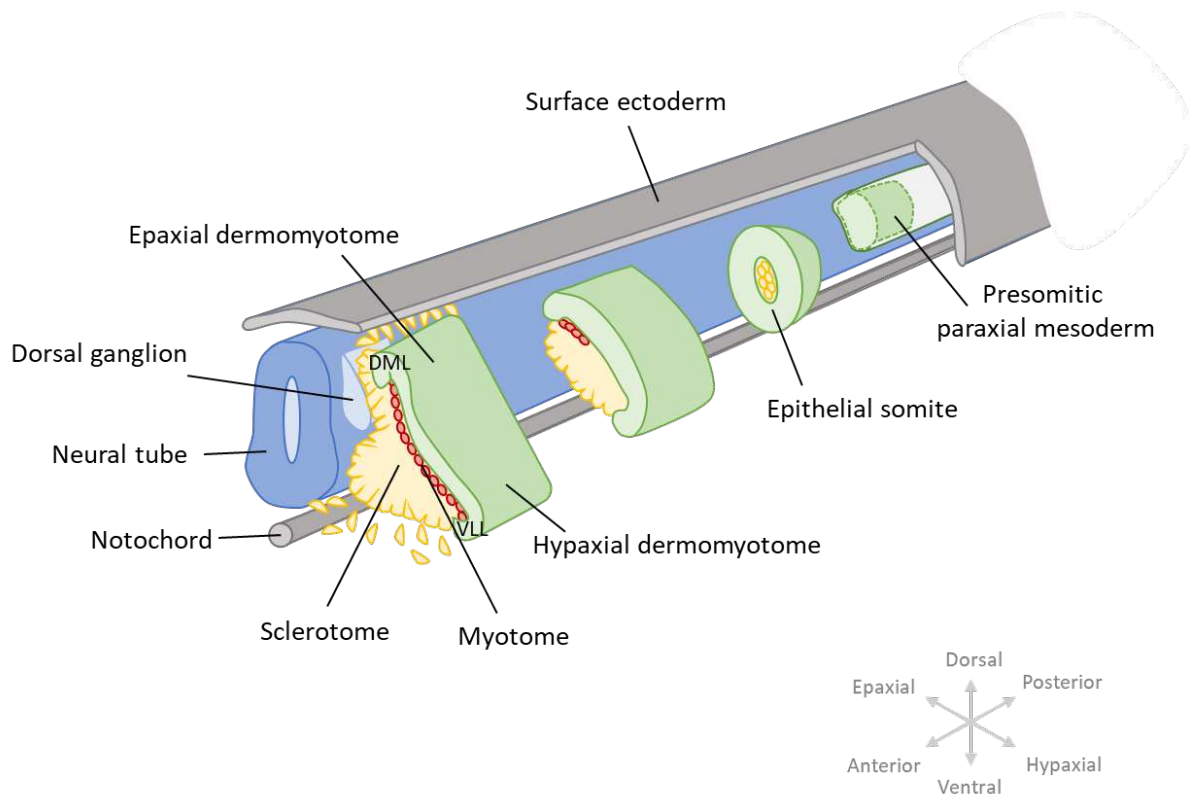
of repeating units called sarcomeres, which are the functional units of muscle contraction. The sarcomeres contain thick and thin filaments, giving the striated aspect to the fiber, composed of contractile proteins: myosin and actin, respectively. The interaction between these filaments enables muscle contraction (Squire, 1997). Furthermore, at the periphery of muscle fibers, there are cells known as satellite cells, which are skeletal muscle stem cells. These cells remain in a quiescent state outside the muscle fibers yet within the muscle tissue. Satellite cells play a vital role in the processes of muscle regeneration and repair. In response to injury or exercise-induced stress, satellite cells become activated, undergoing a phase of proliferation, and subsequently differentiating into new muscle progenitor cells. These newly formed muscle progenitors will either merge with preexisting muscle fibers or generate entirely new fibers, thereby contributing to the recovery and adaptation of the muscle (Mauro, 1961; Collins et al., 2005).

Muscles are formed during the development through a process called myogenesis. This complex journey involves the transformation of the muscle precursor cells into mature and functional muscle fibers (Bentzinger et al., 2012). Despite sharing a common structural organization, each muscle has a unique morphological identity defined by a specific size, number of nuclei, orientation, and attachment sites to tendons. This specificity in morphology enables specific functions crucial for the precise movement of appendages. While the transcriptional core of myogenesis appendages is well characterized, the molecular process controlling the development of specific individual muscle morphologies remain unknown.

## **A.2. Muscle development**

### **A.2.a. Overview of muscle development**

Vertebrate myogenesis is a highly complex and precisely regulated process that underlies the development and functionality of muscle tissue. Diseases of the locomotor system are at the origin of several handicaps like for example muscular dystrophy affecting



**Figure 3 - Schematic representation of vertebrate somitogenesis**

Segmentation of the paraxial mesoderm into somites occurs along the dorsal-ventral axis and in a rostral to caudal direction. In response to signals from the notochord and the neural tube. The presomitic paraxial mesoderm is located on either side of the notochord. Then, the somites differentiate and subdivide to give rise to the dermomyotome and the sclerotome. The dermomyotome subdivide into hypaxial and epaxial dermomyotome. Then, cells of the Dorsal Medial Lip (DML) migrate under the dermomyotome to form the epaxial myotome while cells of the Ventral Lateral Lip (VLL), form the hypaxial myotome.

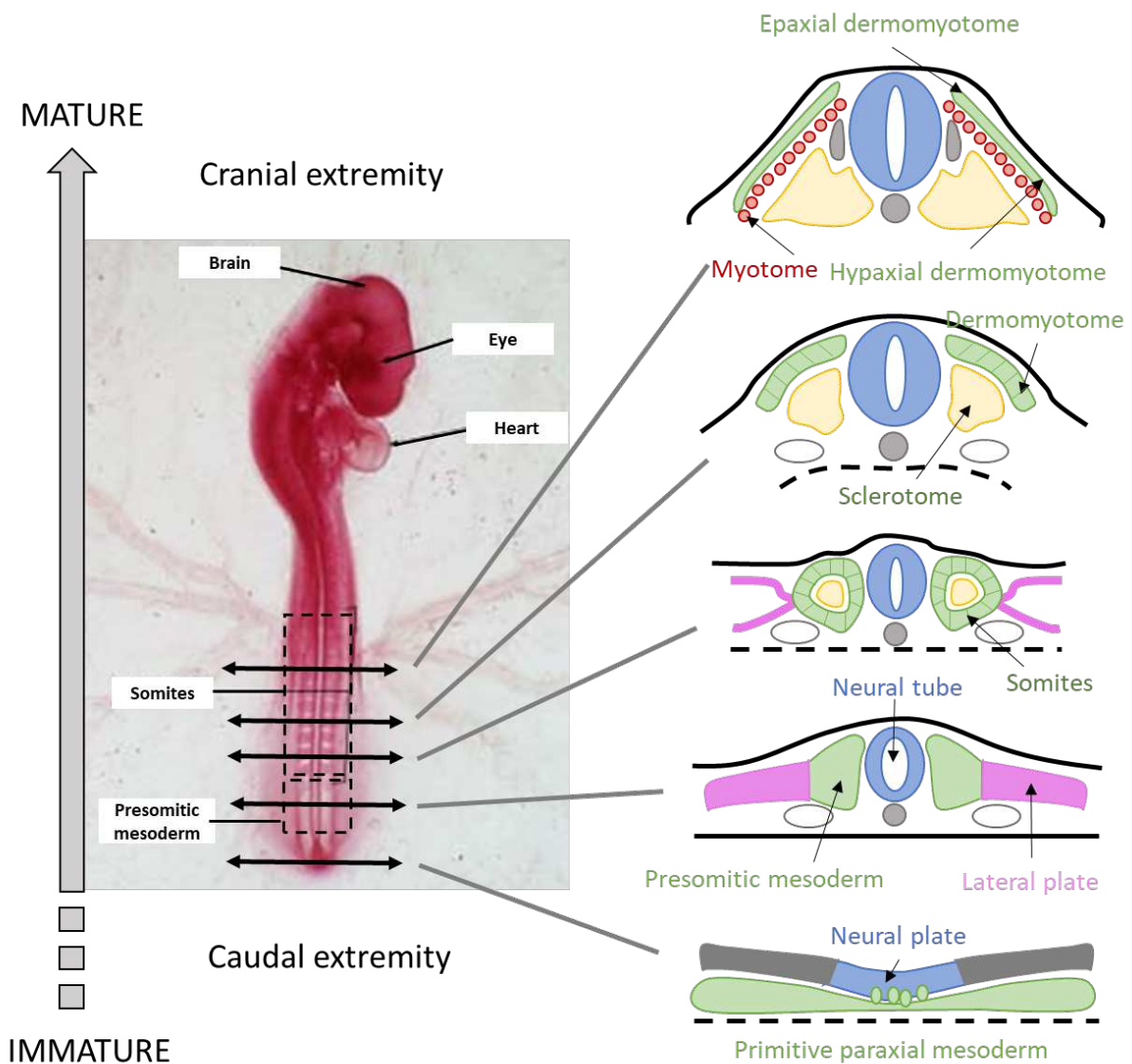
*Adapted from Buckingham et al., 2003*

different subpopulation of muscles. These types of diseases suggest that muscles are heterogeneous populations. Understanding the intricate mechanisms and regulation of the development of specific muscle morphology identity is a key challenge for the scientific community to develop efficient therapies.

The myogenic process entails a series of steps involving muscle cell specification, proliferation, differentiation, migration, fusion, and maturation (**Figure 2**) (Amthor et al., 1999). Initiated during embryonic development, groups of cells are assigned to become myogenic precursor cells, arising from mesodermal tissues in the developing embryo. These precursor cells undergo proliferation and differentiation, ultimately transforming into myoblasts. Once committed to the muscle lineage, these myoblasts express specific molecular markers associated with muscle development. Then, myoblasts migrate to their target locations within the developing embryo. The migration of myoblasts is guided by several molecular cues and signals. Upon reaching their designated sites, myoblasts differentiate into myocytes which align and fuse together, creating multinucleated muscle fibers. The fusion process involves the merging of plasma membranes and the incorporation of multiple nuclei into a single muscle fiber. Following fusion, the newly formed muscle fibers continue their maturation, undergoing structural and functional changes. They elongate and develop highly organized arrangements of contractile proteins, such as actin and myosin filaments, which produce the characteristic striations observed in mature muscle tissue (Wigmore and Evans, 2002). Throughout development, muscle fibers become categorized into distinct types based on their contractile properties and metabolic characteristics, influenced by genetic and environmental factors. The specification of fiber types determines their functional capabilities, ranging from endurance to strength.

### **A.2.b. Somitogenesis**

Before the myogenesis, during gastrulation, the embryonic cells undergo reorganization into three germ layers: the ectoderm, the endoderm, and the mesoderm (**Figure 3 and 4**) ( Buckingham et al., 2003; Arnold and Robertson, 2009). These germ layers are fundamental for the subsequent development of specific primitive systems during



**Figure 4 - Gastrulation and neurulation in the chick embryo**

Picture in the left is Chick embryo stage 16 (3 days old), somites, brain, eye, and heart are visible.

Schematics in the right represent different transversal slides along the cranio-caudal axis of the embryo showing the different stages of somites development.

The presomitic mesoderm coming from the primitive paraxial mesoderm will segment into somites that will give the epaxial and hypaxial dermomyotome, the sclerotome, and the myotome.

*Embryo picture is from Judy Cebra-Thomas, Swarthmore College, Developmental Biology 24*

*Adapted from pictures by H. Sternsheim*

organogenesis. First, the ectoderm, the outermost germ layer, gives rise to the epidermis and nervous system. Second, the endoderm, the innermost germ layer, is at the origin of the digestive system and lungs. Third, the mesoderm, the middle germ layer, contributes to the development of a wide range of tissues, including muscles, bones, connective tissues, blood vessels, and muscles. More specifically, mesoderm layer divide into axial, paraxial, intermediate, and lateral plate mesoderm. The axial mesoderm is at the origin of the notochord while the paraxial mesoderm give rise to the somites. The intermediate mesoderm connects the paraxial mesoderm with the lateral plate. Finally, the lateral plate mesoderm serve for the development of the circulatory system, as well as to the mesodermal components of the limbs (**Figure 5**) (Ferretti and Hadjantonakis, 2019).

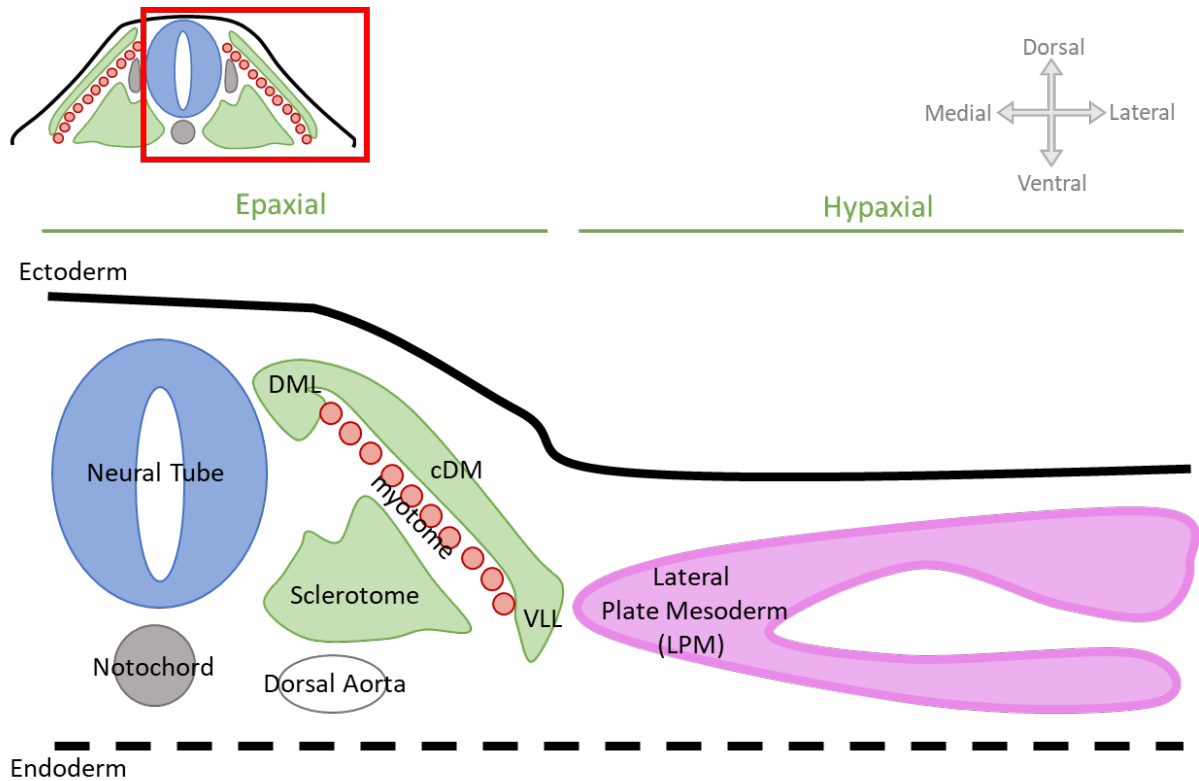
The paraxial mesoderm undergo the process of somitogenesis characterized by its segmentation along the antero-posterior axis into repeating units called somites (**Figure 4**) (Bumcrot and McMahon, 1995). Afterward, tissues adjacent to the somite express signals such as Wnt, Bone Morphogenetic Protein (BMP), Sonic Hedgehog (Shh) and Noggin that will induce the formation of the dermomyotome (in dorsal position) and the sclerotome (in ventral position) (Bryson-Richardson and Currie, 2008). The sclerotome is the precursor of the trunk skeleton, while the dermomyotome give rise to both the dermatome, contributing to the skin dermis, and the myotome, responsible for the trunk muscles (Christ and Ordahl, 1995; Buckingham, 2001).

### **A.2.c. Compartmentalization of the dermomyotome**

Within the dermomyotome distinct domains are establish. The more medial part of the somite gives rise to the epaxial dermomyotome, while the lateral part contributes to the formation of the hypaxial dermomyotome (Christ and Ordahl, 1995; Buckingham, 2001). Muscle precursors for the limb, tongue, and larynx originate from the hypaxial dermomyotome, located in the latero-ventral part of the dermomyotome, while other head muscle precursors derive from the paraxial mesoderm (Noden, 1983a, 1983b, 1986).

Previous studies showed that the signals Wnt, BMP and Shh secreted by tissues adjacent to somites induce the regionalization of somites with the formation of the





**Figure 5 - Vertebrate myotome formation during embryonic development**

In dorsal position, somite differentiates into the dermomyotome divided into 3 regions (central dermomyotome (cDM), Dorsal Medial Lip (DML) and Ventro Lateral Lip (VLL)). The dermomyotome is at the origin of the myotome. More specially, the VLL of the dermomyotome give limb myogenic progenitors that will migrate into the LPM.

In ventral position, somites undergo epithelial-to-mesenchyme transition to form the sclerotome. BMP signaling (not shown) produced by the LPM transiently inhibits somitic lineage differentiation. At the same time, neural crest delaminates from the dorsal neural tube and, while migrating ventrally, contacts dermomyotomal cells to promote myogenic induction.

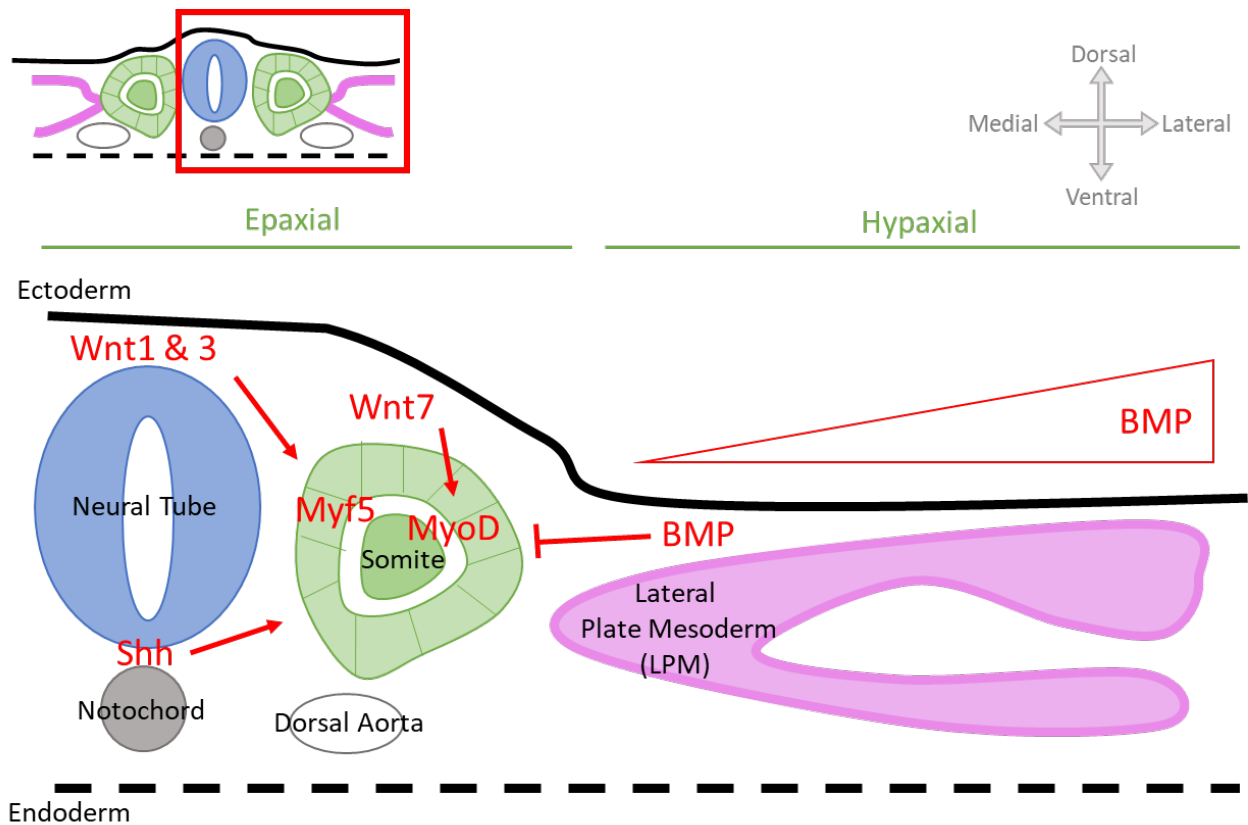
*Adapted from Chal and Pourquié, 2017*

dermomyotome and the sclerotome (**Figure 6**). Indeed, the combined action of Wnt1 or Wnt3 (which interact with frizzled receptors), and Shh from the dorsal neural tube and notochord allows the differentiation of the epaxial dermomyotome, in dorsal position. These signals induce the myogenic differentiation of cells of the dorsal part of the dermomyotome at the origin of intercostal muscles and back muscles (Ikeya and Takada, 1998). On the other hand, BMP4, from the paraxial mesoderm, and Wnt7, by the ectoderm, control the differentiation of the hypaxial dermomyotome in ventro-lateral position (Miyazono et al., 2005). Moreover, some myoblasts of the hypaxial dermomyotome migrate into the limb bud. The closest to the neural tube give rise to the extensor muscles of the limb while the most distant give rise to the flexor muscles of the limb (Musumeci et al., 2015). Thereby, the signals Wnt1 and 3, Shh, BMP are at the heart of the myoblast specification process (Bryson-Richardson and Currie, 2008). Moreover, Shh is secreted by the notochord and the floor plate induced the formation of the sclerotome in the ventral region of the somite. The notochord also produces Noggin, which inhibits BMP signaling and supports sclerotome development (Fan and Tessier-Lavigne, 1994; Hirsinger et al., 1997; Marcelle et al., 1997).

#### **A.2.d. Development of limb muscles**

##### **A.2.d.1. Cellular description of limb muscle development**

Limb myogenesis is divided into two waves representing distinct temporal phases of muscle cell development and allow the formation, growth, and refinement of the musculature during vertebrate development. The primary wave, embryonic myogenesis, serves as the foundation for early muscle development, while the secondary wave, fetal myogenesis, adds to muscle size, complexity, and adaptability, shaping the final architecture of the musculature (Wigmore and Evans, 2002).



**Figure 6 - Compartmentalization of vertebrate somites during embryonic development**

Each somite is patterned into different compartments, dorsoventral, mediolateral and anteroposterior, thanks to factors secreted by the surrounding tissues.

In dorsal position, Wnt pathway specified the dermomyotome and BMP (from the LPM) inhibits the differentiation.

In ventral position, Shh (from the midline) induce sclerotome differentiation.

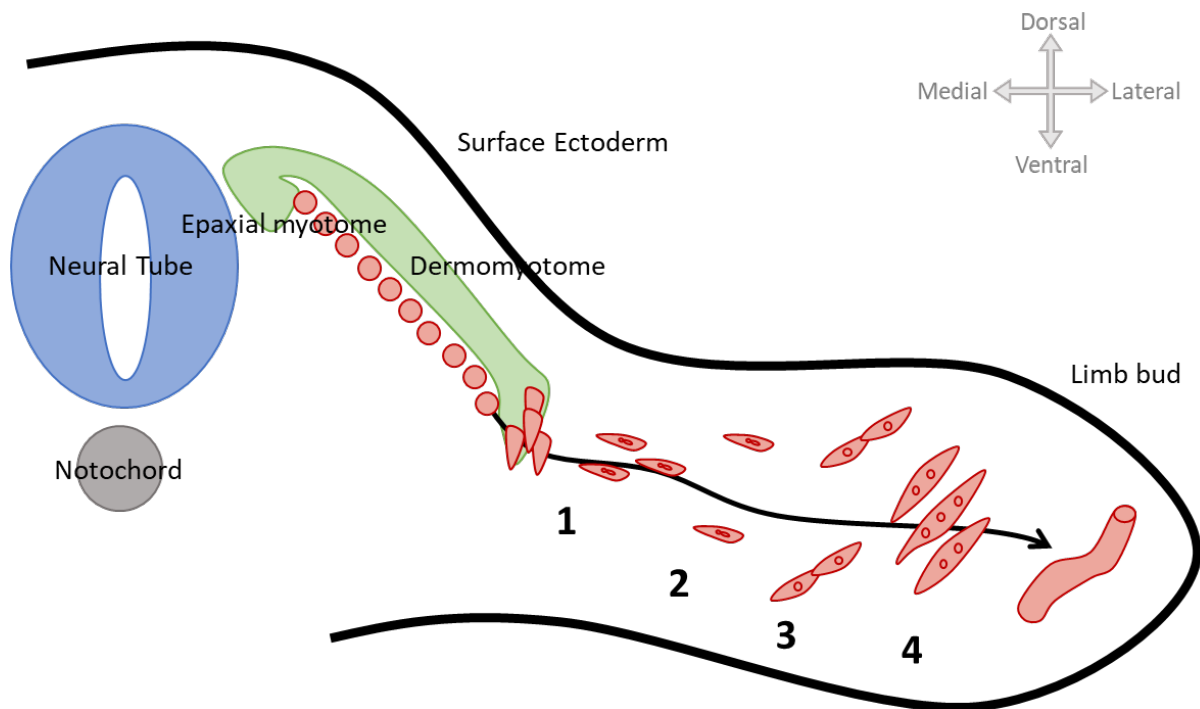
*Adapted from Chal and Pourquié, 2017*

#### **A.2.d.1.1. Delamination and migration**

During the primary wave of myogenesis, which takes place during the initial stages of limb bud formation, muscle precursor cells are specified within the latero-ventral part of the dermomyotome known as the hypaxial dermomyotome (**Figure 7**). These undifferentiated muscle precursor cells undergo a process known as delamination from the ventrolateral dermomyotomal lips. This delamination represents a crucial step in the migration of muscle precursors from the dermomyotome to colonize the developing limb buds (Bladt et al., 1995). During the delamination process, muscle precursors detach from the ventrolateral lips of the dermomyotome and undergo a process known as epithelial-to-mesenchymal transition (EMT). This transition allows the cells to detach from their original location in the dermomyotome and acquire a migratory mesenchymal phenotype (Buckingham et al., 2003). Then, muscle precursors respond to various guidance cues and signaling molecules that guide their migration towards the developing limb bud (Chevallier et al., 1977) (*see molecular basis paragraph*).

#### **A.2.d.1.2. Proliferation**

Proliferation of muscle precursors is a fundamental process crucial for expanding the myogenic cell population, ultimately leading to the formation of a sufficient number of muscle fibers. Once positioned within the limb bud, myoblasts enter a phase of proliferation to increase their numbers. This proliferation is precisely regulated by a complex interplay of various signaling molecules, growth factors, and transcription factors that tightly control cell cycle progression and cell division (Buckingham and Relaix, 2007) (*see molecular basis paragraph*).



**Figure 7 - Development of vertebrate limb muscles**

Schematic representation of limb development with the different steps of myogenesis.

Cells from the dorsomedial lip (DML) migrate under the DM to form the epaxial myotome, while cells originating in the ventrolateral lip (VLL) migrate under the DM to generate the hypaxial myotome.

1/ Delamination & Migration

2/ Proliferation

3/ Determination & Differentiation

4/ Fusion

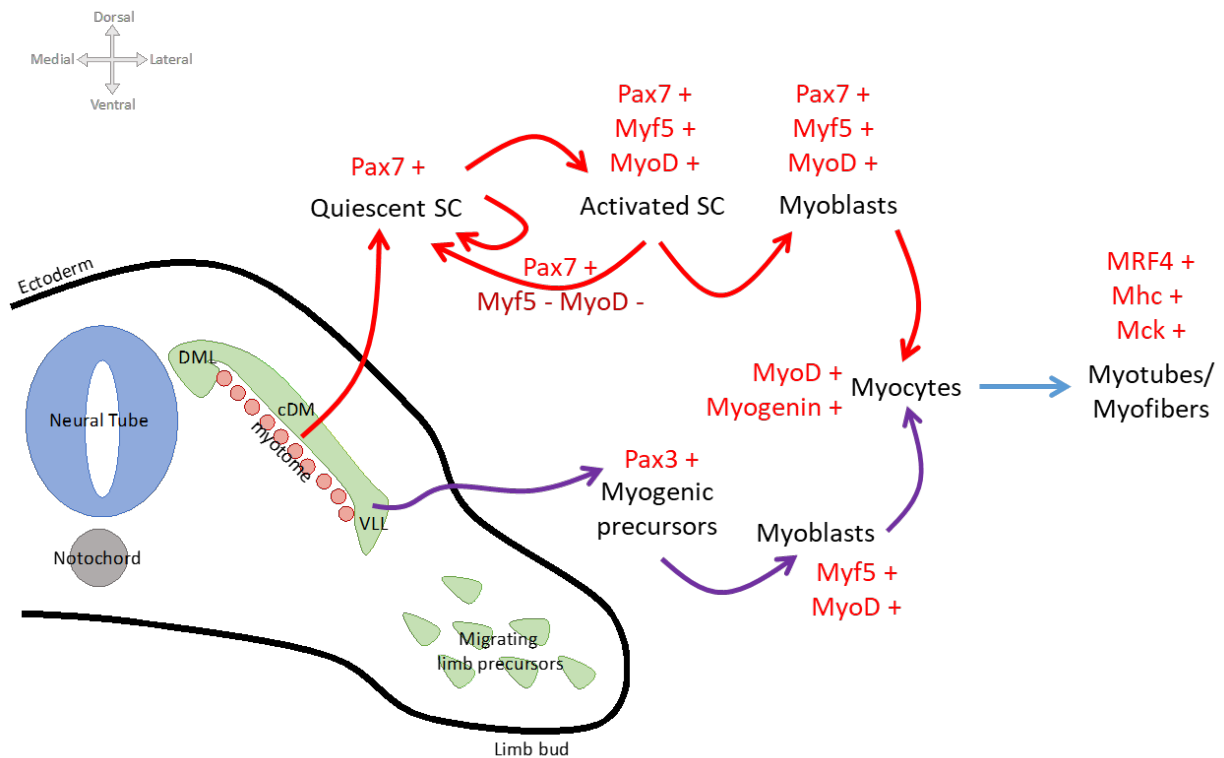
*Adapted from Buckingham et al., 2003*

#### **A.2.d.1.3. Determination and differentiation**

Prior to differentiation, muscle precursors undergo a crucial transition as they exit the cell cycle and withdraw from active cell division. They will extend along the antero-posterior axis over the entire length of the somite and form mononuclear fibers called myocytes (Biressi et al., 2007). This step is essential for muscle precursors to progress through the subsequent stages of differentiation (Davis et al., 1987; Montarras et al., 1991). As muscle precursors exit the cell cycle and become myocytes that will differentiate, they undergo significant changes in gene expression, driven by key regulators from the Myogenic Regulatory Factor (MRF) family. The MRFs are crucial helix-loop-helix transcription factors with sequential expression, playing a central role in muscle precursors differentiation (Pownall et al., 2002) (*see molecular basis paragraph*).

#### **A.2.d.1.4. Fusion**

Upon exiting the cell cycle, muscle precursors initiate the fusion process, where multiple myocytes fuse together to form multinucleated structures called primary myotubes (Biressi et al., 2007). These primary myotubes represent the initial muscle structures within the developing limb and serve as the basis for subsequent muscle development. The fusion process is a hallmark of muscle differentiation and is vital for the subsequent development of functional muscle fibers. Within the myotubes, the assembly of sarcomeres begins. Sarcomeres, the contractile units of muscle cells, are composed of actin and myosin filaments that interact to produce muscle contractions. The organization of these filaments in a highly ordered pattern allows for efficient and coordinated muscle function (Lehmacher et al., 2012). As myotubes mature, they undergo further remodeling and refinement processes. This includes the alignment and organization of sarcomeres along the length of the myotubes, leading to the formation of functional myofibrils within the muscle fibers. Additionally, myotubes establish neuromuscular junctions with motoneurons, enabling the transmission of nerve impulses to initiate muscle contractions. The maturation process also involves the fine-tuning of muscle fiber properties, such as size and fiber type, to ensure proper functioning within the developing limb.



**Figure 8 - Molecular description of MRFs during limb muscle development**

Schematic representation of the role of MRFs family during limb development. First, the Pax3+ cells undergoing epithelial–mesenchymal transition delaminate and migrate to the region of presumptive limb muscle development and form the myotome. Then, Pax3/Pax7 positive cells migrate from the central region of the dermomyotome (red arrows).

Pax3 activate the expression of Myf5 and MyoD and initiate the myogenic program that leads to the development of myoblasts. Although myogenic precursor cells express both Pax3 and Pax7+ cells are more prominent in the central region of the DM that gives rise to the major pool of satellite cells (SC, cells that are mitotically quiescent)

Myoblasts generated from both embryonic and postnatal myogenesis undergo extensive proliferation, which leads to the generation of Myogenin+ myocytes. Finally, these myocytes fuse to form myotubes and subsequently myofibers, which continue to express the terminal differentiation marker MRF4, structural and metabolic genes such as the Myosin heavy chain (Mhc) and the Muscle Creatinine Kinase (MCK) genes.

*Adapted from Singh and Dilworth, 2013*

#### **A.2.d.1.5. Second wave of myogenesis**

The secondary wave of myogenesis occurs during fetal development and extends into postnatal stages. During this wave, secondary myoblasts are generated and, like primary myoblasts, undergo proliferation and differentiation. This leads to the formation of secondary myotubes, which subsequently fuse with existing primary muscle fibers. This fusion process contributes to the growth and enlargement of the muscles, adding to their overall size and complexity ( Duxson et al., 1989; Dunglison et al., 1999). Additionally, this fusion process contributes to the refinement and addition of muscle fiber types, further shaping the musculature. As the secondary myotubes continue to mature and develop, they acquire specific contractile properties and characteristic fiber types.

#### **A.2.d.2. Molecular description of limb muscle development**

The development of limb muscle cells involves a series of distinct steps leading to the formation of functional muscles: specification, delamination, migration, proliferation, determination and finally differentiation. Throughout this intricate process, the coordinated action of signaling pathways, transcription factors, and epigenetic regulators is crucial for precise and successful muscle development (Vasyutina and Birchmeier, 2006).

##### **A.2.d.2.1. Specification of muscles precursors**

Within the ventrolateral lips of the dermomyotome, a subset of cells becomes specified as muscle precursors. These muscle precursors express key transcription factors (TFs), such as Pax3 and Pax7 (Jostes et al., 1990; Goulding et al., 1991; Kassar-Duchossoy et al., 2005; Relaix et al., 2006). Pax3 is expressed in the ventrolateral lip of the dermomyotome and contributes to the determination of the myogenic lineage, while Pax7 is essential for the survival, self-renewal, and proliferation of muscle stem cells (satellite cells) derived from hypaxial precursors (**Figure 8**).





Moreover, two key signaling pathways, Shh and Wnt, play critical roles in the specification of hypaxial precursors (Fan and Tessier-Lavigne, 1994; Bryson-Richardson and Currie, 2008). Shh, secreted by the notochord and floor plate of the neural tube, acts as a morphogen, creating a concentration gradient. The ventral concentration of Shh is essential for inducing the expression of Pax3 and Pax7 (Cairns et al., 2008). By selectively regulating Pax3 expression, Shh restricts its presence to the dorsomedial lip, where myogenic precursor cells are specified, while suppressing Pax3 expression in the ventral dermomyotome. This differential regulation of Pax3 by Shh helps establish distinct domains within the dermomyotome and promotes the formation of the myogenic lineage within it. In addition to Shh signaling, the canonical Wnt/ $\beta$ -catenin pathway of Wnt signaling also plays a role in the specification of hypaxial precursors by initiating and maintaining the myogenic program in these cells. Wnt signaling acts synergistically with Shh signaling to promote the expression of Pax3 and the myogenic program in hypaxial precursors (Marcelle et al., 1997; Bryson-Richardson and Currie, 2008).

#### **A.2.d.2.2. Delamination of muscle precursors**

The receptor c-Met is a key player in the specification of hypaxial precursors during limb development. It is a target of Pax3 and is expressed in all hypaxial precursors present in both the ectoderm and mesenchyme of the limb bud. On the other hand, its ligands, scatter factor/hepatocyte growth factor (SF/HGF), are expressed in the delaminated cells of the axial region and play a role in the delamination process of muscle cell precursors from the dermomyotome (Dietrich et al., 1999). It was demonstrated that ectopic expression of SF/HGF can induce delamination of cells that typically do not undergo this process. These signaling events promote cell survival, proliferation, and differentiation into myogenic cells, ultimately contributing to proper muscle formation and patterning within the limb.



### A.2.d.2.3. Migration of muscle precursors

These undifferentiated muscle precursors then delaminate and migrate to their final positions, where they will differentiate into mononuclear myocytes. The migrated cells retain proliferating characteristics due to myogenic determination factors, which increase the number of precursors in their final position (Buckingham and Rigby, 2014). SF/HGF and c-Met signaling, also play crucial roles in the migration and organization of muscle precursor cells. They promote migration of myoblasts and their integration into the developing limb musculature. The activation of c-Met by SF/HGF triggers intracellular signaling cascades, including the Mitogen Activated Protein Kinase (MAPK) and the phosphoinositide 3-kinase/protein kinase B (PI3K/Akt) pathways, which regulate cytoskeletal dynamics and cell migration. Additionally, Fibroblast Growth Factors (FGFs), particularly FGF4 and FGF8, secreted by the ectoderm and limb mesenchyme, create concentration gradients that guide the migration of muscle precursor cells. FGFs also activate downstream pathways that regulate the directional migration of muscle precursors towards the developing limb (Kahane et al., 2001). FGF2 and FGF4 are secreted by the mesenchyme and provide mitogenic signals to myoblasts, activating downstream pathways like the MAPK pathway, which stimulate cell cycle progression and myoblast migration.

Furthermore, the pool of cells that migrate expresses Ladybird Homeobox 1 (Lbx1), which allows the expression of genes involved in the migration process. Lbx1 cells possess specific axial identity derived from Hox patterning (Alvares et al., 2003). Lbx1 expression is induced in migrating cells to facilitate their lateral migration from the dermomyotome to the limb bud. Moreover, the Six-Eya-Dach complex has been shown to play a role in the migration and survival of muscle precursors (Viaut and Münsterberg, 2020).

It is also important to highlight the role of surface molecules in the migration capacity of muscle precursors. Molecules like N-cadherin and Integrin have been demonstrated to contribute to the migration and guidance of muscle precursor cells during limb muscle development (MacKrell et al., 1988).



#### **A.2.d.2.4. Proliferation of muscle precursors**

The homeodomain transcription factor, Muscle segment homeobox 1 (Msx1) play a role in the proliferation phase. Msx1 is expressed in migrating progenitors at the level of the upper limb. In cell culture, Msx1 is found in dividing myoblasts, and its overexpression in differentiated muscle cells induces a reversion to proliferative cells (Houzelstein et al., 1999).

Moreover, Insulin-like Growth Factors (IGFs) serve as essential regulators of muscles precursors proliferation. Produced by multiple cell types, including mesenchymal cells and muscles precursors themselves, IGFs bind to IGF receptors on muscles precursors, initiating intracellular signaling cascades that promote cell growth and proliferation. It was shown that IGF-1 significantly induces the proliferation of cultured muscles precursors in a dose-dependent manner (Yu et al., 2015). Furthermore, SF/HGF and c-Met signaling pathways contribute to muscles precursors proliferation by acting as mitogens (Gal-Levi et al., 1998).

#### **A.2.d.2.5. Determination and differentiation**

The MRFs form a family of basic helix-loop-helix Transcription Factors (TFs) with sequential expression composed of Myogenic factor 5 (Myf-5), Myogenic Differentiation 1 (MyoD), Myogenin, and MRF4. To begin, Myf-5 and MyoD are expressed in muscle precursors before differentiation and are required for the determination into myocytes (**Figure 8**) (Tajbakhsh et al., 1997; Relaix et al., 2004). Indeed, knockout studies in mice have shown that in their absence, skeletal muscle fails to form. In the opposite, ectopic expression of MRFs initiate the myogenic program. Moreover, it was shown that Shh regulates the expression of MyoD by directly controlling the expression of Myf5 in the epaxial domain (Borycki et al., 1999). Finally, Myogenin and MRF4 are necessary for terminal differentiation (Venuti et al., 1995). Myogenin allow the differentiation of myocytes in myotubes while MRF4 play a role in the myofibers maturation.

Furthermore, the MRFs promote the expression of other factors like the Myocyte Enhancer Factor 2 (Mef2), which cooperates with MRFs to enhance the expression of muscle



genes and the formation of functional muscle fibers (Molkentin et al., 1995; Potthoff and Olson, 2007). Mef2 is expressed in the mesoderm and then its expression decrease but its stay high in the somatic mesoderm and the developing muscle (Lilly et al., 1994).

In parallel, transforming Growth Factor-beta (TGF- $\beta$ ) family members, including TGF- $\beta$ , myostatin, and activin, negatively regulate myoblasts differentiation by inhibiting MRF expression and promoting the expression of inhibitors of muscle differentiation. For instance, Myostatin suppresses myoblast differentiation and inhibits muscle fiber formation (Bentzinger et al., 2010).

Moreover, components of the Extracellular Matrix (ECM), such as fibronectin, laminin, and collagen, provide structural support and signaling cues for myoblast differentiation. Interactions between myoblasts and ECM molecules influence cytoskeletal rearrangements, cell adhesion, and the activation of intracellular signaling pathways involved in differentiation (Chaturvedi et al., 2015).

Finally, the Notch-Delta Signaling pathway regulates the balance between myoblast proliferation and differentiation. Notch activation suppresses myoblast differentiation, while Delta ligands promote differentiation. The downregulation of Notch signaling allows the expression of myogenic regulatory factors, leading to myoblast differentiation (Kopan et al., 1994).

### **A.3. Generation of muscles morphological diversity**

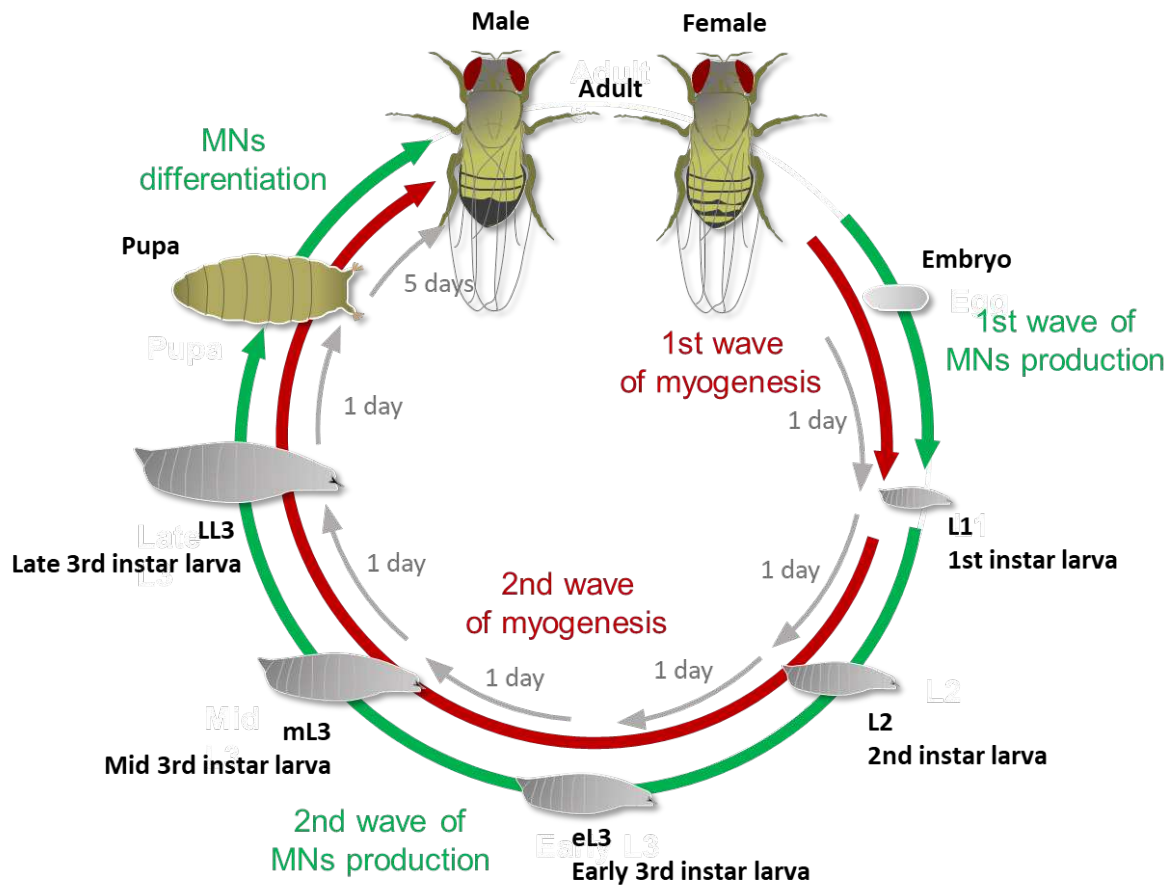
Currently, there is limited knowledge about the comprehensive gene regulation underlying muscle diversity. Lots of studies have been done on the general myogenesis programs regulating gene expression in a spatially and temporally controlled manner during limb development. Several transcription factors, play essential roles in establishing regional identities within the limb. However, how the morphological diversity is established is still not understood.





The vertebrate body is partitioned into domains that exhibit different identities conferring an antero-posterior identity. The acquisition of these identities is controlled by a class of genes called Hox genes which belongs to the homeobox genes superfamily and encodes homeodomain transcription factors (Krumlauf, 1994; Pearson et al., 2005). Hox genes are differentially expressed along the anterior–posterior axis of the embryo allocating distinct morphological identities to each body part (Mallo, 2018). Indeed, precursors expressing more 5' Hox genes ingress later than those expressing more 3' Hox genes and, hence will be positioned more posteriorly along the antero-posterior axis. This sequential expression pattern establishes a link between the gene expression order in the Hox gene complex and their spatial expression along the antero-posterior axis. As a result, each somite acquires a unique combination of Hox gene expression since inside the same complex, genes are shifted in expression by at least one somite. Hox roles have been confirmed by manipulation of their expression resulting in homeotic transformations, where the morphology of one given body part is transformed into that of another (Nowicki and Burke, 2000). Moreover, studies show in *Drosophila* that the homeotic genes Antp, Ubx, AbdA, and AbdB control the formation of somatic muscles by autonomous function in the mesoderm (Michelson, 1994; Roy et al., 1997). Indeed, when these genes are overexpressed in the mesoderm of segments more anterior, the muscle pattern of these segments take a more posterior identity.

This molecular difference for the antero-posterior identity is illustrated by the head muscles myogenesis. Indeed, myogenesis in the first two branchial arch (embryologic structures that develop into the lower face, neck, and part of the upper thorax) is not same despite the expression of the same TFs code. Muscle formation in the first branchial arch is controlled by the TF Pituitary homeobox 2 (Pitx2, bicoid–related homeobox gene) which is expressed in muscle precursors. Pitx2 controls myogenesis of the first branchial arc by regulating the expression of other genes such as Tbx1, MyoR/Msc and TCF21 (capsiline) which with Pitx2 control the expression of MyoD and Myf5 (Shih et al., 2007). *tbx1* mutants show malformations of only few muscles of the first two branchial arcs while double mutants MyoR/Msc and TCF21 cause malformations of the first branchial arc with only a loss of specific muscles (Kelly et al., 2004). These data suggest that myogenesis induction could be regulated differently in each branchial arc (only *tbx1* affects the second branchial arc).



**Figure 9 - *Drosophila melanogaster* life cycle**

*Drosophila* development follows seven stages: egg, L1, L2, L3 (divided into early, mid and late L3), pupa, and adult.

Motoneurons (MNs) are produced in two waves of neurogenesis: one during the embryonic development, and another one during larval and pupal phases.

Muscles are derived from two waves of myogenesis: an early embryonic wave providing the larval muscles, and a second wave, during the metamorphosis, giving rise to the adult muscles.

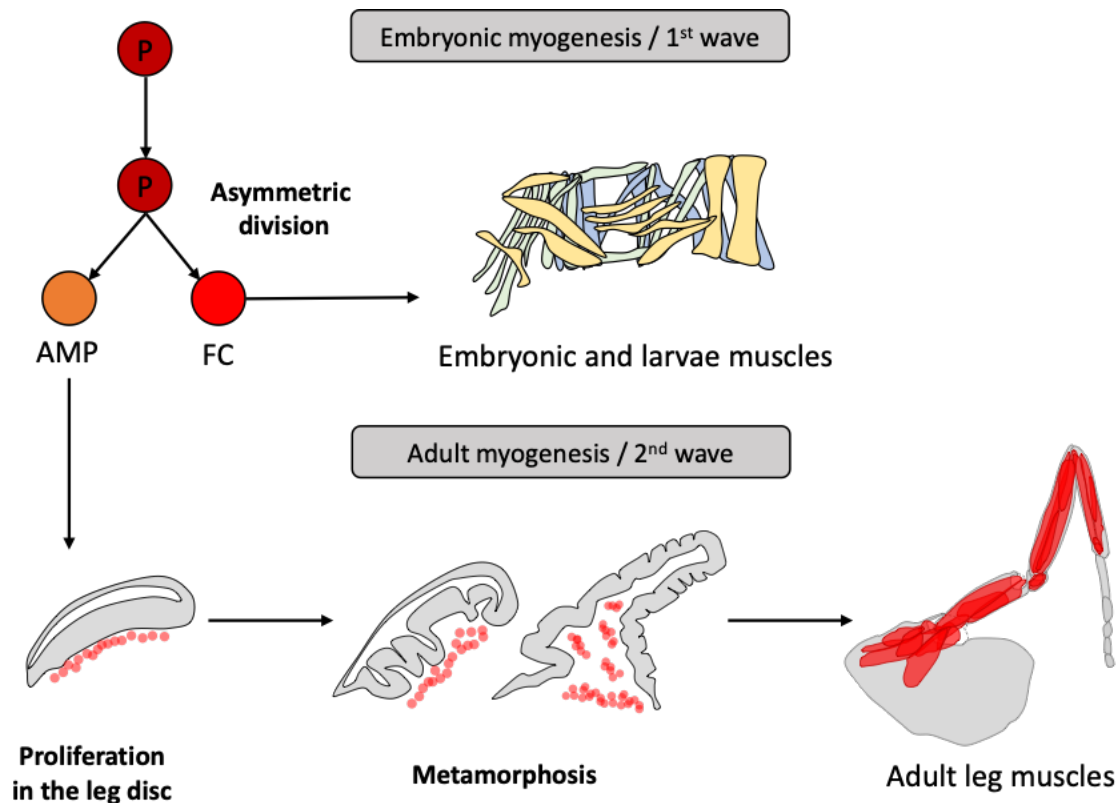
Moreover, recent studies show the importance of surrounding connective tissue cells in orchestrating muscle morphogenesis by regulating the differentiation and organization of muscle precursor cells (Besse et al., 2020). Indeed, the extracellular matrix and its stromal cells forming the Muscle Connective Tissue (MCT) creates a developmental pre-pattern that orientates and controls the positioning of myogenic precursors that differentiate into myofibers forming individual limb muscle bundles. Indeed, if the TF Tbx5 is conditionally deleted from forelimb MCT cells, muscle patterning is disrupted, demonstrating the role of this gene, acting in MCT cells, in developing muscles with correct organization and function. The expression pattern of the MCT serve to define the size and shape of muscles with the orientation of its myofibers.

However, how muscles derived from the same somites acquire different morphologies is still not understood. The work of this thesis is to understand how the morphological diversity is generated at the level of the limb development. For this purpose, I used *Drosophila melanogaster* as a model system to study the development of leg muscles morphologies.

## **B. *Drosophila melanogaster***

*Drosophila melanogaster* is a highly powerful model system for investigating the intricate genetic, molecular, and cellular processes that govern muscle development (**Figure 9**). One of the main advantages of using *Drosophila* is its well-annotated genome, which provides comprehensive understanding of the genes involved in development. Additionally, the availability of numerous genetic tools allows for precise manipulation and study of gene function, enabling targeted investigations into specific genes' developmental roles.

In *Drosophila*, myogenesis is divided into two waves (**Figure 10**) (Bate M. 1993). The first wave takes place during early embryonic development, giving rise to the larval somatic muscles (Baylies et al., 1998; Laurichesse and Soler, 2020; Rout et al., 2022). These muscles play essential roles in the larval stage, contributing to locomotion and feeding functions. This wave also give rise to the Adult Muscle Precursors (AMP) that will be use during the second



**Figure 10 - Drosophila adult muscle development: two waves of myogenesis**

The first wave takes place during early embryonic development, giving rise to the larval somatic muscles. One selected muscle progenitor (P) undergoes asymmetric division leading to either two founder cells (FC) to generate the larval muscles, or one FC and an adult muscle precursor (AMP) to form the adult muscles.

The second wave of myogenesis occurs during metamorphosis, giving rise to the adult skeletal muscles from the AMP originated in the first wave.

These AMP migrate to colonize the leg imaginal disc and form a layer on the surface of the disc. Then, during larval stages and metamorphosis, the leg disc evaginates out from its center and elongates along the proximo-distal axis to give the final adult leg shape. During this process, the myoblasts fuse between themselves to give rise to multinucleated adult muscle fibers.

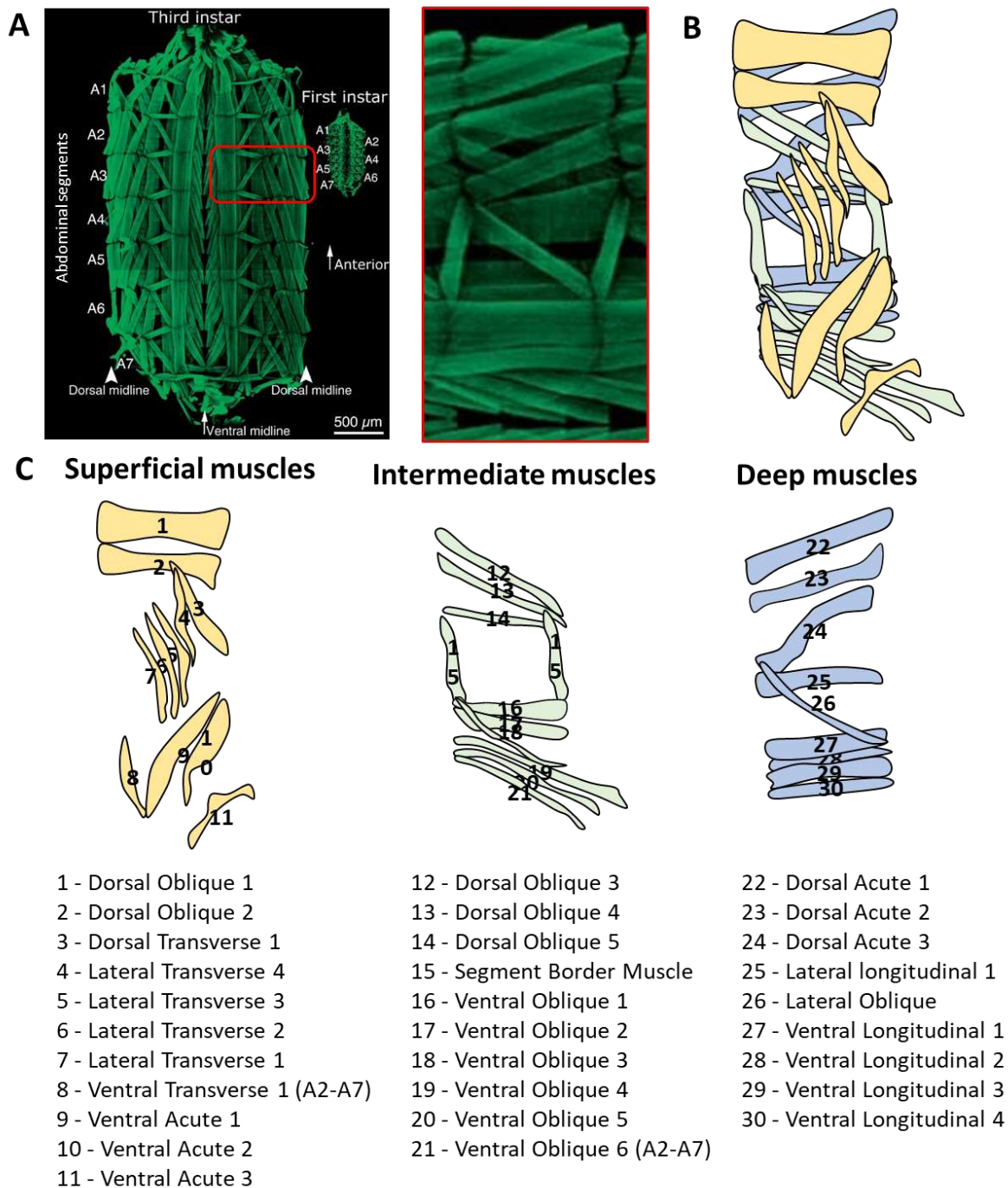
wave of myogenesis. This second wave occurs during metamorphosis and give rise to the final adult skeletal muscles. Most adult muscle are *de novo* formed while some adult muscles originated from a remodeling of larval muscles (Fernandes et al., 1991). The adult skeletal muscles are vital for enabling the fly to perform essential locomotor activities required for adult-specific behaviors.

## **B.1. The larval body wall muscles**

### **B.1.a. Organization of the larval body wall muscles**

The organization of muscles in *Drosophila* larvae follows a highly stereotyped pattern and arrangement of muscle fibers within a two-dimensional system. The somatic muscles are primarily responsible for the locomotion and movement of the larvae. Each somatic muscle is composed of a single multinuclear muscular fiber, and they are attached to specific cuticular attachment sites called tendon cells (differentiated cells derived from the epidermis), which serve as anchor points to the body wall cuticle.

Somatic muscles are arranged in a segmental pattern, with muscles found in each body segment. Within the larval body, the abdominal segments A2 to A7 follow an identical, stereotyped pattern consisting of 30 muscles organized into three layers. However, the thoracic segments T2-T3 and abdominal segment A1 exhibit distinct organizational patterns. On the other hand, the first thoracic segment (T1) and the eighth abdominal segment (A8) have fewer muscles with more variation among them (Bate, 1990) (**Figure 11**). This stereotyped pattern and precise positioning of these muscles are crucial for ensuring proper function and contribute significantly to the overall stability of *Drosophila* larval locomotion (Bate M. 1993).



**Figure 11 - The larval body wall muscles**

A/ 3rd instar larva compared to 1st instar larva, FITC-conjugated Phalloidin staining of larval body wall (BW) muscles (*Picture from Gorczyca and Budnik 2006*)

B/ Schematic representation of the BW muscles of one abdominal segment.

C/ Schematic representation of the BW muscles in the three spatial position: superficial, intermediate and deep muscles.

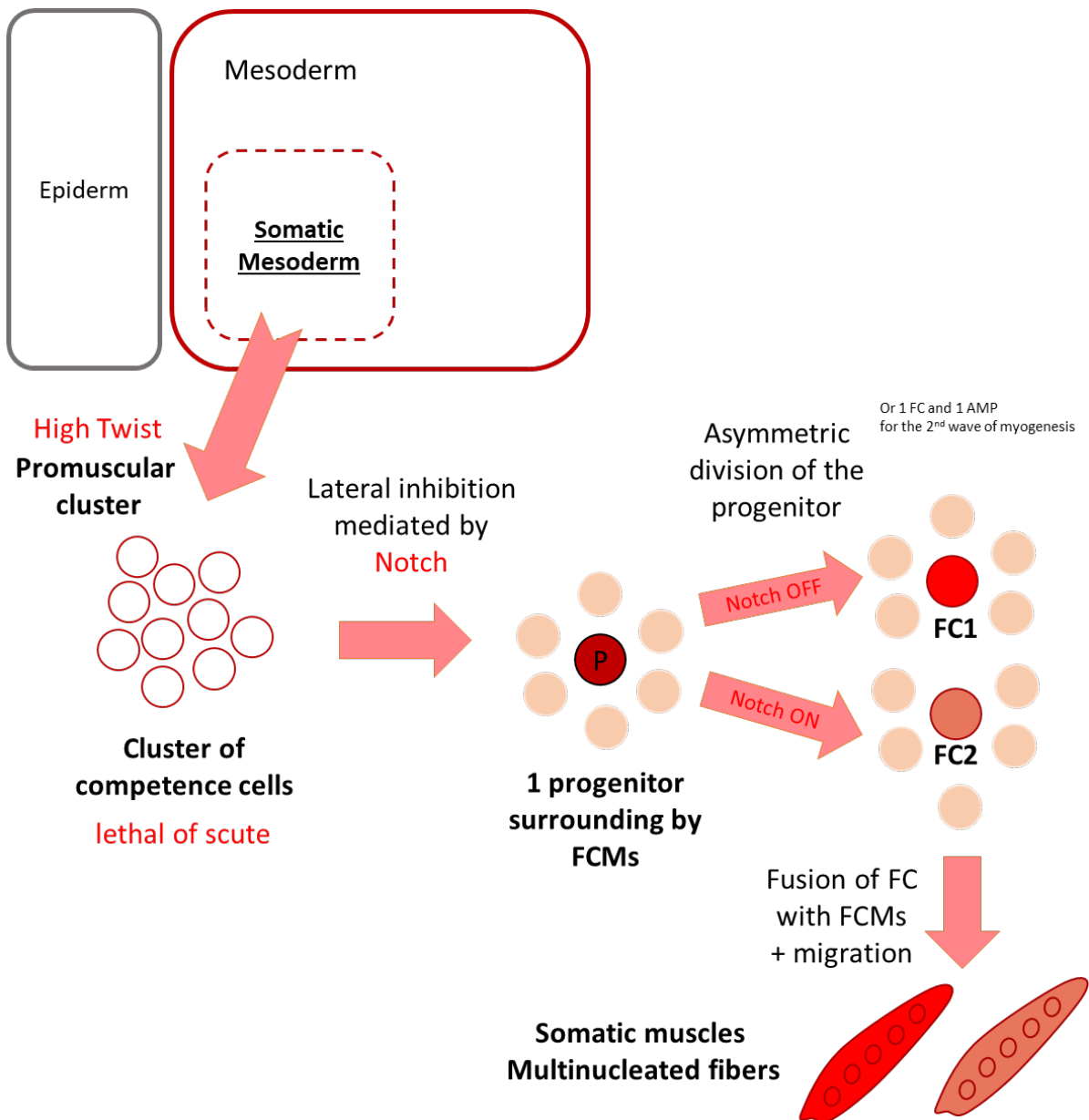
## **B.1.b. Larvae body wall muscles development**

### **B.1.b.1. Cellular basis of larvae myogenesis**

In *Drosophila* larvae, as in vertebrates, the mesoderm is divided into segmental units under the influence of signaling molecules such as TGF $\beta$ , Wnt, and Hh family members. From these segmental units, different myogenic lineages emerge. The origin of the embryonic mesoderm in *Drosophila* occurs during gastrulation in which cells at the ventral side of the embryo move inward, resulting in the formation of a furrow-like structure called the ventral furrow that undergoes invagination (Baylies and Bate, 1996). Within the ventral furrow, a group of cells undergoes an epithelial-to-mesenchymal transition (EMT), a process where cells lose their epithelial characteristics and adopt a mesenchymal phenotype. Following EMT, these mesenchymal cells undergo delamination, detaching from the epithelial layer and moving into the interior of the embryo. The cells that delaminate from the ventral furrow form the primary mesoderm, which serves as the source of mesodermal cells in the developing embryo. The embryonic mesoderm is defined very early during the development by the expression of a basic helix-loop-helix protein domain, Twist (Baylies and Bate, 1996). Subsequently, the primary mesoderm undergoes extensive cell division and migration to populate different regions of the embryo. Within the mesodermal tissue, the expression of Twist becomes higher in a region that gives rise to the somatic mesoderm (the lower twist segment gives rise to cardiac and visceral mesoderm) (Bate, 1993; Roy and Vijayaraghavan, 1999).

After specification of the mesoderm, groups of cells, called pro-muscular clusters, are specified by the expression of the protein Lethal of scute (Carmena et al., 1995). From each of this cluster, one progenitor cell is selected by lateral inhibition mediated by Notch, while the other cells become Fusion Competent Myoblasts (FCM or naïve myoblasts). Then, each progenitor undergoes asymmetric division to generate two Founder Cells (FC) or one FC and one Adult Muscle Precursor (AMP, cell precursor of adult muscle that stays quiescent until the metamorphosis and is used during the second wave of myogenesis, see next sections).





**Figure 12 - Drosophila myogenesis first wave**

Somatic mesoderm is determined in the anterior part of each segment by a high level of twist. Competence domains are determined by the expression of the proneural gene *lethal of scute* (*l'sc*) that restrict its expression *l'sc* to form promuscular clusters. In each cluster, one cell is determined as a muscular progenitor (P) that induces higher Notch activity in the other cells of the cluster, which give the fusion-competent myoblasts (FCMs). Then the P divides asymmetrically, giving rise to either two founder cells (FCs) or one FC and one adult muscle precursor (AMP) via an asymmetric division. Finally, different FCs fuse with a defined number of FCMs giving rise to the formation of myotubes that give functional muscles.

*Adapted from Tixier, et al. 2010*

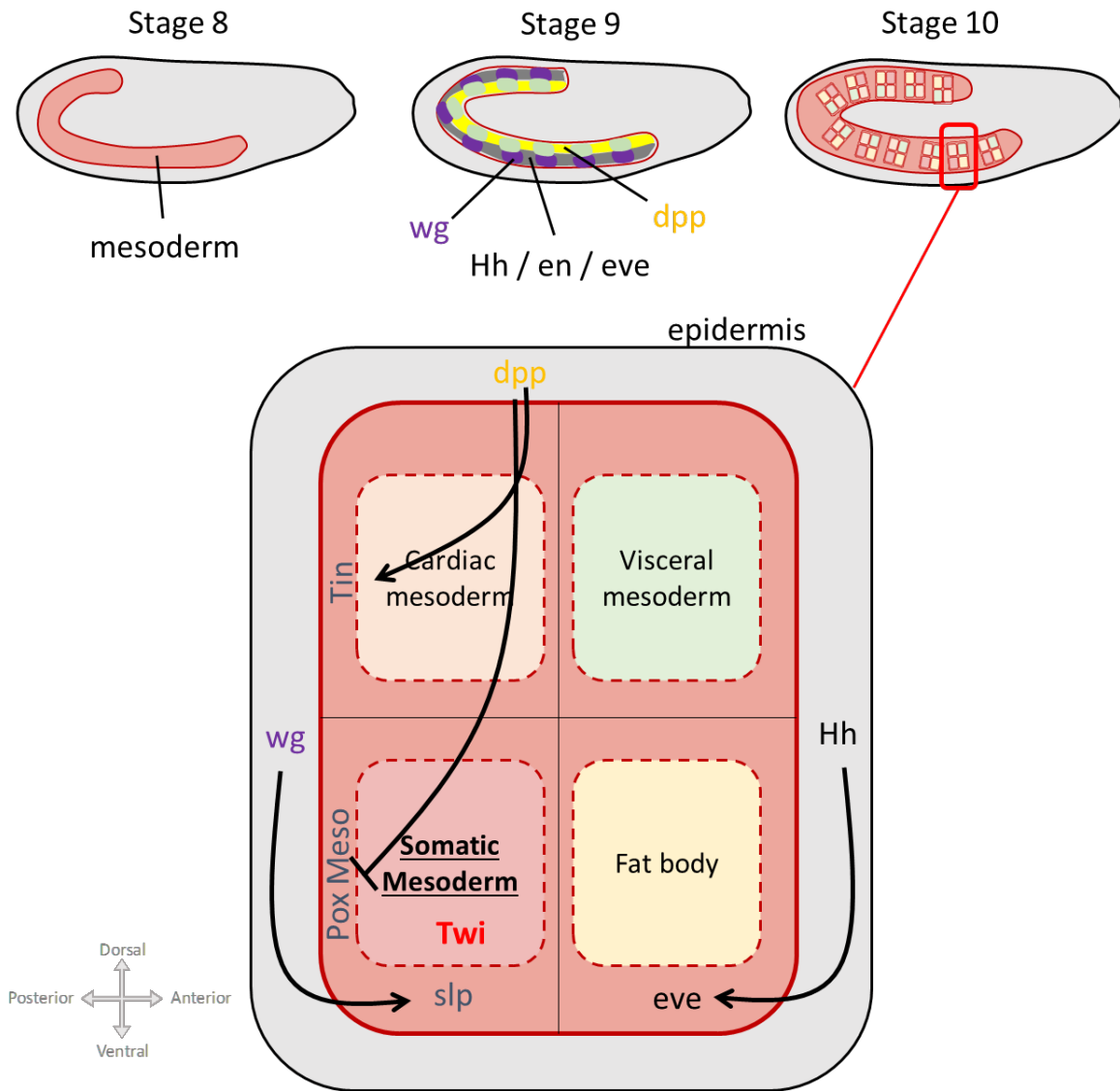
Finally, each FC fuse with a define number of FCM and form multinucleated muscle fibers (Baylies et al., 1998; Rochlin et al., 2010; Rout et al., 2022). The myoblast fusion process involves a series of coordinated events. First, myoblasts recognize and adhere to each other through cell surface receptors and adhesion molecules present on their cell membranes. These interactions facilitate the physical alignment of myoblasts and create a bridge between them. Following the initial adhesion, signaling pathways are activated within the myoblasts. These signaling pathways trigger a cascade of intracellular events that promote cytoskeletal rearrangements, cell membrane remodeling, and ultimately, the merging of individual myoblasts. As a result, the myoblasts fuse together to form multinucleated myotubes (**Figure 12**).

### **B.1.b.2. Molecular basis of larvae myogenesis**

#### **B.1.b.2.1. Segmentation of the mesoderm**

The patterning of the mesoderm allows the proper development and organization of larval somatic muscles into intra-segmental patterning with the dorso-ventral and antero-posterior axis (**Figure 13**).

The dorso-ventral patterning of the mesoderm is under the control of Decapentaplegic (Dpp), a member of the transforming growth factor-beta (TGF- $\beta$ ) superfamily, secreted by the ectoderm, that plays a role in patterning the identity of muscle cells along this axis (like the BMP in vertebrates). Indeed, Dpp signaling is involved in the specification of muscle progenitors in the somatic mesoderm. Dpp, secreted from the dorsal ectoderm, acts as a morphogen, forming a concentration gradient in the developing embryo. Different levels of Dpp signaling induce distinct cellular responses, specifying different mesodermal fates, to establish the segment-specific patterns of muscle development (Staebling-Hampton et al., 1994). Moreover, Dpp act on the expression of several mesodermal gene. Indeed, Dpp restrict the expression of the TF Tinman in the dorsal region of the mesoderm where it will specify cardiac cells, visceral mesoderm, and dorsal muscles (Yin et al., 1997; Zaffran et al., 2006).



**Figure 13 - Molecular segmentation of the embryonic mesoderm**

At stage 8, the presumptive mesoderm is homogeneous. Then, at stage 9, a characteristic pattern is generated with stripes of Hh, eve and en that delimit anterior domain, alternating with stripes of Wg and Slp that delimit posterior border. Moreover, Dpp further delimits ventral versus dorsal mesoderm domains by influencing the expressing of Tinman and Pox Meso. Finally, mesoderm specification finished at stage 10, when wg/slp domains give rise to the dorsal cardiac mesoderm and ventral somatic muscles (with high level of Twist). Additionally, en/eve/Hh domains, give rise to the dorsal visceral mesoderm and the ventral fat body.

*Adapted from Mbodj et al., 2016*

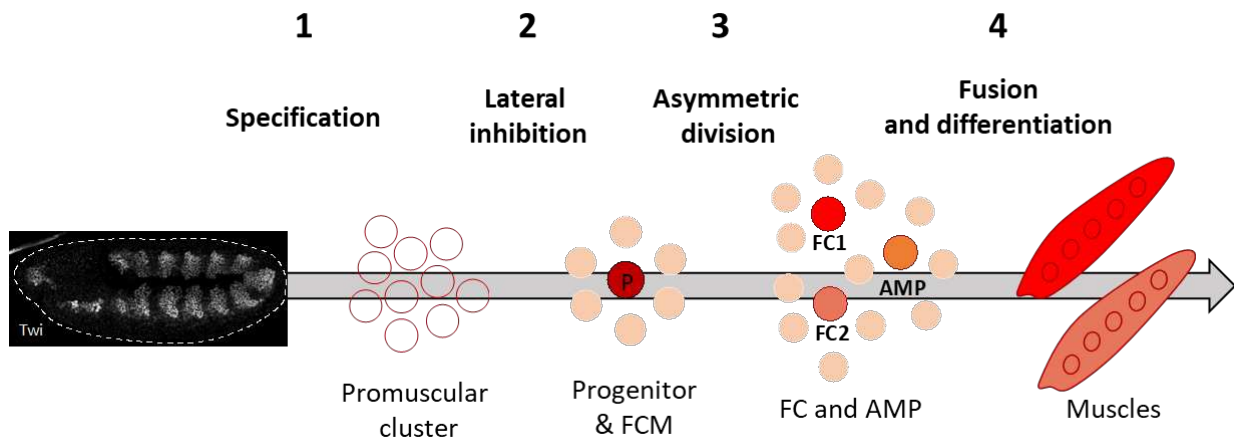
Moreover, Dpp dorsally represses Pox meso and restricts its expression to the ventral mesoderm where it will act on the ventral and some lateral muscles specification (Duan et al., 2007).

In parallel, the establishment of the antero-posterior axis in the mesoderm is tightly controlled by several signaling pathways, including wingless (wg, homologous of Wnt), hedgehog pathway (Hh, homologous of Shh in vertebrates), and two pair rule genes, even-skipped (eve) and sloppy paired (slp) (Azpiazu et al., 1996; Riechmann et al., 1997). These genes act on the gradient of expression of Twist in the mesoderm along the antero-posterior axis. In the posterior domain, Hh positively regulates eve, resulting in a lower level of Twist expression, while, in the anterior domain, wg positively regulates slp, leading to a higher level of Twist expression. This gradient of Twist is essential for the development of muscles from the somatic mesoderm (Dobi et al., 2015).

This gene network along the different axis allows the specific segmentation of the mesoderm into distinct domains. The visceral mesoderm (at the origin of the visceral muscles) and the fat body develop within the expression territory of eve and Tin, under the influence of dpp and Hh. The heart and the dorsal somatic mesoderm form within the expression territory of slp and tinman with the regulatory control of dpp and wg. Finally, the ventral somatic mesoderm is under the control of the expression of slp and Pox meso with regulatory cues from wg.

#### **B.1.b.2.2. Molecular regulation of larvae myogenesis**

After the formation of the somatic mesoderm, the larvae skeletal muscle development can be divided into four steps (**Figure 14**). First, the specification inside the mesoderm of promuscular clusters. Second, the selection of a progenitor cell by lateral inhibition. Third, the asymmetric division of each progenitor cell to generate either two FC or one FC and one AMP. Finally, each FC fuses with the FCM to form the final multinuclear muscle fibers. The molecular regulation of larvae myogenesis is well known.



**Figure 14 - Principal steps of *Drosophila* skeletal myogenesis**

- 1/ Specification of a Promuscular cluster with acquisition of myogenic competence (*Immunostaining with  $\alpha$ -Twist antibody, stage 11 embryo picture, from Borges Pinto et al., 2022*).
- 2/ Lateral inhibition with segregation of Progenitor and Fusion Competent Myoblasts (FCM).
- 3/ Asymmetric division with generation of Founder Cells (FC) and Adult Muscle Precursors (AMP).
- 4/ Each founder form a specific muscle by fusion with FCMs.

#### **B.1.b.2.2.1. Specification of the promuscular clusters and selection of progenitors**

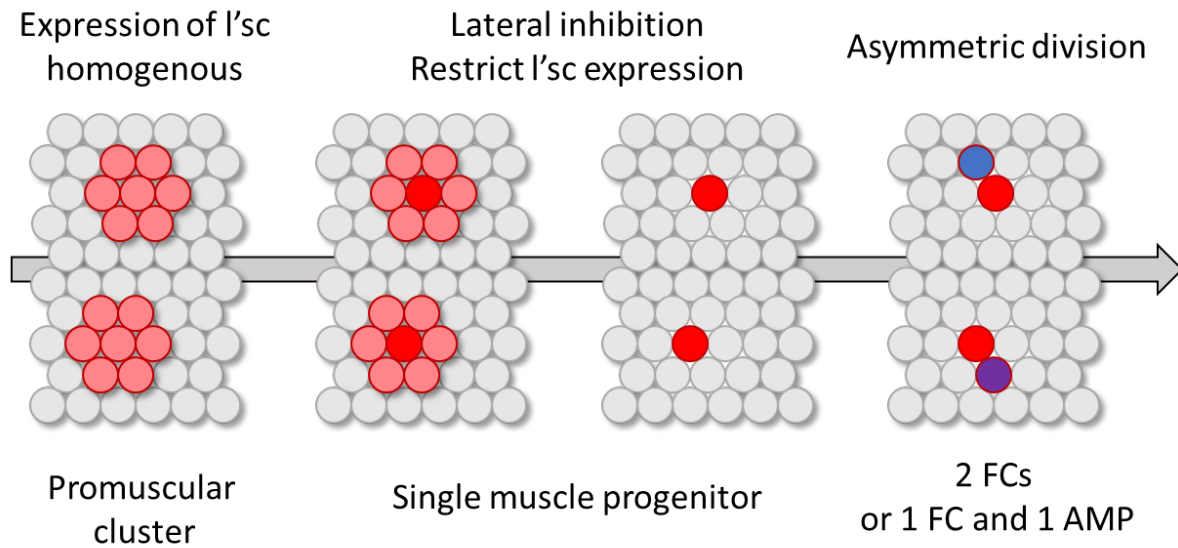
Inside the somatic mesoderm, cluster of cells express the proneural gene lethal of scute (*l'sc*), from which individual muscle progenitors are singled out by progressive restriction of *l'sc* expression (**Figure 15**) (Carmena et al., 1995). 18 promuscular clusters positioned on the dorso-ventral axis within the somatic mesoderm have been identified. Their stereotyped spatial organization is due to the intersection of the antero-posterior *wg* and the dorso-ventral *Dpp* pathways. Within each of this cluster, lateral inhibition mediated by Notch pathway limit the expression of *l'sc* in one or two cells that become progenitor cells.

#### **B.1.b.2.2.2. Formation of founder cells**

Subsequently, each progenitor divide once into two FC or one FC and one AMP by asymmetric division. This process involved the two proteins *Inscuteable* (*Insc*) and *Numb* that are localized oppositely in the progenitor cell. After the division, the Notch pathway become inactive in the FC where *Numb* is expressed and active in the FC where *Insc* is expressed (Guo et al. 1996).

#### **B.1.b.2.2.3. Fusion of founder cells with fusion competent myoblasts**

The founder cells FC subsequently fuse with the FCMs to form multinucleated myotubes, giving rise to the embryonic muscles that form the basis of the larval body wall muscles. Before the fusion process, the FCM migrate and recognize their specific FC. For doing this, FCM have filopodia that allow them to identify the FC. This recognition process is facilitated by various cell adhesion molecules which mediate cell-cell recognition and adhesion (Bour et al., 2000; Ruiz-Gomez, M. 2000; Dworak et al., 2001). Indeed, the FC expresses membrane proteins including *Dumbfounded* (*Duf*) and *Roughest* (*Rst*) while the



**Figure 15 - Specification of the promuscular clusters and selection of progenitor**

Inside the somatic mesoderm, cluster of cells express the proneural gene lethal of scute (I'sc), from which individual muscle progenitors are singled out by progressive restriction of I'sc expression. Inside each of this cluster, I'sc will be expressed only in one or two cells that will become progenitor cells. Subsequently, each progenitor divide once into two Founder Cells (FC) or one FC and one Adult Muscle Precursor (AMP) by asymmetric division.

*Adapted from Dobi et al., 2015*

FCM expresses Sticks and stones (Sns), Hibris (Hbs) and Rst. This asymmetry in the expression of the immunoglobulin domain proteins allow the specific recognition between the FC and the FCM necessary for the fusion process into functional muscle fibers (Bour et al., 2000; Strünkelnberg et al. 2001). Indeed, in Sns mutant there is no fusion between the cells, the same observation is done in the Duf mutant. This cell interaction triggers actin polymerization and the formation of fusion pores, leading to the mixing of cytoplasmic contents and the formation of multinucleated muscle fibers (Doberstein et al., 1997).

## **B.2. The adult leg muscles**

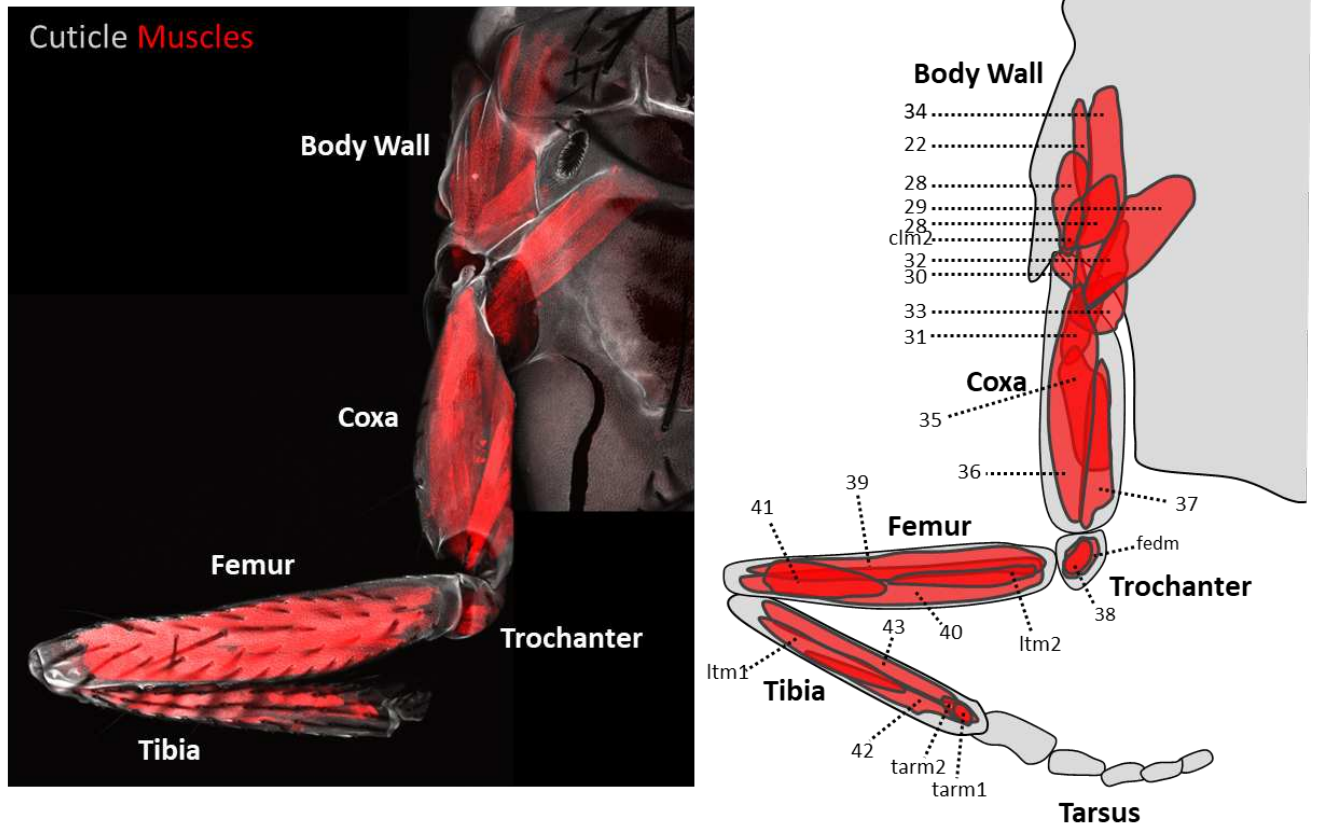
### **B.2.a. Organization of the adult leg muscles**

In *Drosophila*, each leg consists of 14 leg muscles that play a crucial role in coordinating the movement of different leg segments during walking. These muscles are composed of multiple muscle fibers and are localized in four segments of the leg: coxa, trochanter, femur, and tibia (Miller A., 1950; Baek and Mann, 2009; Brierley et al., 2012; Soler, 2004). Remarkably, each muscle exhibits a unique morphology that is stereotyped across individuals (**Figure 16**). While most studies on the development of the muscular system have focused on *Drosophila* larval body wall muscles, it is important to note that the larval system differs from the adult system. *Drosophila* adult legs serve as a structural and developmental analog to vertebrates' upper and lower limbs, characterized by tubular and striated multi-fiber muscles (Spletter and Schnorrer, 2014). In the adult leg, the muscle organization is three-dimensional, each muscle is composed of several muscle fibers, and distal tendons are independent structure from the epidermis (Soler, 2004).

### **B.2.b. Generality about adult leg myogenesis**

Two groups of adult muscles can be distinguished: those that result from larval muscle remodeling, such as the dorso-longitudinal muscles of flight, and those that originate from the





**Figure 16 – *Drosophila* adult leg muscles**

Left: confocal picture of the T1 leg, in red the muscles (Mef2>Tomato) and in gray the cuticle.

Right: schematic of the different muscles of the leg:

Body Wall muscles:

- 22 Head dorsal muscle (hdm)
- 28 Coxa levator 1 (clm 1)
- 29 Coxa depressor muscle (cdm)
- Coxa levator 2 (clm2)
- 30 Coxa abductor muscle (cabm)
- 31 Coxa anterior rotator muscle (carm)
- 32 Coxa posterior rotator muscle (cprm)
- 33 Coxa adductor muscle (cadm)
- 34 Trochanter depressor muscle 1 (tdm1)

Coxa muscles:

- 35 Trochanter depressor muscle 2 (trdm 2)
- 36 Trochanter levator muscle (trlm)
- 37 Trochanter reductor muscle (trrm)

Trochanter muscles:

- Femur depressor muscle (fedm)
- 38 Femur reductor muscle (ferm)

Femur muscles:

- Long tendon muscle 2 (ltm2)
- 39 Tibia levator muscle (tilm)
- 40 Tibia depressor muscle (tidm)
- 41 Tibia reductor muscle (tirm)

Tibia muscles:

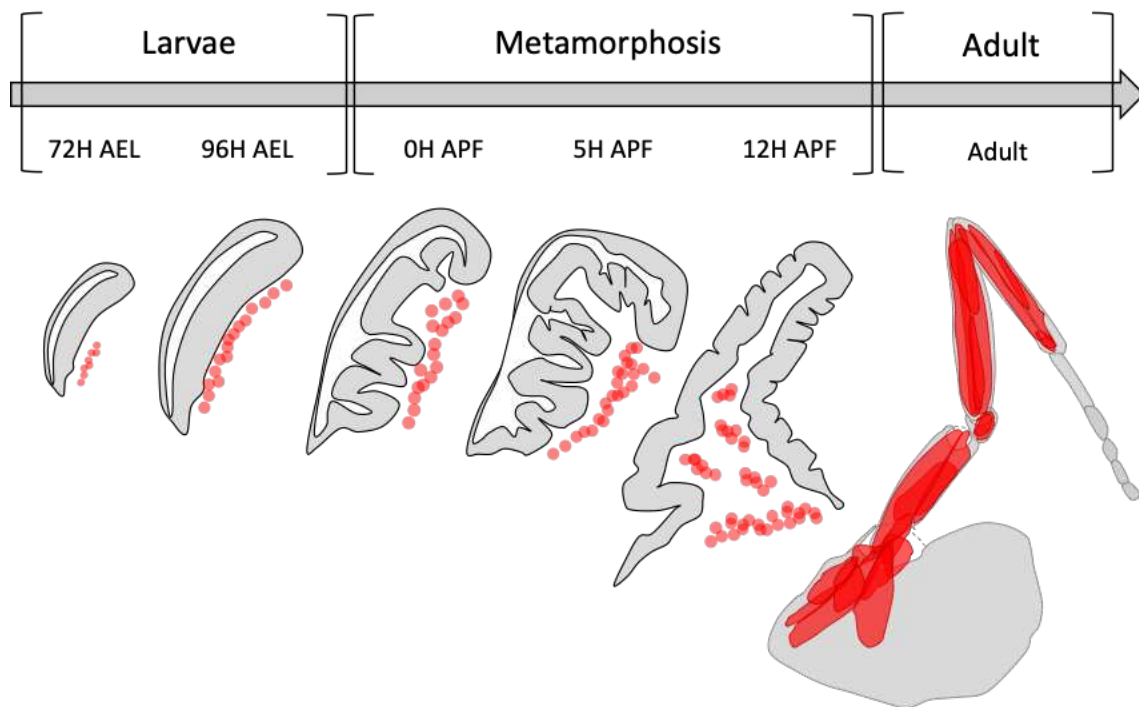
- Long tendon muscle 1 (ltm1)
- 42 Tarsal levator muscle (talm)
- 43 Tarsal depressor muscle (tadm)
- Tarsal reductor muscle 1 (tarm1)
- Tarsal reductor muscle 2 (tarm2)

proliferation and differentiation of the AMP (Fernandes et al., 1991). Indeed, during the first wave of myogenesis, some progenitors divided into one FC and one AMP which are mitotically quiescent until the metamorphosis (Rochlin et al., 2010). Then, the second wave of myogenesis begin when these resident AMPs, Twist positive, migrate to colonize the leg imaginal disc and form a layer on the surface of the disc. The leg disc is a flat epithelial sheet of cells during the larval instar stages. During larvae stages and metamorphosis, this structure evaginate out from its center and elongate along the proximo-distal axis to give the final adult leg shape (**Figure 17**) (Condic et al., 1991).

During metamorphosis, between 20h and 25h After Pupal formation (APF), once in the disc, myoblasts fuse between themselves to give rise to multinucleated adult muscle fibers (body wall musculature, flights, and legs muscles) (Broadie and Bate, 1991; Soler, 2004). After fusion to form multinucleated muscle fibers, they undergo terminal differentiation to become mature muscle fibers with specific identities and function. This process involves organization of contractile proteins and the development of specialized muscle structures. During terminal differentiation, sarcomeres, the contractile units of muscles, undergo further maturation and alignment, enabling efficient muscle contraction (Ghazi and Vijayraghavan, 2003; Gunage et al., 2017). Within the multinucleated muscle fibers, muscle-specific proteins are organized into sarcomeres, the contractile units responsible for muscle contraction. Several structural and contractile genes ensure the proper assembly of contractile proteins such as actin and myosin, and the development of regulatory proteins essential for muscle contractions.

### **B.2.c. The immature leg**

Fly legs are derived from imaginal discs that are structures composed from the sheet of monolayered epithelial cells that invaginated during embryogenesis and proliferate during the larval stages (Estella et al., 2012). The disc epithelium is morphologically a flat structure during the early larva stages that evaginate as a telescope from its center and elongates to form the adult leg. It possesses a fate map where cells at its center give rise to distal structures, furthest from the body, while the more distant from the center are at the origin of proximal structures, closest to the body, during metamorphosis (**Figure 18**). This correspondence



**Figure 17 - Drosophila leg disc development**

Schematic cross section of a leg disc during development at 5 time points:

- Larvae stages: 72h and 96h After Egg Laying (AEL)
- Metamorphosis/Pupae stages: 0h, 5h and 12h After Pupal Formation (APF)
- Adult stage

In gray is represented the epithelium and in red the myoblasts.

requires the establishment of different territories in the leg disc during the larval stages by gene regulation pathways whose have specific expression patterning (**Figure 19**). Therefore, each leg segment is specified as a concentric domain in the leg disc (E. B. Lewis, 1978). Finally, the expression of morphogens gradients allows the regionalization of the leg disc by initially specifying a posterior and anterior domain, then a ventral and dorsal domain (Estella et al., 2012).

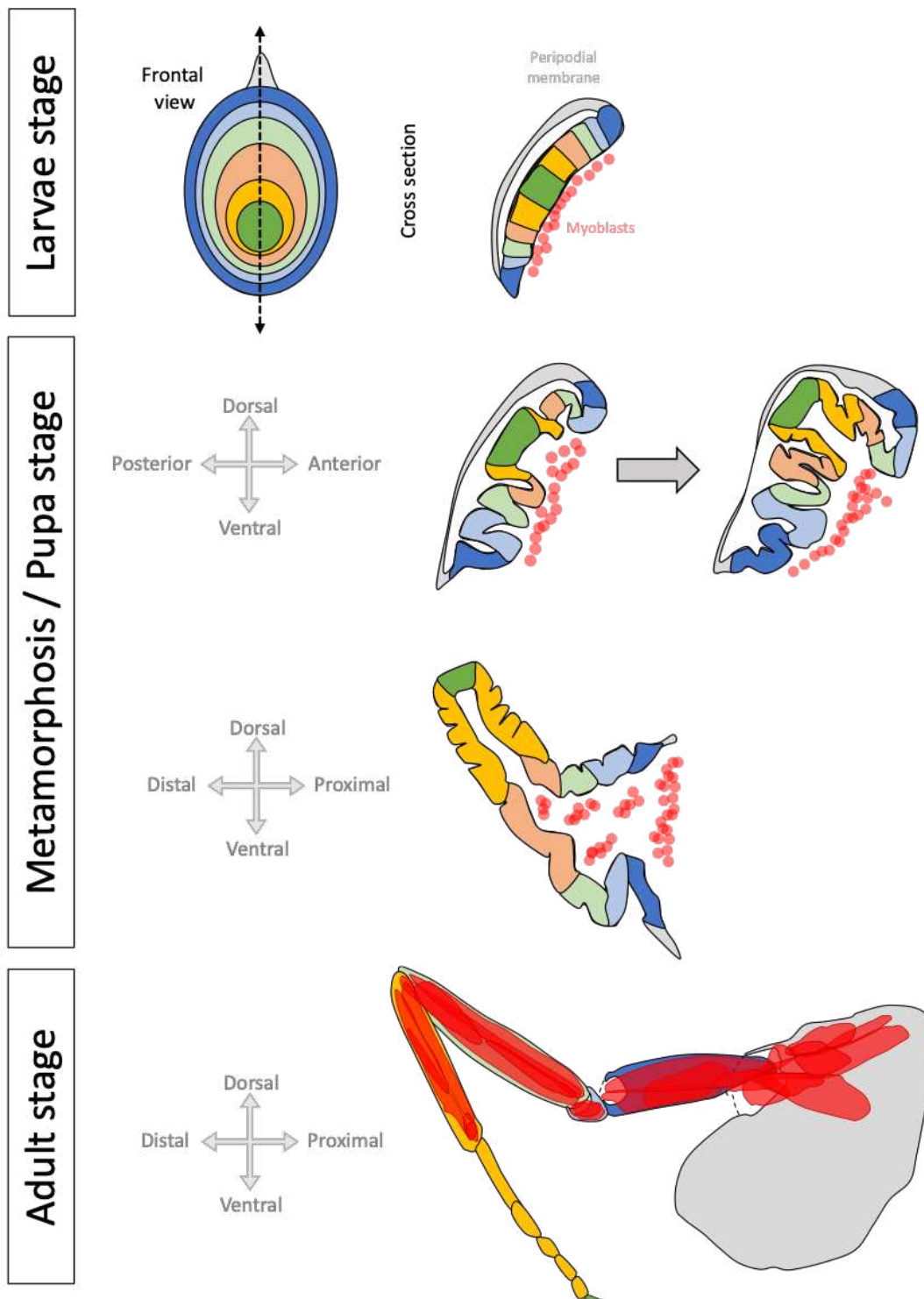
#### **B.2.d. Patterning of the leg disc**

##### **B.2.d.1. The antero-posterior axis of the leg disc**

The establishment of the antero-posterior axis in the leg disc relies on the expression of morphogens that create concentration gradients along this axis, such as engrailed (*en*), Hedgehog (*Hh*), and wingless (*wg*) (Basler, 1994; Neumann and Cohen, 1996). Indeed, *Hh* is secreted from a localized source at the anterior end of the leg disc, and it forms a gradient extending posteriorly. *en* define the posterior domain and activate the expression of *Hh* that allow the expression of *wg* and *dpp* defining the ventral region and the dorsal region respectively.

##### **B.2.d.2. The Dorso-ventral axis of the leg disc**

The dorso-ventral axis is established early in leg disc development through the action of morphogens such as *dpp* and *wg* (Ferguson and Anderson, 1992; Lecuit and Cohen, 1997). *Dpp* and *wg* possess an antagonistic action providing positional information along the dorso-ventral axis (Penton and Hoffmann, 1996). *Dpp* forms a gradient that extends dorsally, with higher concentrations in the dorsal region and lower concentrations in the ventral region. At the opposite, *wg* possess a decreasing gradient from the ventral region to the dorsal region of the disc. Studies show that this dorso-ventral specification involves mutual repression by the *wg* and *dpp* signaling systems (Brook and Cohen, 1996). Moreover, *wg* and *dpp* are diffusible morphogens, and thus their gradients are also established in the disc along the antero-posterior axis (Kojima, 2004).



**Figure 18 - Corresponding between larvae domains in the disc and the adult leg segments**

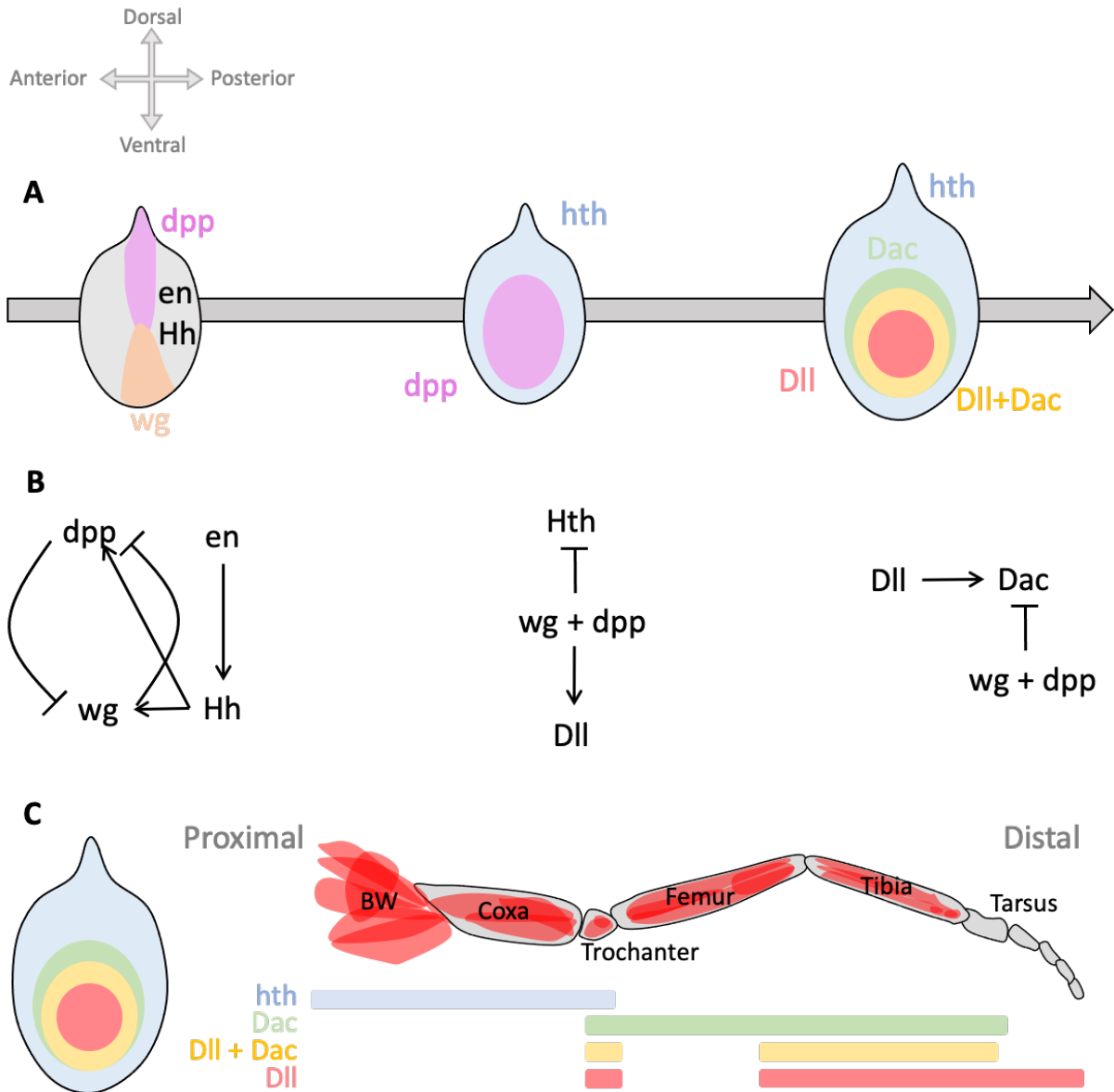
In gray the Peripodial membrane, in color the different domains/segments of the leg disc during development.

*Adapted from Kojima, 2017*

### **B.2.d.3. The proximo-distal axis of the leg disc**

The leg disc is also patterned along the proximal-distal axis, which determines the identity of the different leg segments. The morphogens *wg* and *dpp* play a role in this proximo-distal axis development. While they have antagonist action in the dorso-ventral axis establishment they act cooperatively as morphogens for the proximo-distal axis development (Diaz-Benjumea et al., 1994; Lecuit and Cohen, 1997). Removal of *dpp* and *wg* in larvae stages result in legs with dorso-ventral patterning defects but with normal organization along the proximo-distal axis (Galindo et al., 2002). Moreover, *wg* and *Dpp* expression domains meet at the center of the leg disc, corresponding to the most distal point of the presumptive leg (Campbell et al., 1993). This is this combinatory activity of both signals that is necessary to define the spatial domains of target gene expression along the proximo-distal axis of the leg as the disc develops.

The expression pattern of *dpp* in *wg* induce the expression of genes such as *distal-less* (*Dll*), *dachshund* (*Dac*) and *homothorax* (*hth*) helping to establish the proximal and distal domains within the leg field from larvae stages (Abu-Shaar and Mann, 1998; Wu and Cohen, 1999). Indeed, higher level of *dpp* and *wg* positively activate *Dll* while their low expression activity activates *Dac*. Moreover, *hth* expression is restricted to the periphery after repression by *dpp* and *wg*. These genes have an expression in concentric circles with *hth* express at the periphery of the disc while *Dll* is find in the center of the disc. The presence of *Dll* activate the expression of *Dac* defining the medial portion of the developing leg. The culmination of these genes' expression orchestrates the final leg formation process. This sequence begins with the tarsus development from the center of the disc and extends outward to the leg's integration with the body wall at the periphery of the disc. This intricate cascade of events collectively shapes the precise patterning and morphogenesis of the leg.



**Figure 19 - Drosophila epithelium patterning**

A/ Gradient of morphogens: en, Hh, dpp and wg that allow the expression of hth, DII and Dac.

B/ Network establishing the different axis

C/ Relationship between the gene expression domain establishing the Proximo-distal axis and the leg segments

*Adapted from Estella 2012*

### **B.3. Generation of muscle identity**

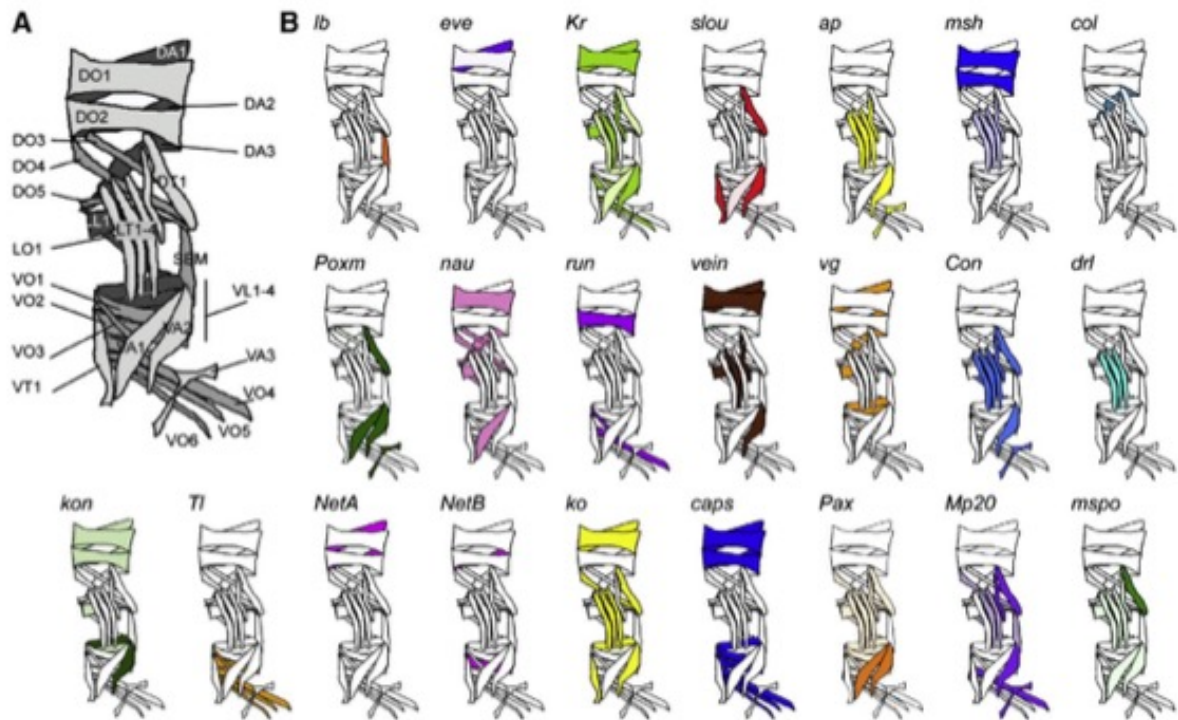
The muscle identity can be characterized by the position of the muscle within a segment and its morphology, defined by its number of nuclei, its orientation, and its number of fibers for instance. In the context of somatic muscle development, two types of identity can be distinguished: intra-segmental identity and inter-segmental identity. Intra-segmental identity defines the identity of one muscle within a segment. It is regulated by specific TFs known as identity factors, which exhibit differential expression patterns within each FC. On the other hand, inter-segmental identity refers to the identity of muscles across each segment and is under the control of Hox proteins. For instance, some muscles may or may not be present depending on the segment or can have different morphological properties.

#### **B.3.a. The intra-segmental identity**

The intra-segmental identity of adult muscles is controlled by the differential expression of TFs codes in specific muscle FC during development. Studies has revealed that unique molecular identities of each muscle result from the combinatorial expression of multiple genes with overlapping domains of expression referred to as identity factors (Bourgouin et al., 1992). In the *Drosophila* embryonic somatic musculature, twenty-three identity factors have been identified (**Figure 20**) (Tixier et al., 2010; Dobi et al., 2015). These identity factors are TFs expressed in specific muscle progenitors or FCs. Loss-of-function mutations in these TFs lead to defects in the specific muscles where they are normally expressed. They have been shown to regulate muscle morphogenesis, including size, shape, orientation, attachment sites, innervation, and number of fusion events.

For example, research has revealed the role of the segmentation gene Krüppel, a zinc finger TF expressed in a subset of muscle progenitor (Ruiz-Gómez et al., 1997). Indeed, gain and loss of Krüppel expression in sibling FC cells have the potential to alter the fate of the progenitors and the muscles they ultimately form. Krüppel mutants can lead to modifications in muscle morphology, orientation, or even complete muscle loss. These phenotypes are due to the loss of the Krüppel's targets in muscle progenitors, such as Slouch. Indeed, in Krüppel





**Figure 20 - Restricted gene expression patterns in somatic muscles**

Figure from Tixier et al., 2010

A/ Schematics of the larval muscle pattern of abdominal segment A2–A7 (Internal muscles in dark grey, intermediate muscles in grey, and external muscles in light grey).

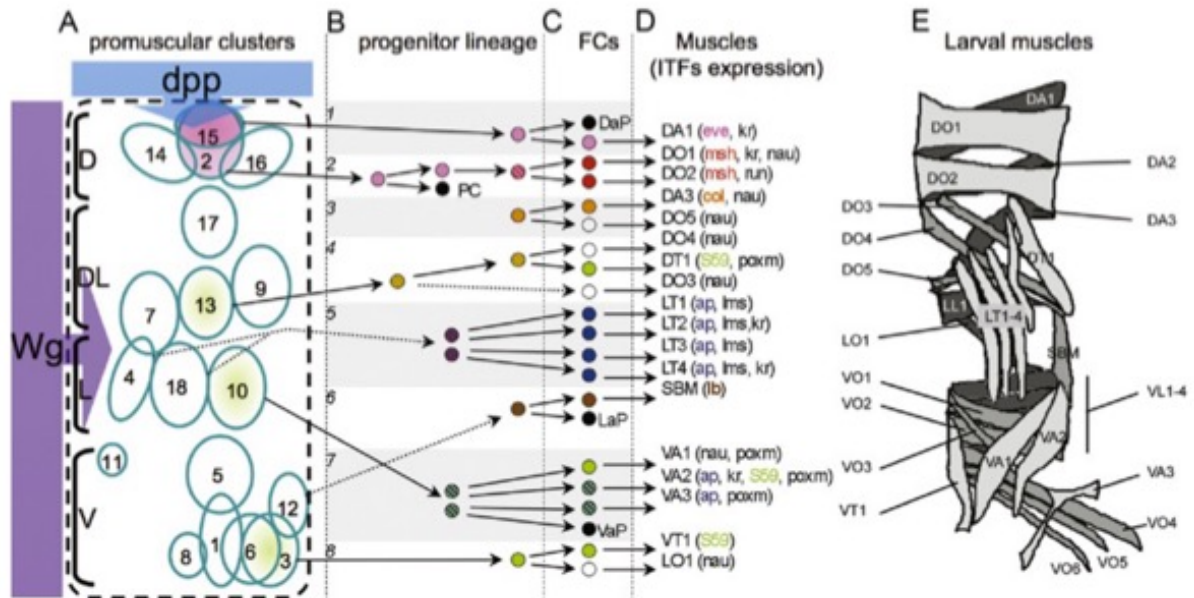
B/ Expression patterns of genes expressed in a muscle subset (light coloring corresponds to transient expression or lower expression levels). Unique molecular identities of each muscle arise through the combinatorial expression of several genes with overlapping domains of expression called identity factors. Twenty-three identity factors have been identified in the *Drosophila* embryonic somatic musculature.

mutant, the expression of Slouch is lost, and the Slouch positive muscle is duplicated. This study underscores the critical role of Krüppel in maintaining the expression of other genes in muscle progenitors, which is essential for muscle development and identity.

Moreover, it is the combinatorial expression of these identity factors in muscle progenitors that allow to define the identity of specific muscle. Indeed, the precise mutation of these factors leads to loss, duplication, or morphology modifications of its muscle sub-populations. For example, the loss of function of both Collier and Nautilus lead to different modifications of the DA3 muscle (Enriquez et al., 2012). Each of these genes control different properties of the DA3 muscle development. This TF code of the muscle progenitor is propagated into the nuclei of the naive myoblasts that fuse to form the final muscle fiber. In this way, all the nuclei of the muscle fiber express the same TF code. The sequence of expression and combinatorial activities of Collier and Nautilus control the pattern and morphology of specific muscles. It is the combinatorial pattern of expression of the identity factors in each segment that is necessary for the acquisition of specific muscle characteristics (**Figure 21**). However, while all the studies on the identity factors code show their role in the intra-segmental identity, their effectors remain unknown.

### **B.3.b. The inter-segmental identity**

The expression of specific Hox genes within the somatic mesoderm is segment specific and determine the positional identity of myoblasts (LaBeau et al., 2009; Lehmacher et al., 2012). For example, Sex combs reduced (Scr) is specific to the first thoracic segment (T1), Antennapedia (Antp) to the second (T2) and third (T3) thoracic segments, Ultrabithorax (Ubx) to the abdominal segments A1 to A7, Abdominal A (Abd-A) to the segments A1 to A8, and Abd-B for segments A8 and A9. The Hox genes control the somatic muscle with an autonomous function within the mesoderm. Indeed, studies show that mutations of Hox genes can change the posterior identity of one segment into the identity of a more anterior segment. In the opposite, their overexpression led to the transformation of an anterior segmental identity into a posterior identity (E. B. Lewis, 1978; Bender et al., 1983).



**Figure 21 – Expression of identity factors code in FC to determine specific muscles**

Figure from Joussineau et al., 2012

A/ Early patterning by *wg/dpp* and others induce definition of four domains in the presumptive somatic mesoderm: D (dorsal), DL (dorsolateral), L (lateral), and V (ventral).

B and C/ Probable (plain arrows) or supposed (dotted) connections between promuscular clusters, progenitors (P), and founder cells (FCs).

D/ Identity factors expressed in muscle fibers.

E/ Schematic representation of the larval somatic muscles.

Moreover, Hox genes act on the expression of other factors and of the inter-segmental identity at progenitor stage by influencing the selection of the progenitor cell from the promuscular cluster (Enriquez et al., 2012). For example, the expression of two identity factors Nautilus and Collier at very early stages in progenitor cells is under the control of Hox genes. Indeed, despite that Collier is expressed in the promuscular cluster of the DA3/DO5 muscles in each segment, the DA3 muscle is formed only in the T2 to A7 segments. Indeed, in Antp mutant, a high-level expression of Nautilus and Collier in one progenitor in T1 is observed, a pattern normally restricted to T2 and T3. In the same way, mis-expression of Ubx resulted in high-level expression of Nautilus and Collier in two progenitors in T1-T3 segments, a pattern normally typical of A1-A7 segments.

Furthermore, in addition to their role in the selection of the FC, the Hox genes can modulate the properties of each muscle depending on the segment. For example, the DA3 muscle variate in size between segments with a smaller size in T2 and T3 segments than in A1 to A7 segments, directly link to a variation in nuclei number. This variation in nuclei number is link to the activity of Hox genes in FC (Enriquez et al., 2012). The Hox genes contribute to the expression of the combinatory TFs code controlling the specific development of each muscle. However, the nature of the cellular and molecular events acting downstream these codes involved in the specific muscle identity remain undiscovered.



## THESIS PROJECT

In appendages, precise movement relies on the unique morphology of muscles. The study of muscle development has been a cornerstone in scientific research, providing profound insights into the intricate processes governing muscle morphology. As shown in the introduction of this thesis, *Drosophila* adult leg offers a structural and developmental analogous model to vertebrates upper and lower limbs, for unraveling the complexities of muscle morphologies development. Each adult leg contains 14 leg muscles produced from myoblasts localized in the developmental structure called imaginal disc. Each muscle has a unique morphology that is stereotyped between individuals. While the general molecular program of myogenesis is well characterized, the TFs codes and their effectors, controlling individual muscle morphology, identity remain unknown. The central objective of this thesis resides in delineating the molecular mechanisms underpinning the diversity observed in adult muscle morphologies. This thesis is organized around one problematic: Is there, in parallel with the general myogenesis program, a specific muscle program that governs unique muscle morphologies? During my thesis, I seek to address this problematic by challenging the hypothesis that leg myoblasts form heterogeneous population that can be subdivided in different subpopulations expressing different gene networks during development. To investigate this, I employ a multifaceted approach that integrates genetics, confocal microscopy, single-cell sequencing, image post-processing and bioinformatic. This study has led to a significant discovery regarding the genetic program that governs the diversity of adult muscle morphologies.

First, we employ genetic tools, immunostaining, and bioinformatics to characterize the cellular and molecular mechanisms underlying the development of adult leg muscles. Our investigation builds upon the prior observations made by Miller (1950) and Soler et al. (2004) to comprehensively identify all the muscles in the T1 adult prothorax leg. The mechanisms controlling the morphological diversity development of the leg musculature system remain unknown. To understand when myoblasts become specialized, we study myoblast growth in the leg disc at different developmental stages, ranging from early larval development to stages just before the fusion process. Our findings revealed that at 72 hours after egg laying (AEL),



myoblasts position dorsally and split into two distinct groups: one in postero-central position and another one in postero-peripheral location. Then, starting from 96h AEL, myoblasts gradually become specialized from one distal population and one proximal population to the level of the muscle that we study with the example of the tibia levator muscle (tilm) development in the femur. These results suggest that myoblasts undergo a progressive specialization process throughout development. We confirm these cellular observations by studying myoblasts transcriptomes during development thanks to single-myoblast sequencing. The distal myoblasts population is defined by the code of TFs Lim1, Sox100b, *bi* and *ko*, while the proximal myoblasts population express *unpg* and *hth*. These analyses confirm that myoblasts form a heterogeneous population with molecular differences.

We succeed to identify the evolution of different clusters obtained in our single cell by distinguishing the two populations, distal and proximal muscles, that evolved through time. We focus on the myoblasts sub-populations on the center of the disc that will generate the more distal muscles in the adult (muscles in the tibia and femur segments). This central sub-population will divide and give rise to four clusters. One will generate tibia muscles (expressing genes like *bifid*) and another one will generate most of the femur segment muscles (identified by the gene code of *bifid*, *grain* and *shaven*). Then, a third cluster is specific to one muscle (identified by the code of TFs Lim1, Msh, *lbe* and *Doc2*), the *tilm*, part of the femur segment, that is already specified in our stage of analysis. Finally, the last cluster is not identified, and we think that it gives rise to a non-muscle population that is present since the beginning within the disc in the more central part of it. This population specifically expressed genes such as Sox100b and Lim1. We make the hypothesis that these cells, despite they are Twi/Mef2 positives, are not myoblasts but neural lamella-associated cells, newly identified cells discovered after sequencing of adult tarsus (Hopkins et al., 2022). Their role in leg disc development remains to be explored. Finally, at stage 108h AEL, we identified a transitory cluster that seems to be a transitory stage between the first time point stage to the later stages. In summary, these transcriptomic results confirm that sub-population of myoblasts are determined very early to give rise to specific muscle with an organization in clusters precisely localized in the future leg muscle position. We confirm these outcomes with the analysis of the expression pattern of different genes throughout the development.





Finally, we begin to employ computational tools, genetic techniques, and advanced microscopy techniques to assess the impact of genetic manipulations on muscle architecture. We are doing functional analysis by looking at morphological and locomotion phenotypes after abolishing the function with RNAi technique of candidate genes. Since TFs are key players in cell specifications, the good candidates for our phenotypical analysis will be TFs differentially expressed in our subpopulation of myoblasts. We will characterize the morphological characteristics of adult muscle such as volume, shape, orientation, or number of muscle nuclei. This part of the study will allow to find specific genes involved in the specification of the muscle identity to be able to find epistatic relationship between genes playing role in the development of myoblasts morphological identity.

In summary, my research findings reveal that myoblasts initially possess a naive state upon joining the epithelial cells of the leg disc. During early stages, myoblasts are specified into two subpopulations. One population, located in the center of the disc, will give rise to distal muscles, while the other one, at the periphery, will generate proximal muscles. Then at later stages, myoblasts giving rise to the same muscle are specified and spatially cluster together in the region prefiguring the position of the future adult muscle. The analysis of the single cell sequencing perform at different stages highlights the potential genes controlling this specification. The current analyze of the functional role of these genes in shaping muscle architecture will give us the intrinsic and extrinsic factors controlling this multistep regulation of muscle morphology.



## PAPER IN PREPARATION

### Spatio-temporal control of muscle diversity in *Drosophila leg*

Camille GUILLERMIN<sup>1</sup>, Mathilde BOUCHET<sup>1</sup>, Violaine TRIBOLLET<sup>1</sup>, Laurent GILQUIN<sup>1</sup>, Sergio SARNATARO<sup>1</sup>, Isabelle STEVANT, Guillaume MERCY<sup>2</sup>, Emeric TEXERAUD<sup>2</sup>, Anne LAURENCON<sup>1</sup>, Yad GHAVI-HELM<sup>1</sup>, Benjamin GILLET<sup>1</sup>, Sandrine HUGHES<sup>1</sup>, and Jonathan ENRIQUEZ<sup>1</sup>

<sup>1</sup>*Institut de Génomique Fonctionnelle de Lyon, ENS de Lyon, CNRS, Université Lyon 1, 46 Allée d'Italie 69007 Lyon, France*

<sup>2</sup>*Univ Lyon, Université Claude Bernard Lyon 1, Bioinformatic Platform of the Labex Cortex, 69008 Lyon, France.*

## INTRODUCTION

The wide range of muscle morphologies found across the animal kingdom enable intricate movements essential for various behaviors, such as feeding, mating, or even tasks like composing articles. Many animals rely on appendages for their locomotion. These appendages feature muscles with diverse morphologies that enable specific functions crucial for the precise movement of appendage segments. However, the precise mechanisms governing the development of these unique muscle morphologies remain unknown.

In vertebrates, muscle precursors originating from the hypaxial region of the dermomyotome delaminate and subsequently migrate to populate the limb bud<sup>1</sup>. Once within the limb bud, muscle progenitors undergo active proliferation and become determined to muscle production. These myoblasts then differentiate into myocytes, which fuse together to form multinucleated fibers<sup>2</sup>. As this differentiation process unfolds, myofibers align to generate muscle bundles prefiguring the future muscles. The overall progression of myogenesis is intricately regulated by a hierarchical cascade of transcription factors (TFs) expressed at various stages of development.



For example, during the delamination, muscle progenitors express TFs, such as Pax3/Pax7, while during the phase of migration they express TF such as Lbx1, Mox2<sup>3,4</sup>. It is only during the proliferative phase that muscle precursors express the transcriptional core of myogenesis, known as the MRFs (Myogenic Regulator Factors). These MRFs hold the unique capability to induce myogenesis when overexpressed in differentiated cells<sup>5</sup>. The two MRFs, MyoD and Myf5, are expressed during the proliferation and determination of the myoblast while Myogenin, Myf6 and MyoD are expressed during the differentiation of myocytes. Additionally, another crucial group of TFs in the broader myogenesis program consists of members of the Mef2 TF family, totaling four members, which work synergistically with MRFs to orchestrate the complex process of muscle development<sup>6</sup>.

In *Drosophila*, leg muscle development closely mirrors a pattern seen in vertebrate limbs. The adult *Drosophila* leg is structured into segments, each comprising 14 muscles, with multiple muscle fibers<sup>7</sup>. During embryogenesis, a specific group of mesodermal progenitor cells undergoes an asymmetric division. This division results in the formation of one founder cell (FC) at the origin of a larval muscle and one adult muscle precursor<sup>8</sup>. The adult muscle precursors, known as AMPs, colonize the leg disc, analogue structure to the limb bud in vertebrates<sup>9</sup>. Within the leg disc, they undergo a phase of proliferation during the larval stages. In early pupal stages, myoblasts fuse with one another, forming the adult leg muscles. The transcriptional control of myogenesis is relatively simpler in *Drosophila*. The analog of MRFs in *Drosophila* is the transcription factor Twist, which can induce myogenesis when overexpressed in embryonic epithelial cells<sup>10</sup>. AMPs has been described to express Zfh1, a TF, maintained in satellite cells of flight muscles<sup>11</sup>. Although this hasn't been demonstrated in leg myoblasts, Twist is known to regulate the expression of Mef2 in embryonic and wing myoblasts<sup>12</sup>.

While the general program of muscle development has been extensively studied and understood, the specific program that controls the diversity of muscle morphology in multifiber muscles remains unknown. Is there in parallel to the general program of myogenesis a specific program controlling muscle morphology? Several studies on the body wall muscles of the *Drosophila* larva revealed that TFs expressed differentially in specific myoblasts control muscle morphologies<sup>13</sup>. It is important to note that larval body wall muscles are very different in terms of morphology and development compared to appendage muscles



in vertebrates and invertebrates. The larval body wall muscles are composed of a single muscle fiber, while leg muscles are tubular and striated multi-fiber muscles akin to vertebrate muscles<sup>14</sup>. The development of the body wall muscles is also very unique. *Drosophila* muscle pattern is seeded by a special class of myoblasts, the FCs, which display the unique property of being able to undergo multiple rounds of fusion with another class of myoblasts, the Fusion Competent Myoblasts (FCMs)<sup>15</sup>. FCs originate from the asymmetric division of progenitors cells (PCs) that are specified at fixed positions within the somatic mesoderm<sup>16</sup>. How are multifiber muscles in leg morphologically specified?

In this study, we have comprehensively outlined the architecture of adult muscles within the T1 leg and prothorax in a 3D context, showcasing the diverse range of muscle morphologies. Notably, we identified and described two previously undocumented muscles in the prothorax and coxa leg segment. Our investigation also delved into the behavior of Adult Muscle Precursors (AMPs), referred to as myoblasts in this study. We observed that approximately 20 Twist+ AMPs colonize the epithelial cells of the leg disc from the dorsal region. Over time, these myoblasts organize into two distinct populations within the disc: one central and another peripheral.

During the late larval and early pupal stages, the myoblasts from the central population form morphologically distinguishable subpopulations. To ascertain whether myoblasts are specified before the fusion process and, if so, when they are specified to produce specific muscles, we generated myoblast clones at various stages during embryogenesis and larval development. Our findings revealed that myoblasts enter the disc in a naive state and become specified during mid-larval stages to produce proximal versus distal muscles. The proximal and distal muscles exhibit clear boundaries, with distal muscles corresponding to those in the femur and tibia, while proximal muscles correspond to those in the prothorax, coxa, and trochanter. Lineage tracing of myoblasts demonstrated their progressive specification to produce subpopulations or individual muscles.

Single-cell sequencing at different time points allowed us to discover the transcription factors (TFs) that progressively specify myoblasts into specific muscles. The initial step of myoblast specification occurs at 96 hours after egg laying, where myoblasts segregate into two groups





expressing distinct sets of TFs. The central group expresses a combination of Lim1, Bi, Ko, and Sox100B, whereas the peripheral group expresses Hth and unpg. Both groups are localized in the presumptive territory of the tibia/femur and prothorax/coxa trochanter, respectively. Utilizing a myoblast tracing system activated by drivers expressed in these groups, we revealed that both central and peripheral groups of myoblasts prefigure muscle positions and are specified to produce muscles in the tibia/femur and prothorax/coxa/trochanter respectively.

Subsequent single-cell sequencing at later time points, focusing on the specification of myoblasts in the central group, showed that they are progressively specified during late larval and early pupal stages into subgroups of myoblasts expressing different combinations of TFs, giving rise to unique or subpopulations of muscles. These myoblasts subgroups give rise to three muscles in the tibia, three muscles in the femur, or a single muscle. These groups are localized in the presumptive territory where the future muscles will form. The groups of myoblasts giving rise to tibial muscles express bi, while the group giving rise to tilm, tidm, and tirm express sv, bi and grn, and the group giving rise to tilm express Lim1, Msh and Doc2.

Collectively, these results demonstrate that myoblasts are progressively specified before the fusion process, giving rise to specific muscles. They express a distinct set of TFs parallel to the general core of myogenesis. Currently, we are focused on determining the function of these TFs and understanding how these codes are established during development.

## **RESULTS**

### **Diversity of muscle morphologies in the T1 thoracic segment**

The muscle architecture of the T1 leg has been previously documented in earlier studies. In this study, we have chosen to adopt the nomenclature used by Soler et al. (2004)<sup>7</sup>, which is based on muscle function. However, it's worth noting that the muscles responsible for leg movement, particularly for coxa movement, are situated in the prothoracic region of the thorax. The architecture of the muscles in this prothoracic region has not been reexamined



since Miller's work in 1950<sup>17</sup>. This region contains a total of eight muscles responsible for head movement and six muscles responsible for coxa segment movement.

In this research, we have reexamined the architecture of the muscles located in the thorax responsible for moving the prothoracic leg. We have developed a new protocol for dissection, fixation, and mounting, allowing us to label and visualize the architecture of the previously described eight muscles. These muscles have been renamed according to the new nomenclature proposed by Soler et al. (**Figure 1A**). Additionally, during our investigation, we identified two previously undocumented muscles in the prothorax (the coxa levator muscle 2, *clm2*) and the coxa (the trochanter reductor muscle 2, *trrm2*) (**Figure 1B**). The mechanisms specifying such muscle diversity in terms of morphologies remain unknown.

### **Description of the myoblast behavior during development.**

In our pursuit of comprehending how muscle diversity emerges during development, we have delved into a detailed examination of myoblast behavior at the cellular level within the leg imaginal disc (**Figure 1I**).

Our investigation commences in the late embryo stage when the epithelial cells of the leg disc have already evaginate from the ectoderm. At this stage, approximately 20 myoblasts, marked by Twist expression, are closely associated with the dorsal region of the Dll+ epithelial cell (**Figure 1C**). As time progresses, by 72 hours after egg laying (AEL), around 50 myoblasts begin to colonize the dorsal region of the disc (**Figure 1D**). By 84h AEL, around 60 myoblasts predominantly occupy the dorsal region of the disc but are distinguishable as two distinct populations: one postero-central and one more postero-peripheral (**Figure 1E**). By 96h AEL, the number of myoblasts has drastically increased (~450), and they have entirely colonized the leg imaginal disc, forming two distinct clusters: one in the central region and another in the peripheral region (**Figure 1F**). Subsequently, the number of myoblasts remained constant until the formation of pupae. During the late larval stages, around 120h AEL and the white pupal stage, 0 hours after pupa formation (APF), the central clusters of myoblasts subdivide into smaller clusters forming one layer of myoblast in close contact with epithelium. At 5 hours APF, the tarsal segment and the tibia undergo evagination, while the femoral segment partially protrudes. The coxa/trochanter and prothorax segments remain unevaginated (**Figure**



**1G, and 1H**). Notably, the myoblasts corresponding to the central cluster observed at 96h AEL form groups that might prefigure the future muscle positions in their respective segments. Myoblasts at the outermost edges of the clusters make contact with the apodeme, a further tendon corresponding to the invagination of the epithelium, as described in Soler's work. The myoblast localized in the prothorax and coxa/trochanter segment remain organized in two rings.

### **Myoblasts are progressively specified to produce specific muscles**

The specification of leg AMPs for the production of specific muscles may initiate during embryogenesis. These myoblasts originate from progenitors that divide once to produce a FC cell already determined to generate a particular larval muscle fiber and an AMP. It is conceivable that AMPs follow a similar developmental trajectory than the FC and acquire their specifications during this phase. Another possibility is that myoblasts remain undifferentiated until they infiltrate the disc. Peter Lawrence's 1982 research suggests that myoblasts possess during early larval stages the potential to generate all muscles in the prothorax. Lawrence generated minute clones during early larval stages and analyzed the progeny of these myoblasts only in the thorax since the legs were removed during the experiment.

To pinpoint when leg AMPs are specified to produce specific muscles, we generated mCherry+ myoblast clones at various developmental stages and dissected adult flies to determine which muscles these clones produced (**Figure 2A**). First, we generated clones during the second half of embryogenesis, 12-24 hours after egg laying. In our examination of 900 flies, we identified 17 clones in the prothoracic segment. The majority of these clones (70%, N=12) labeled all leg muscles across all leg segments, including the prothorax (a similar proportion was observed in the metathoracic and mesothoracic segments, although not detailed here) (**Figure 2B**).

Interestingly, in some cases (N=4), these clones extended into the contralateral hemisegment, a phenomenon also noted by Peter Lawrence in 1982<sup>18</sup>. Notably, the intensity of muscle labeling in the contralateral segment was relatively weak. These observations are unlikely to result from a double event (two independent clones) due to the infrequency of such events and their absence in the metathoracic and mesothoracic segments. The midline crossing likely



occurs specifically in the prothoracic segment probably due to the physical contact between the T1 leg discs.

Both observations, where clones label all muscles in the prothoracic segment and have the potential to cross the midline boundary, strongly suggest that myoblasts are naive before colonizing the leg disc.

We also observed two distinct types of myoblast clones that were labeling muscles in the femur and tibia leg segments (18%, N=3) or muscles in the trochanter, coxa, and prothorax (12%, N=2). These findings may suggest a couple of possibilities. One hypothesis is that myoblasts begin to be specified for producing these two subpopulations of muscles embryogenesis, implying that the majority of myoblasts remain uncommitted are in process of being specified. Another scenario is that myoblast specification occurs at later developmental stages.

To differentiate between these two scenarios, we generated myoblast clones during larval stages. Similar results were obtained compared to embryonic induced clones when clones were generated at 24h AEL (N=26, labeling trochanter, coxa, prothorax: 17%, N=4; Femur and Tibia: 13%, N=3; all: 73%, N=19) and 48h AEL (N=39, labeling trochanter, coxa, prothorax: 18%, N=4; Femur and Tibia: 13%, N=3; all: 70%, N=19). However, when clones were generated at 84h and 96h AEL, we did not observe any myoblast clones labeling all muscles (N=2000, N=1600 respectively). Instead, the majority of the clones labeled either the trochanter, coxa, and prothorax (84 h AEL: 47%, N=18; 96h AEL: 48%, N=23), or the tibia and femur (84 h AEL: 21%, N=8; 96h AEL: 27%, N=13). Interestingly, at these time points, we identified clones that labeled only a subset of muscles in the femur (84h AEL: 31%, N=12; 96h AEL: 25%, N=13). Further examination through confocal analysis revealed that these clones were, in fact, primarily two clones labeling three muscles in the femur or one single muscle (the tilm) (**Figure 2C**).

In summary, these findings collectively indicate that myoblasts are undifferentiated upon their entry into the disc and undergo specification in two populations to produce distal muscles (found in the tibia and femur) and proximal muscles (located in the prothorax, coxa, and trochanter) after 82 hours AEL. These distinct populations may correspond to myoblasts localized in the center of the disc versus those situated in the periphery. The presence of two





types of clones at 84 and 96 hours AEL, one labeling three muscles and another labeling a single muscle in the tibia, suggests that myoblasts undergo gradual specification to generate subpopulations of muscles and individual muscles.

### **The initial step in muscle specification involves defining distal muscles versus proximal muscles.**

We began by investigating how myoblasts are progressively specified to produce unique muscles by conducting single-cell sequencing (SCS) of leg myoblasts at various time points (**Figure 3A**).

The earliest feasible time point for sufficient disc dissection before Myoblast GFP+ FACS sorting was at 96 AEL. The SCS analysis unveiled two distinct clusters of myoblasts (Clusters 0 and 1) (**Figure 3B**). To recapture the spatial information lost during FACS sorting, we utilized GFP expression under the control of enhancer or gene trap lines specifically active in one or the other cluster. We observed that myoblasts located at the disc's center exhibited GFP expression when driven by genes expressed in cluster 0 (such as *tnc*, *bi*, *Lim1*, or *Ko*) (**Figure 3C**), while those at the disc's periphery displayed GFP expression with genes expressed in cluster 1 (like *unpg*) (**Figure 3D**).

Immunostaining against *Bi*, *Lim1* (for the center of the disc) or *hth* (for the periphery of the disc) confirmed that clusters 0 and 1 correspond to two populations differentially localized spatially. We then labeled the peripheral part of the epithelium giving rise to proximal segments (prothorax, coxa, and trochanter) and the central part of the epithelium giving rise to distal segments (femur, tibia, and tarsal segments) with *Hth* and *Dll/Dac*, respectively, to better characterize the position of the two myoblast clusters compared to the epithelium (**Figure 3E**). The central cluster localized in the *Dll+* and *Dac+* regions, while the peripheral cluster localized in the *Hth+* regions. If the position of the myoblasts prefigures the final position of the muscles, the peripheral and center clusters should be specified to produce distal leg muscles (femur, tibia) and proximal leg muscles (prothorax, coxa, and trochanter), respectively. This would hypothesis aligns with the two groups of myoblasts identified in our lineage tracing experiments.

To substantiate this hypothesis, we employed a lineage muscle tracing system to maintain



GFP expression until the adult stage. When we activated the lineage muscle tracing system using Gal4 under the control of *Lim1*, *tnc*, *bi*, or *ko* (cluster 0), GFP expression was observed in tibial and femoral muscle (**Figure 3F**). Conversely, activation of the lineage muscle tracing system with *unpg*-Gal4 (cluster 1) resulted in GFP expression in muscles of the prothorax, coxa, and trochanter (**Figure 3G**).

These findings revealed that myoblasts, expressing two different combinations of TFs, lose their potential to produce all types of muscles and are specified into two populations localized in the center and the periphery of the disc, giving rise to distal (femur, tibia) and proximal (prothorax, coxa, and trochanter) muscles, respectively. The position of these two populations prefigure the position of the future muscles and expressed two different combinations of TFs in parallel to the general myogenic TFs characterized at this stage by the expression of *Twist* and *Zfh1*. The central cluster express *Sox100B*, *ko*, *Lim1* and *bi* while the peripheral cluster express *unpg* and *hth*.

### **Myoblasts are progressively specified to produce specific muscles**

We conducted subsequent Single-Cell Sequencing at later time points on FACS-sorted GFP+ myoblasts, aiming to monitor the specification of the central cluster leading to the femoral and tibial muscles (**Figure 4A and 4B**). While the peripheral cluster also gives rise to distinct clusters, we did not analyze their specification within the scope of this article.

To recapture the spatial information lost during FACS sorting, we once again utilized GFP expression under the regulation of enhancer or gene trap lines that are specifically expressed in a particular cluster or subpopulation of clusters. To label clusters 4, 6, 13, 14, and 15, we employed *tnc-Gal4* to express GFP and utilized morphological markers (FAS III) or spatial markers (*Dll*, *Dac* and *Hth*) to label the epithelium and found that *tnc*-GFP labeled the myoblasts at the center of the discs, the presumptive territory of the tarsus, tibia, and femur (**Figure 4C**). We utilized a lineage muscle tracing system to maintain GFP expression until the adult stage, activating it after early pupal stages (0 hour APF). This resulted in the labeling of muscles in the tibia and femur. These findings indicate that myoblasts in the center of the disc at 96 hours After Larval Hatching (AEL) (cluster 6) retain their spatial localization in the presumptive territory of the femur and tibia at later stages, gradually specified into



subpopulations of myoblasts (clusters 4, 6, 13, 14, and 15). We applied similar strategies to determine the identity of individual central clusters.

First, we performed immunostaining against Bi and used *bi-gal4* to label myoblasts of Cluster 14 with GFP at 0 and 5 hours APF (**Figure 4D**). Based on these experiments, we revealed that myoblasts belonging to cluster 14 are localized in the presumptive territory of the tibia. We applied the lineage muscle tracing system to maintain GFP expression until the adult stage, activating it at early pupal stage using *bi-gal4* with *tub-gal80<sup>ts</sup>* (0 hour APF), and successfully labeled only tibial muscles.

We employed a similar strategy to determine the identity of cluster 13 by using *sv-gal4* specifically expressed in this cluster, concluding that myoblasts belonging to this cluster are localized in the presumptive territory of the femur and give rise to three muscles: *ltm2*, *tidm*, and *tirm* (**Figure 4E**).

To recover the spatial information of Cluster 6, we utilized an enhancer trap of *rtGEF* highly expressed in this cluster and specifically labeling a group of myoblasts in the leg disc (**Figure 4F**). To validate *rtGEF-Gal4* as a reliable marker for Cluster 6, we performed co-stainings between GFP under the control of *rtGEF-Gal4* and different TFs expressed in this cluster (*Lim1*, *Msh*, *lbe* or *Doc2*), as well as genes expressed in other clusters (*Sox 100B* or *Bi*). We then labeled both the proximo-distal and ventro-dorsal regions of the discs and found that the myoblasts corresponding to Cluster 6 are localized in the region with high levels of *Dpp* (dorsal region) and high levels of *Dac*, with no expression of *Dll* (proximal region of the femur) (**Figure 4F**). Based on these spatial and morphological markers (including *Fas III*), we conclude that the myoblasts of Cluster 6 are localized in the dorsal region of the future femur. If Cluster 6 prefigures the future position of a muscle, it should be specified to produce the *Tilm* (tibial levator muscle) localized in the dorsal region of the adult femur. To test this hypothesis, we utilized our muscle lineage tracing system with *rtGEF-Gal4* and were able to label only the *Tilm* in the adult.

Myoblasts belonging to cluster 15 express specifically *Sox100B* compared to clusters 6, 13, 14, and 15 (**Figure 4G**). We then used *tnc-gal4* to express GFP and performed co-staining with *Sox100B*. *Tnc+* and *Sox100B+* cells were detected in the presumptive region of the tarsus as well as in the femur and tibia. The localization in the tarsus suggests that these cells are not



myoblasts. While the *tnc*<sup>+</sup> cells in the tarsus do not express Mef2 protein as well as some *tnc*<sup>+</sup> cells in the presumptive territory of the tibia and tarsus we could capture these cells during the FACS sorting because the *mef2* driver used to label myoblasts is weakly expressed in these cells. Cluster 15 also expresses Lim1, which we confirm to be expressed in the *tnc*<sup>+</sup> cells in the tarsus, tibia and femur. Based on these markers, we postulate that these cells are a neural lamella-associated cell type, a new type of cell recently discovered after SCS of the adult tarsus, which are Lim1<sup>+</sup> and Sox100B<sup>+</sup>.

Finally, Cluster 4 is a transitory cluster mostly found at 108h AEL. The SCS trajectory analysis using PAGA strongly suggests that this cluster is a transitory cluster between cluster 6, mostly found at 96 hours AEL, and clusters 13, 14, and 15 (**Figure 5**).

In summary, the two populations of myoblasts giving rise to proximal vs. distal muscles are progressively specified into subpopulations of myoblasts or single muscles (**Figure 6**). This data demonstrates that the specification of muscle morphology is established very early, more than 24 hours before the fusion process. Following the specification of cluster 6, we showed that myoblasts specified to produce individual muscles are positioned very early during development in the future location of the adult muscle, and myoblasts specified to give rise to the same muscle fuse only between themselves.

## DISCUSSION

Here, through the integration of various techniques, we demonstrate the stepwise specification of myoblasts. Initially, myoblasts are specified into two distinct populations giving rise to proximal and distal muscles, localized in the center and periphery of the leg imaginal disc, respectively, during mid-larval stages. These populations exhibit unique transcription factor (TF) codes: the central population expresses Lim1, Bi, Ko, and Sox100B TFs, while the peripheral group expresses Hth and unpg. Subsequently, these populations are progressively specified into subpopulations of myoblasts, ultimately giving rise to either single muscles or groups of muscles. Spatial analysis of these myoblast subpopulations reveals that their positioning prefigures the future location of the adult muscles. Furthermore, our data elucidate that only myoblasts sharing a common transcriptional program fuse with one





another. Lastly, our results reveal that these subpopulations, in parallel the general program of myogenesis, express distinct combinations of TFs that may further specify myoblasts diversity.

The arthropod limb can be categorized into two main regions: the coxopodite and the telopodite<sup>19</sup>. The prevailing theory suggests that the telopodite, considered the 'true leg,' has evolved from the coxopodite, which is deemed to be part of the thorax. In the case of *Drosophila*, this theory finds support in numerous subsequent studies, indicating that the molecular signature of the coxopodites (coxa + trochanter) closely resembles that of the trunk, whereas the telopodite possesses its distinct molecular signature.

These molecular signatures, further studied through functional analysis of numerous genes, were only investigated within the epithelium of the leg disc. In our study SCS during mid larval stages and genetic experiments, focusing on derivatives of the mesoderm, revealed two distinct developmental regions corresponding the telopodia and the coxopodia/prothorax. These mesodermal derivative territories are established during mid-larval stages when the proximo-distal patterning of the epithelium is also established. Low-resolution analysis of myoblast single-cell clusters at alter time point highlights two major myoblasts clusters, corresponding to myoblasts in the telopodia versus the coxopodia/prothorax, implying that myoblasts within these structures follow distinct developmental trajectories.

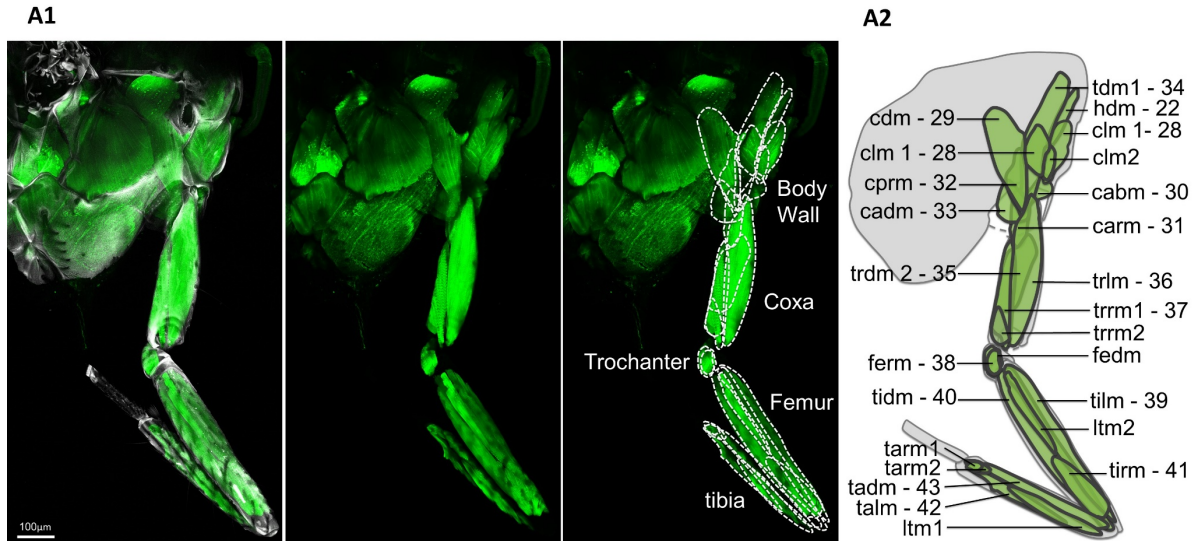
Understanding how derivatives of both ectoderm and mesoderm co-evolve to construct a functional structure is essential. One possibility is that the leg epithelium disc plays a pivotal role in controlling myoblast development. The differentiation of coxopodite and telopodite occurs during larval stages and is regulated by Wg and Dpp morphogens that define the proximo-distal axis. Regions receiving both morphogens are determined to become the telopodite (positive for Dll and Dac), while the other region develops into the coxopodite (positive for Hth). In our clustering analysis, Dll and Dac were not expressed in the myoblasts, and neither were Wg and Dpp morphogens. However, all of them expressed Wg and Dpp receptors. We are currently investigating if Dpp and Wg determine the central cluster of myoblasts.



The two populations of myoblasts then undergo further specification, differentiating into subpopulations that ultimately give rise to either groups or single muscles. We hypothesize that the ventro-dorsal patterning of the epithelium plays a pivotal role in this sub-patterning of myoblasts. For instance, myoblasts giving rise to the Tilm are located just beneath epithelial cells that secrete high levels of Dpp. Dpp morphogen defines the dorsal region of the leg, and the Tilm muscle is situated in the dorsal region of the femur. We are currently conducting experiments where we disrupt Dpp function after the establishment of the proximo-distal patterning, after 84 hours AEL it becomes independent of Dpp and Wg morphogens. Additionally, we are inhibiting the reception of Dpp signaling in myoblasts. Under these genetic conditions, we anticipate that the Tilm muscle may not be specified.

In conclusion, we hypothesize that first the proximo-distal patterning of the epithelium determines the two population of myoblasts giving rise to distal and proximal muscles. Furthermore, at later stages, the ventro-dorsal patterning of the leg further refines myoblasts, specifying them to produce specific muscles.





**Figure 1A – Adult leg muscle development**

**Drosophila adult leg muscles.**

A1/ Confocal picture of the T1 leg, in green the muscles (Mhc>GFP) and in gray the cuticle (each muscle is surrounded by a dotted line).

A2/ Schematic representation of the different muscles within each leg segments of the leg (tibia, femur, trochanter, and coxa) and within the thorax. Numbers used are the one named by Peter A. Lawrence (1982), the fedm, ltm2, ltm1, tarm1/2 by C. Soler (2004). The coxa levator 2 (clm2) in the prothorax and the trochanter reductor muscle 2 (trrm2) in the coxa are two new muscles identified in this study.

**Body Wall muscles:**

- 22 Head dorsal muscle (hdm)
- 28 Coxa levator 1 (clm1)
- 29 Coxa depressor muscle (cdm)
- Coxa levator muscle 2 (clm2)
- 30 Coxa abductor muscle (cabm)
- 31 Coxa anterior rotator muscle (carm)
- 32 Coxa posterior rotator muscle (cprm)
- 33 Coxa adductor muscle (cadm)
- 34 Trochanter depressor muscle 1 (tdm1)

**Trochanter muscles:**

- Femur depressor muscle (fedm)
- 38 Femur reductor muscle (ferm)

**Femur muscles:**

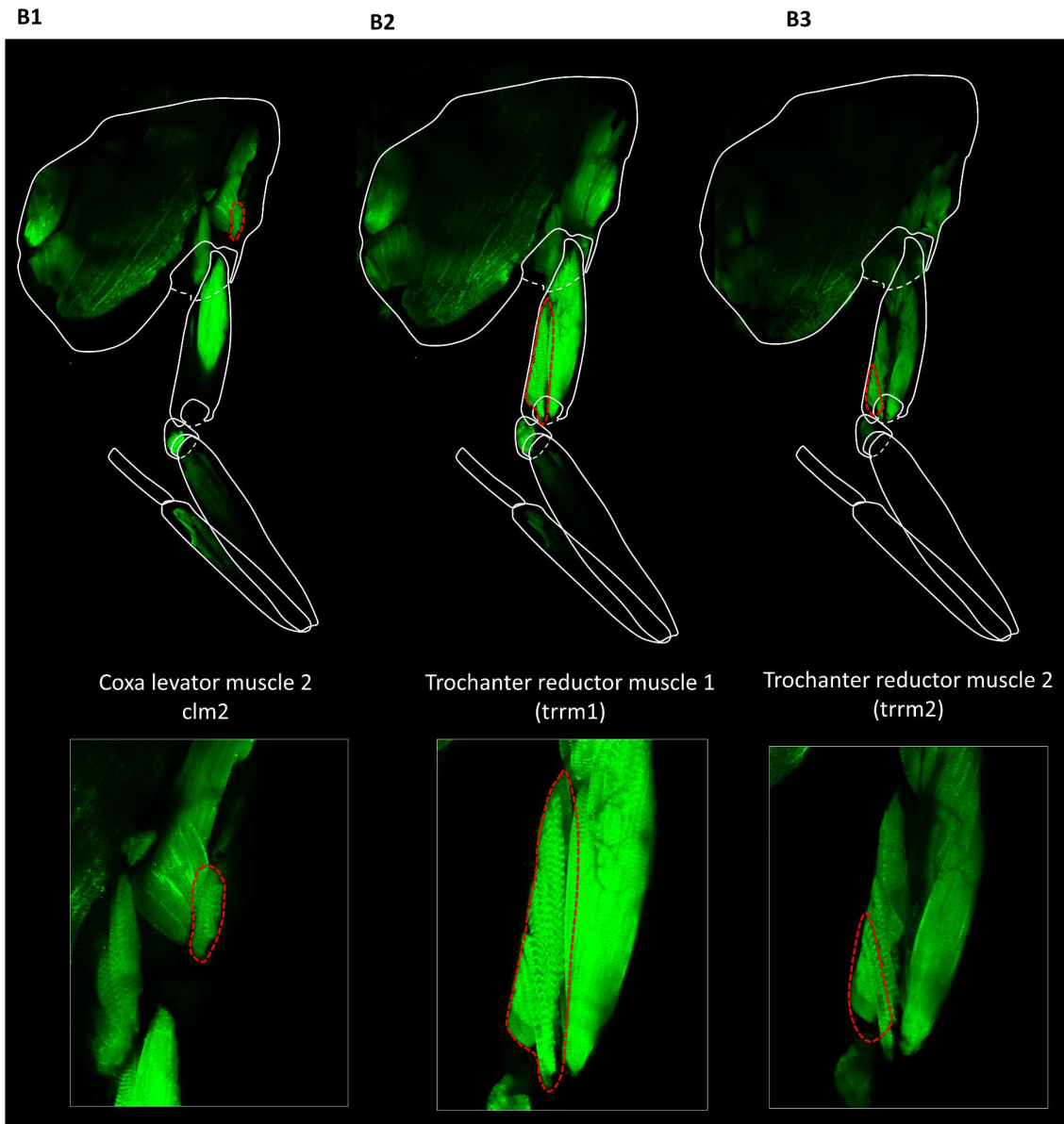
- Long tendon muscle 2 (ltm2)
- 39 Tibia levator muscle (tilm)
- 40 Tibia depressor muscle (tidm)
- 41 Tibia reductor muscle (tirm)

**Tibia muscles:**

**Coxa muscles:**

- 35 Trochanter depressor muscle 2 (trdm2)
- 36 Trochanter levator muscle (trlm)
- 37 Trochanter reductor muscle 1 (trrm1)
- Trochanter reductor muscle 2 (trrm2)
- Long tendon muscle 1 (ltm1)
- 42 Tarsal levator muscle (talm)
- 43 Tarsal depressor muscle (tadm)
- Tarsal reductor muscle 1 (tarm1)
- Tarsal reductor muscle 2 (tarm2)





**Figure 1B – Adult leg muscle development**

**Identification of undocumented muscles.**

B/ Confocal picture of the T1 leg, in green the muscles (Mhc>GFP).

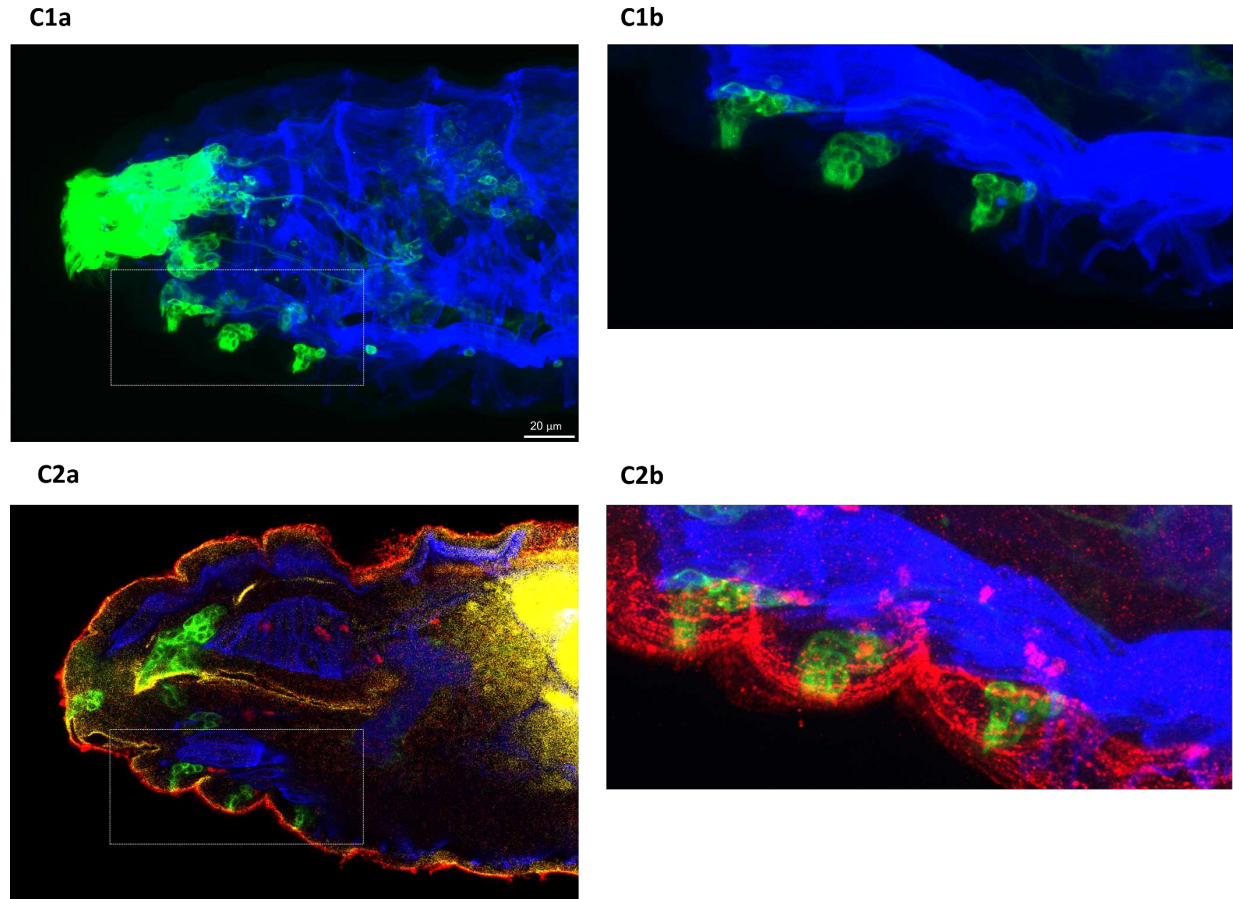
B1/ Identification of one undocumented muscle in the prothorax: the coxa levator muscle 2 (clm2).

B2/ Previous documented muscle in the coxa, renamed after this study: the trochanter reductor muscle 1 (trrm1).

B3/ Identification of one undocumented muscle in the coxa: the trochanter reductor muscle 2 (trrm2).







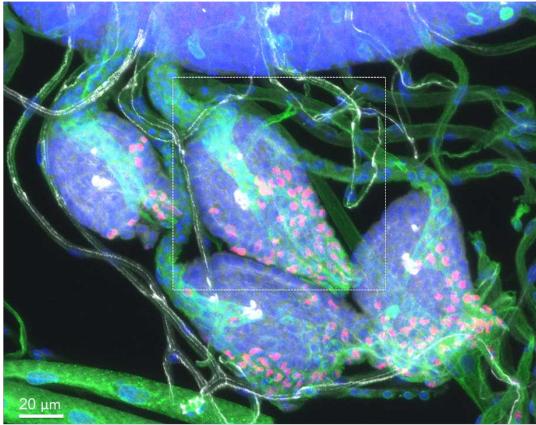
**Figure 1C – Adult leg muscle development**

**Embryo Stage: myoblasts are localize near the future leg disc.**

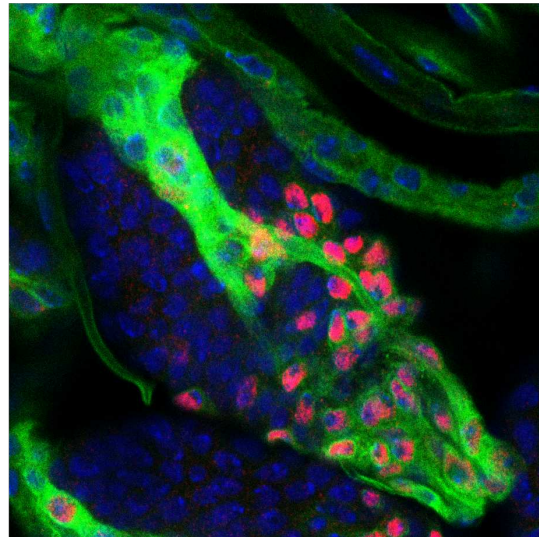
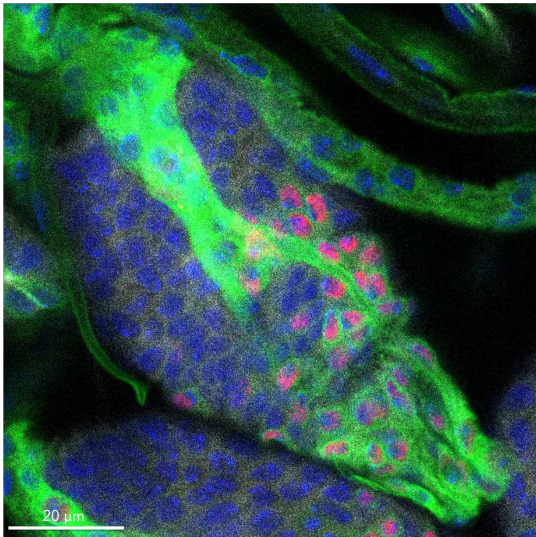
C/ Confocal pictures of embryo stage 16. In blue, the larvae body wall muscles (Phalloïdin), in green the epithelial cells (DII+), in red the myoblasts (Twist+) (C2). The zoom in C1b and C2b are showing the T1, T2 and T3 segments with labeled in green the epithelial cells of the leg disc that have already evaginate from the ectoderm. Around 20 myoblasts, Twist+, are close to the dorsal region of the epithelial cell.



**D1a**



**D1b**



**Figure 1D – Adult leg muscle development**

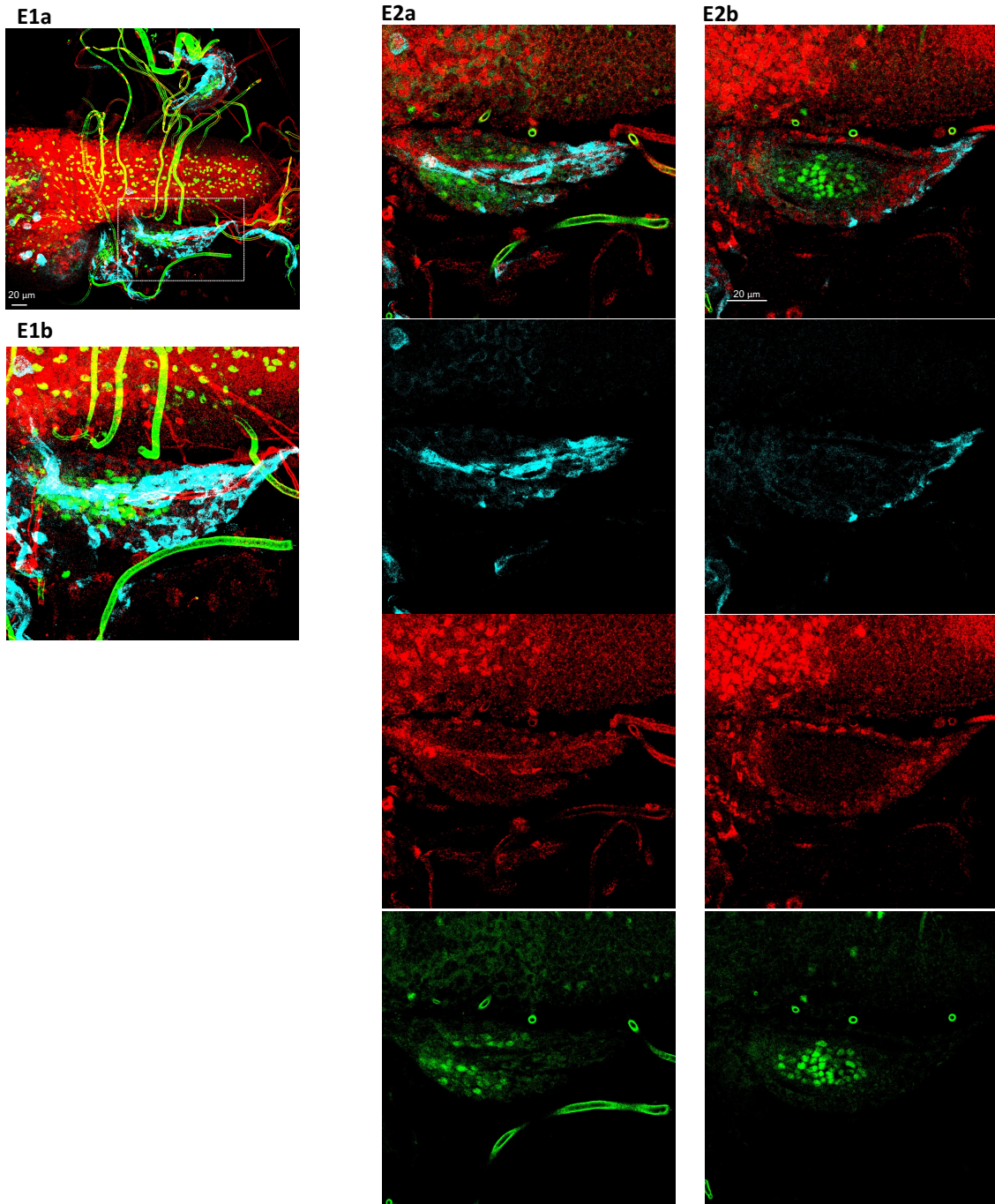
**72h AEL: myoblasts begin to colonize the disc.**

D1a/ Confocal pictures of the four leg discs (2 T1 and 2 T2) attached to the VNC, at 72h AEL (in white the epidermis labeled with Fas3 antibody).

D1b/ Zoom on one T1 leg disc of D1a. Myoblasts are labeled in green with the transgene 24B>GFP and in red with Twist antibody (in blue nuclei with DAPI). Around 50 myoblasts begin to colonize the dorsal region of the developing leg disc.





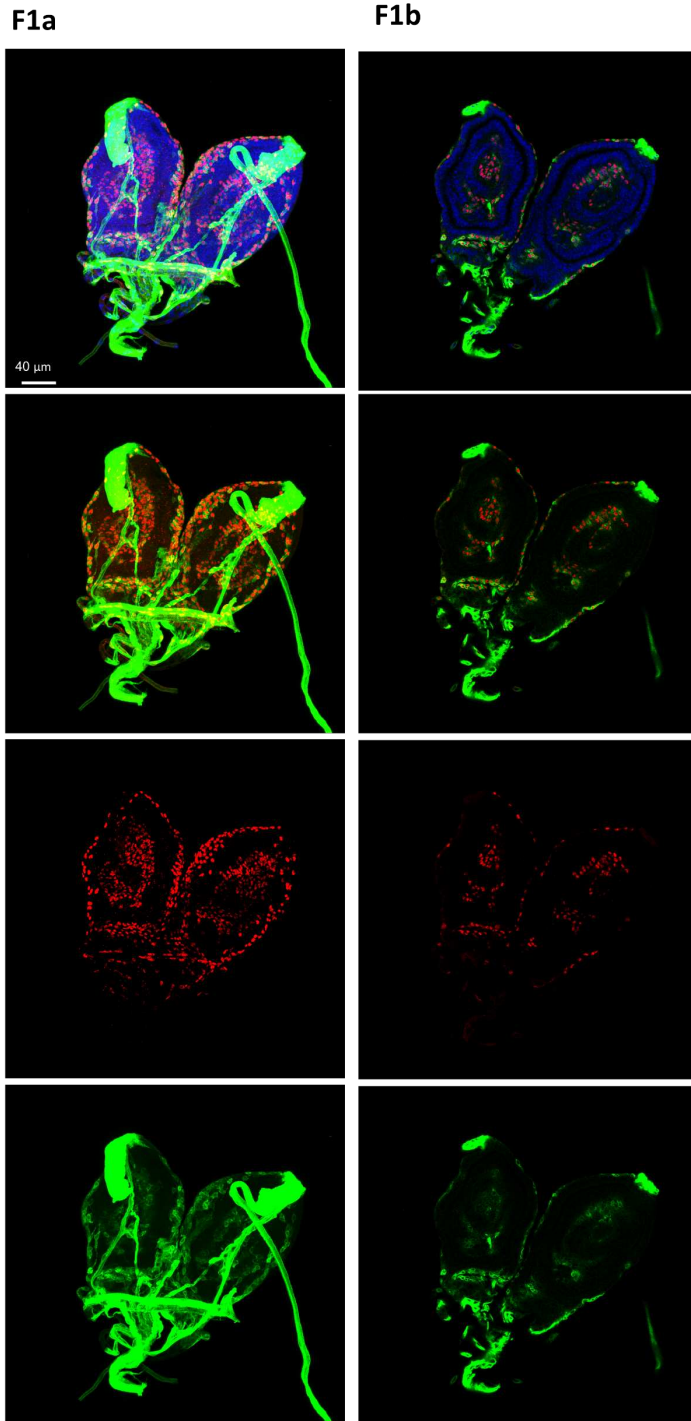


**Figure 1E – Adult leg muscle development**

**84h AEL: myoblasts organized in two populations.**

E/ Confocal picture of a leg disc attached to the VNC at 84h AEL (E1a) and zoom on the leg disc (E1b). Immunostainings are used to show the spatial organization of myoblasts: in red, hth labeled the peripheral part of the disc, while in green, Dll labeled the center part of the disc. Myoblasts, in cyan, are organized in two populations: one postero-peripheral and postero-central. E2a/ Z-projection of the leg disc. E2b/ Cross section of the leg disc.





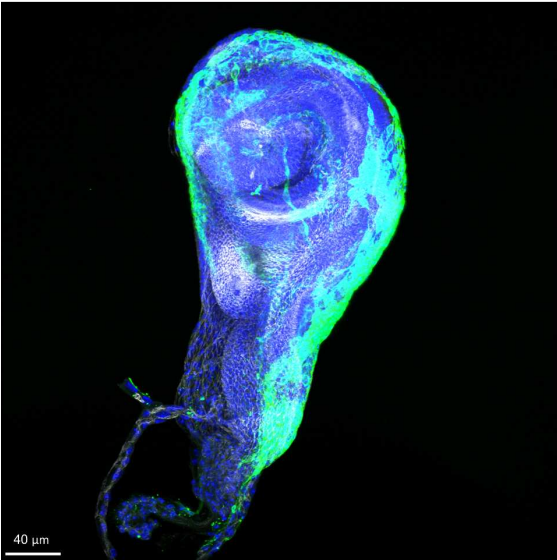
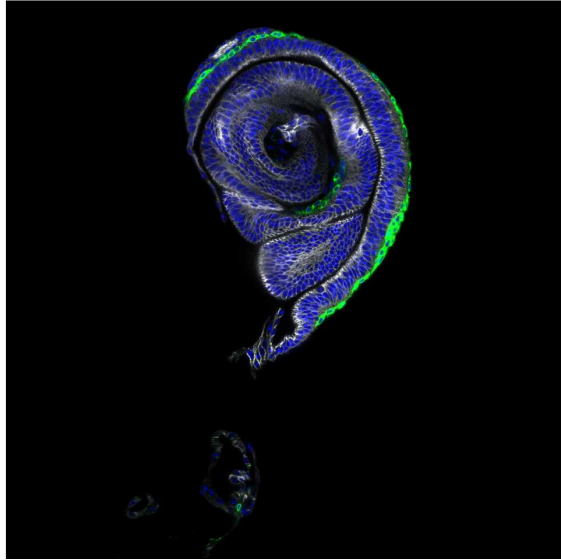
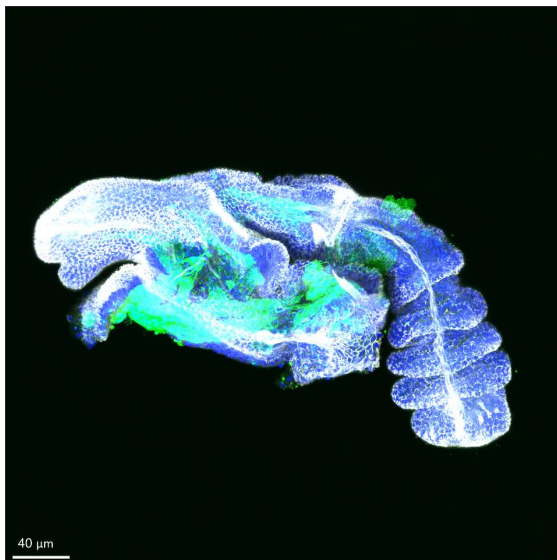
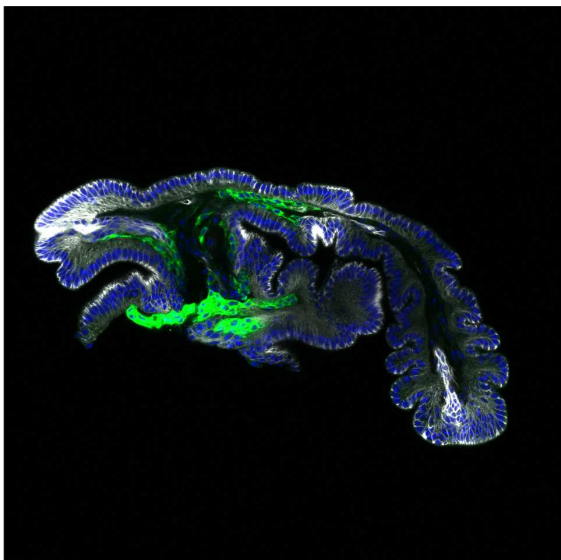
**Figure 1F – Adult leg muscle development**

**96h AEL: myoblasts proliferation and clustering.**

F/ Confocal picture of two T1 leg disc, at 96h AEL, attached together (F1a z-projection, F1b z-stack). Myoblasts are labeled in green with the transgene 24B>GFP and in red with Twist antibody (in blue DAPI). Around 450 myoblasts have colonized the leg disc, forming two distinct clusters: one in the central region and another in the peripheral region.

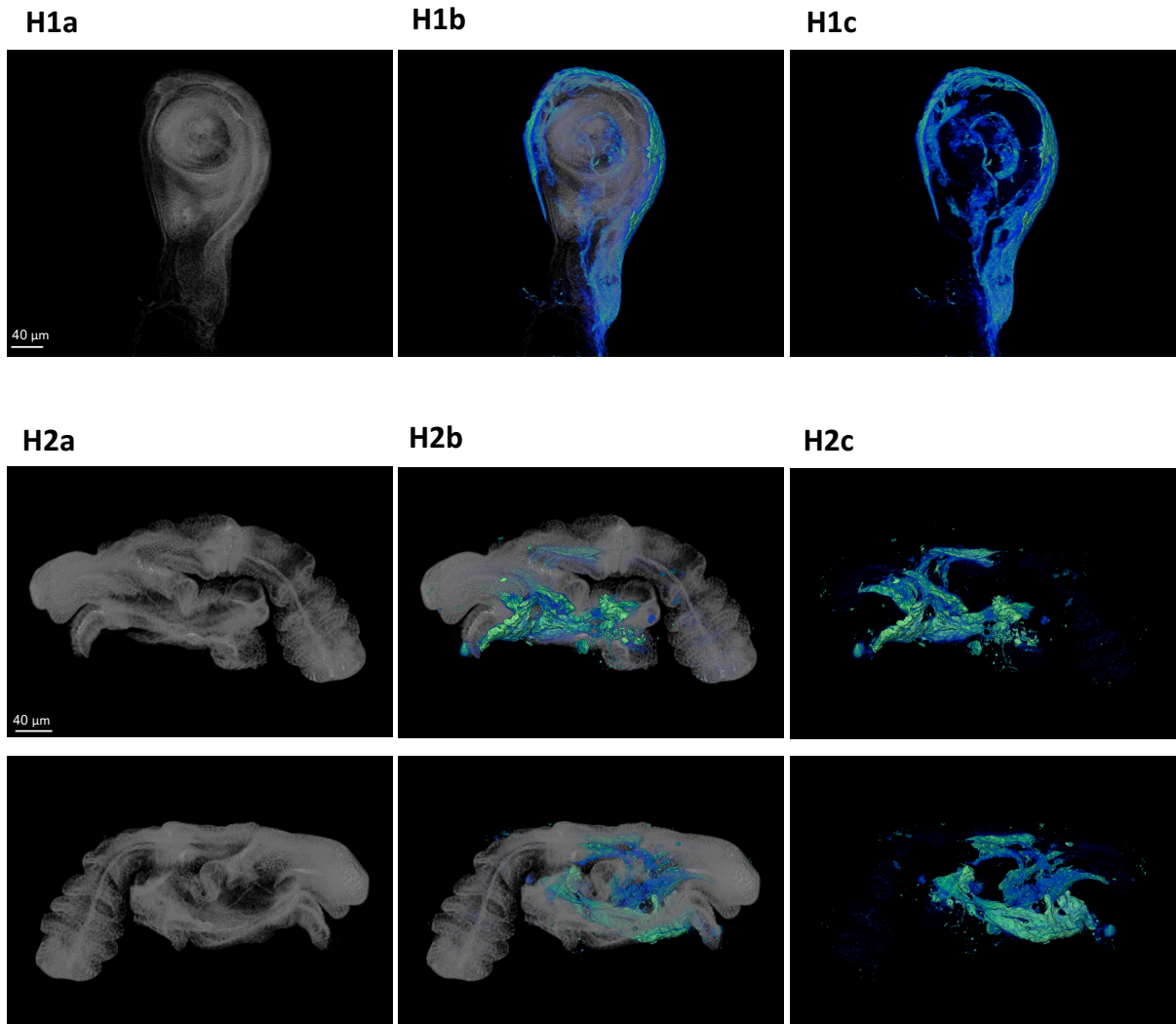




**G1a****G1b****G2a****G2b****Figure 1G – Adult leg muscle development****Metamorphosis: leg disc evagination.**

G/ Confocal picture of a T1 leg disc, at 0h APF (G1) or at 5h APF (G2) with in green the myoblasts (Mef2>GFP) and in white the epithelium (Fas3 staining) (in blue the nuclei with DAPI) (G1a and G2a z-projection, G1b and G2b z-stack). Myoblasts are subdivide into two clusters: one peripheral forming one layer in contact with the epithelium, another one in central position. At 5hAPF, the leg disc begins to evaginate from its center.



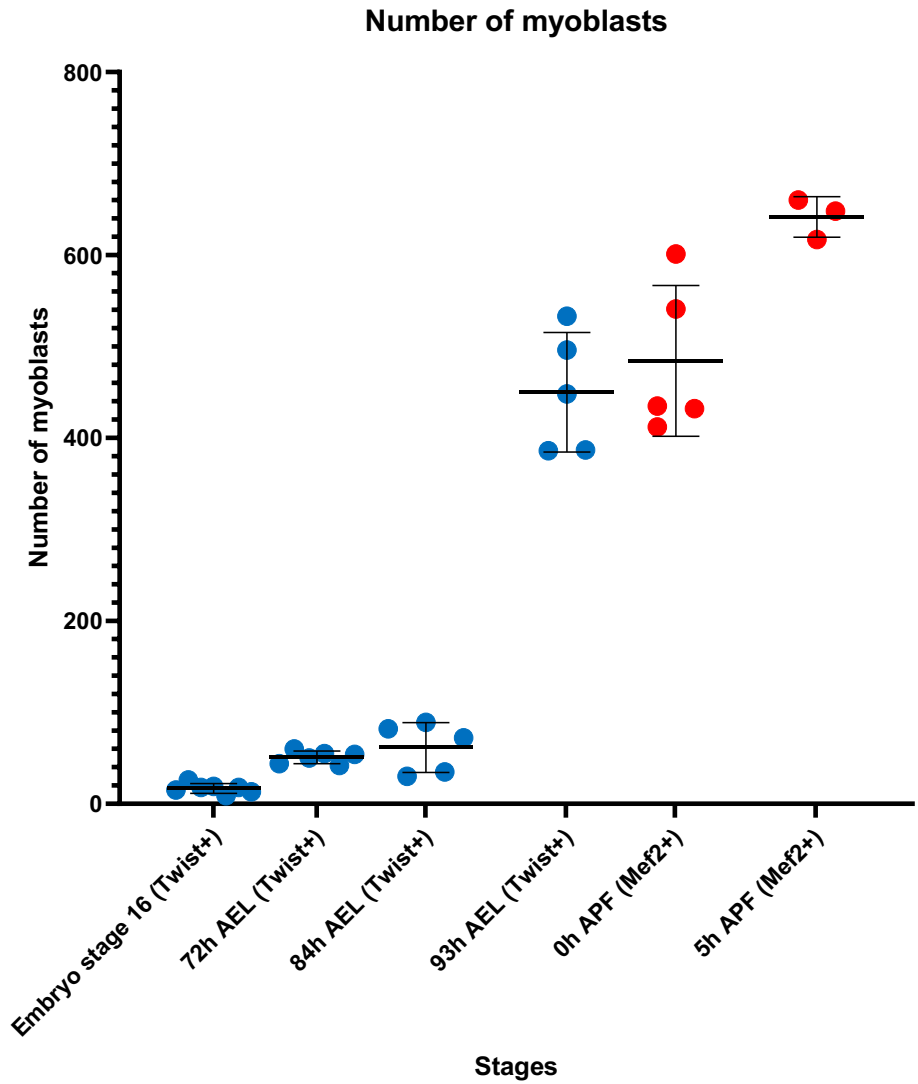


**Figure 1H – Adult leg muscle development**

**Metamorphosis: leg disc evagination.**

H/ 3D representation of confocal picture of a T1 leg disc, at 0h APF (H1) or at 5h APF (H2). In H1a and H2a the epithelium in gray (Fas3 staining), in H1b and H2b the myoblasts in green and blue (gradient depending of the GFP intensity) (Mef2>GFP). In H1c and H2c, 3D projection of the myoblasts without the epithelium. Myoblasts are subdivided into two clusters: one peripheral forming one layer in contact with the epithelium, another one in central position. At 5hAPF, the leg disc begins to evaginate from its center and myoblasts organized in the different leg segments.



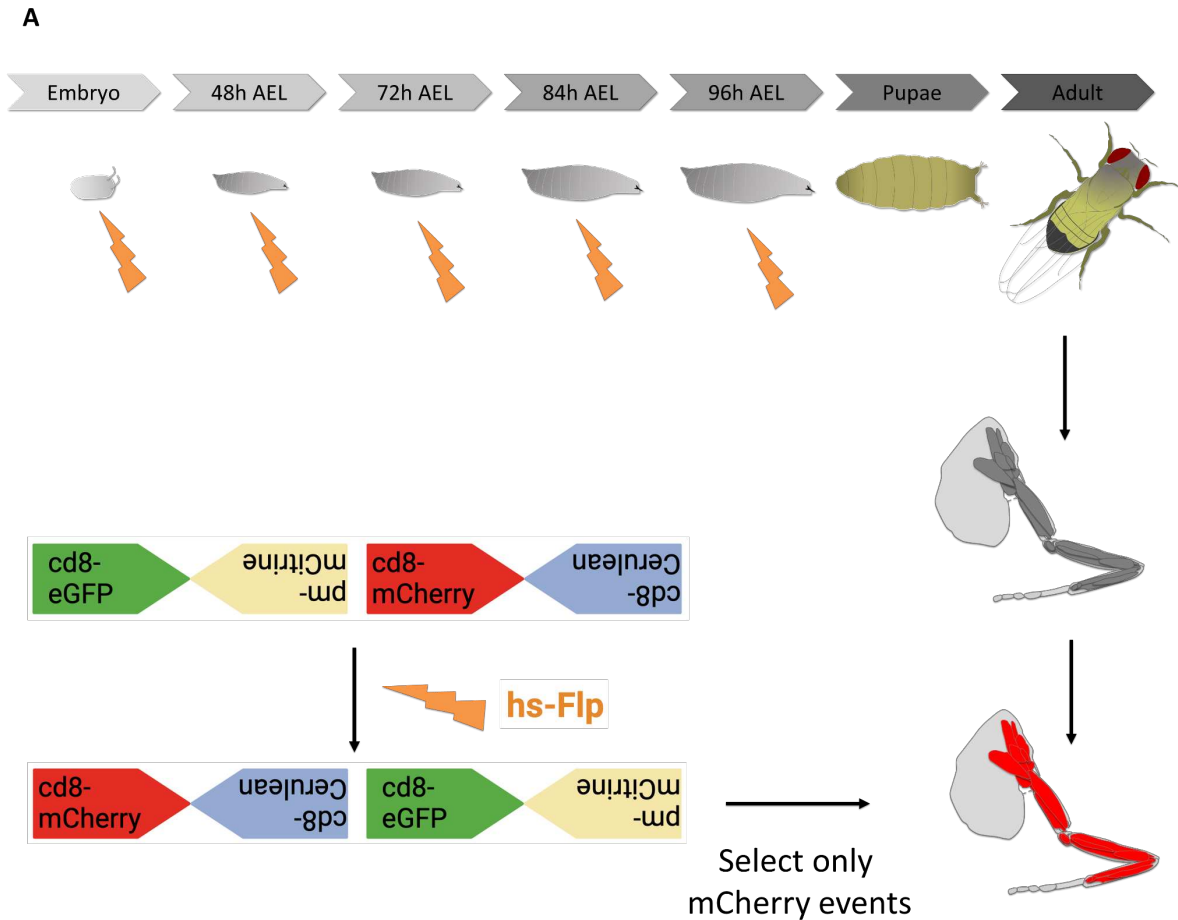


**Figure 11 – Adult leg muscle development**

**Number of myoblasts during the leg disc development.**

At late embryo stages, there are around 17 myoblasts Twist+ (SD 5). At 72 hours after egg laying (AEL), there are around 51 myoblasts Twist+ (SD 7). At 84h AEL, there are around 62 myoblasts Twist+ (SD 27). At 96h AEL, the number of myoblast has drastically increased to 450 myoblasts Twist+ (SD 65). At 0hAPF, there are around 484 myoblasts Mef2+ (SD 83). At 5hAPF, there are around 642 myoblasts Mef2+ (SD 22).





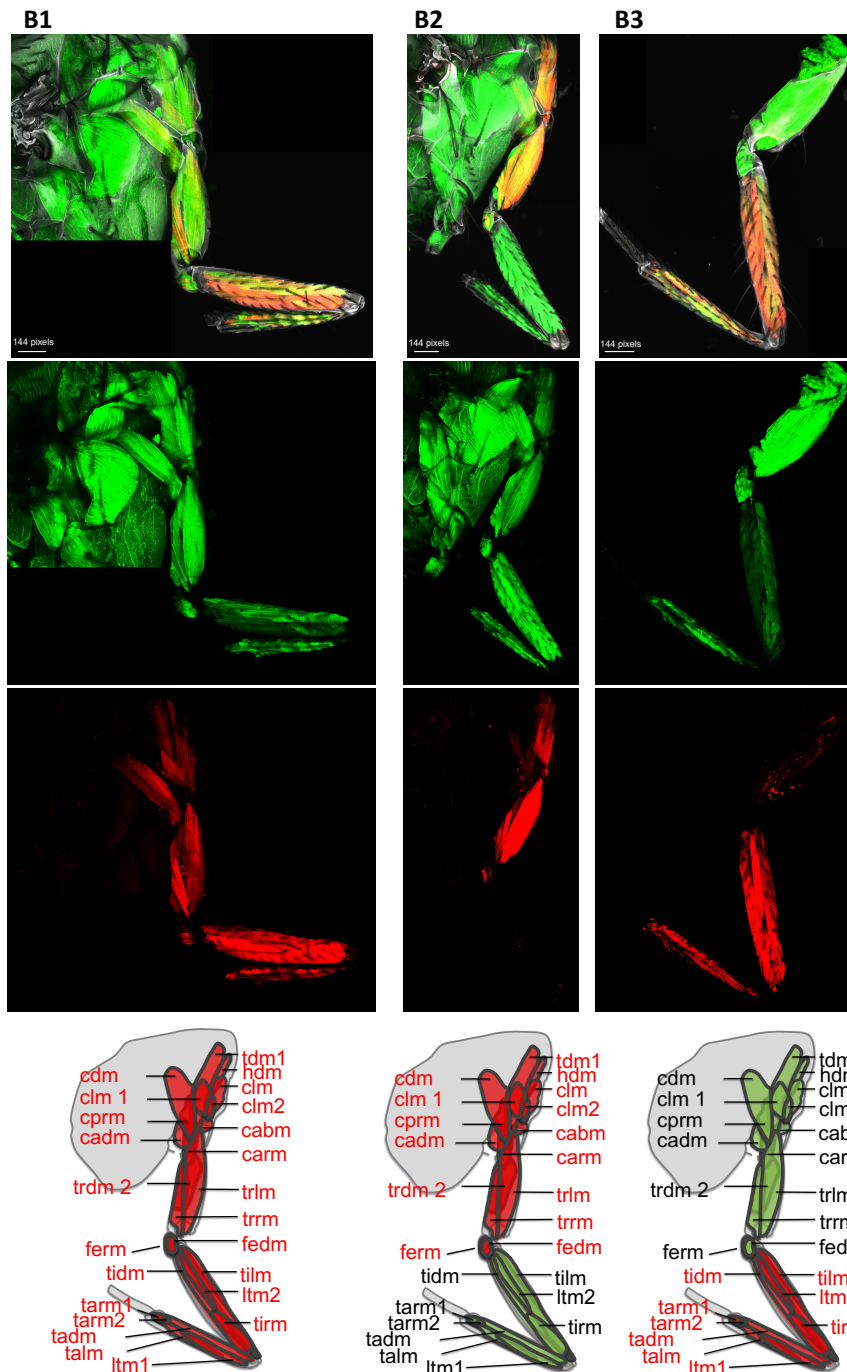
**Figure 2A – Clonal analysis**

**The Flybow system.**

The Flybow system lineage tracing method (UAS-FB1.1) allow to follow the progeny of a single myoblast by permanently labels them with different fluorescent markers (Hadjiconomou et al., 2011). The UAS-Flybow.1.1 allows stochastic expression of the fluorescent protein from a single transgene using excision mediated by the hs-Flp recombinase. It contains UAS regulatory sequences fused upstream of two adjacent invertible cassettes, each flanked by inward-facing FRT sites and each containing a pair of fluorescent proteins in opposing orientations. After activation of the hs-Flp5 at different developmental time point, the cassettes are inverted, the m-Cherry sequence can become juxtaposed with the UAS sequence, resulting in expression of that protein. The system is specifically activate in the myoblasts thanks to the driver Mef2-Gal4 that activated the UAS-Flybow1.1.







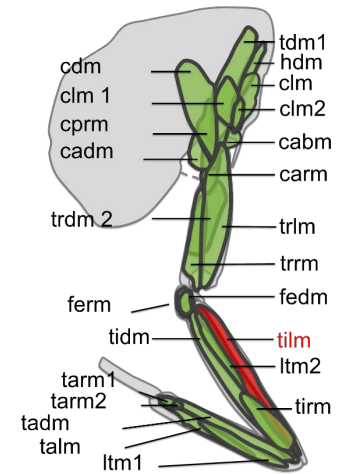
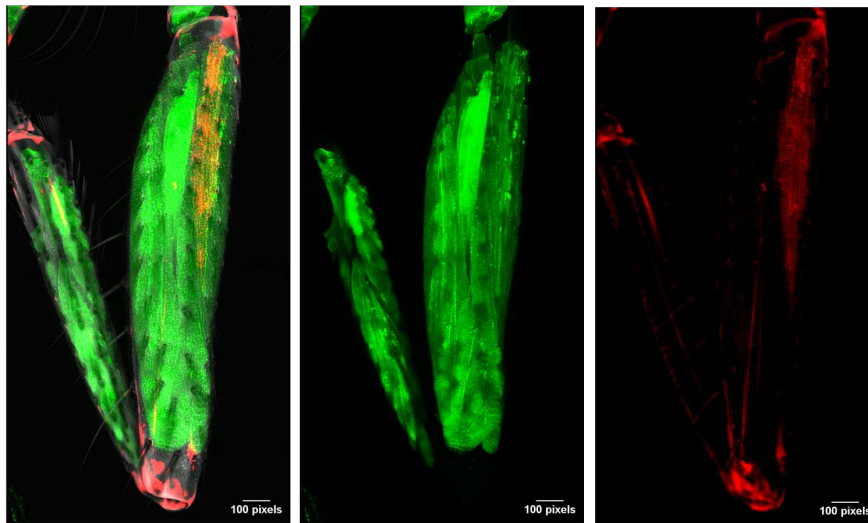
**Figure 2B – Clonal analysis**

**Myoblasts are specify to produce specific muscles**

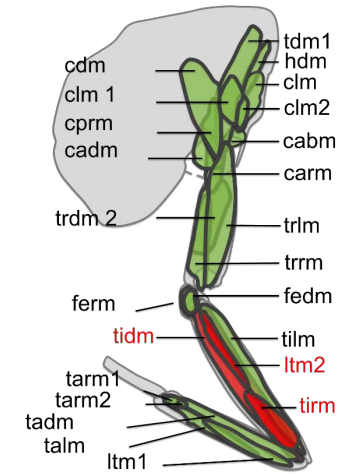
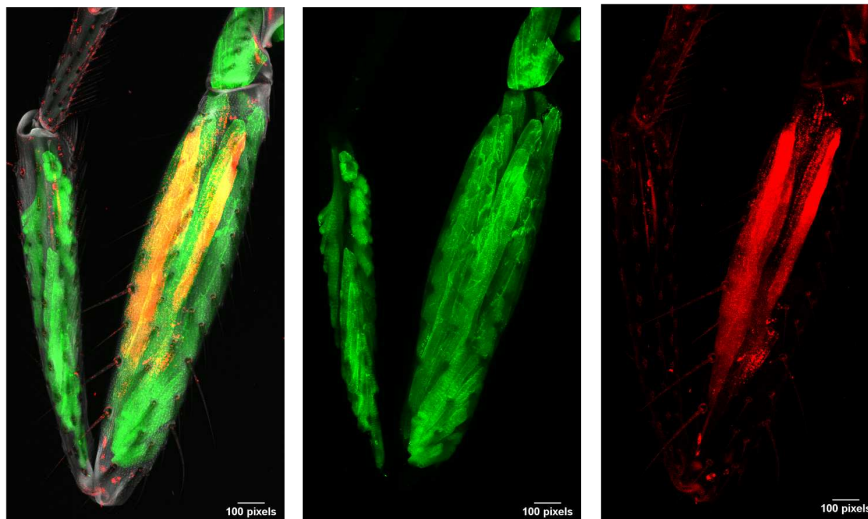
B/ T1 legs from mCherry+ myoblast clones generated by the Flybow technique, at 12-24 hours AEL (late embryo stage). In green all muscles labeled with Mef2>GFP and in red the different clones obtain. Three types of clones are obtain: 70% (N=12) labeled all leg muscles across all leg segments, including the prothorax (B1), 18% (N=3) labeling muscles in the femur and tibia leg segments (B2), and 12% (N=2) labeling muscles in the trochanter, coxa, and prothorax (B3).



C1



C2



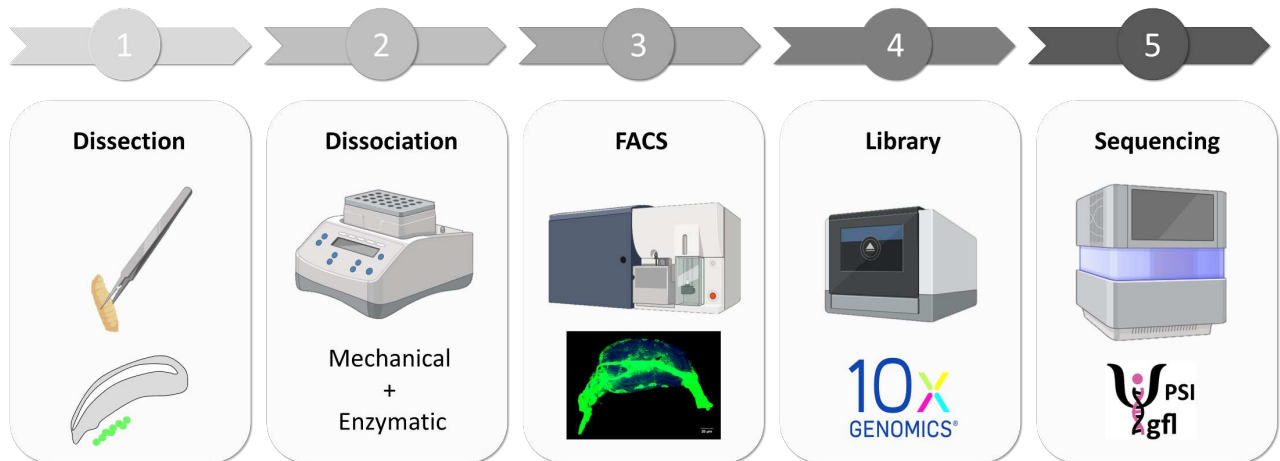
**Figure 2C – Clonal analysis**

**Myoblasts are specify to produce specific muscles**

C/ T1 legs from mCherry+ myoblast clones generated by the Flybow technique, at 84 hours AEL (C1). In green all muscles labeled with Mef2>GFP and in red clone obtain with only the tilm (Tibia levator muscle) labeled (C1) or three of the four femur muscles labeled (C2).



A

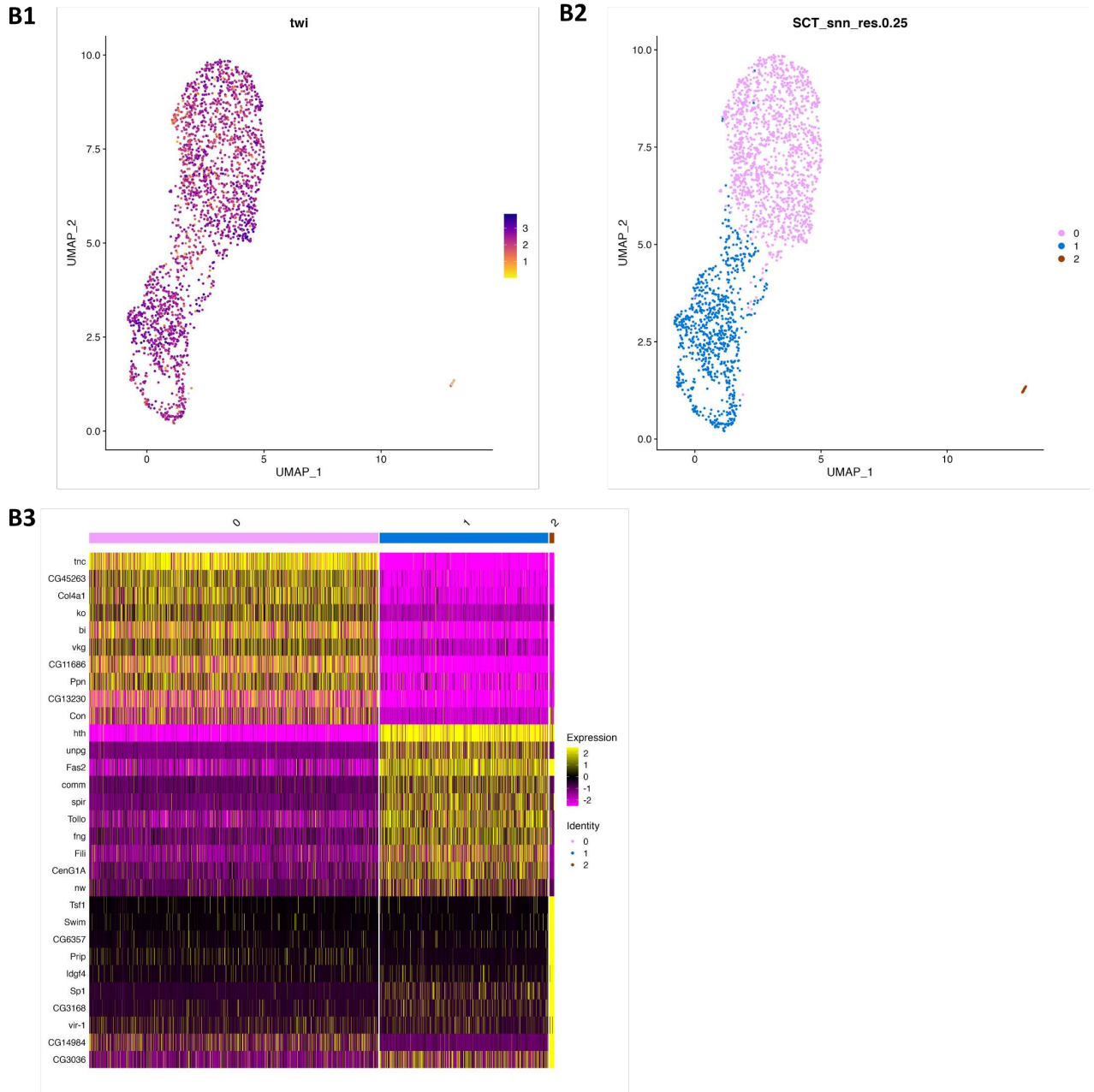


**Figure 3A – Single-cell sequencing protocol**

**The single cell follows five steps.**

A/ First, we dissected the leg discs (Schematic of a cross section of a leg disc at 93h After Egg Laying, AEL) that will be dissociated by enzymatic and mechanical dissociation process. Then, we FACS sorting the GFP+ myoblasts (confocal picture of a 93h After Egg Laying leg disc (AEL) with myoblasts in green (24B-Gal4>UAS-mCD8::GFP) and nuclei in blue(DAPI)) to captured and construct the library by following 10X Genomics protocol. Finally, library is sequencing at the IGFL sequencing platform (PSI). The same procedure was use for all the single-cell sequencing time points.





**Figure 3B – Single-cell sequencing at 93h AEL**

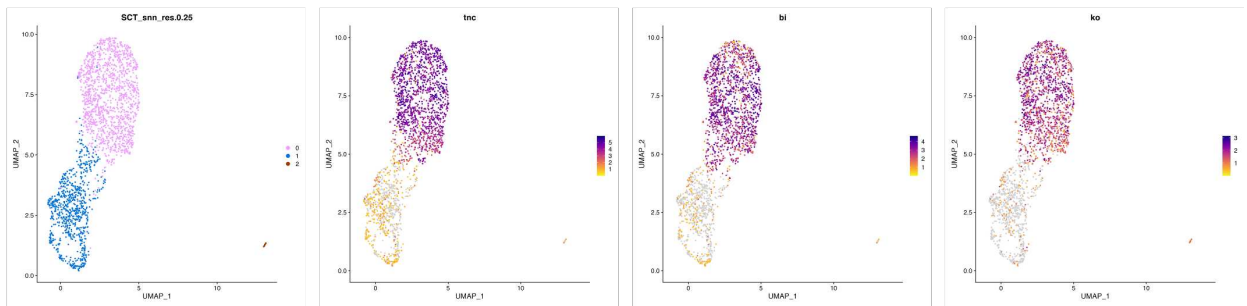
**Analyzing results with Seurat.**

B/ This analyze of the single cell sequencing data of myoblasts (Twist+, B1) at 93h AEL highlight 2 distinct clusters of myoblasts (cluster 0 and 1) at a resolution of 0,25 (B2). B3 show heat map of the top10 genes highly expressed for each cluster. Cluster 0 is marker by the expression of genes such as tnc, ko and bi, while cluster 1 can be identified by the expression of genes such as hth and unpg.

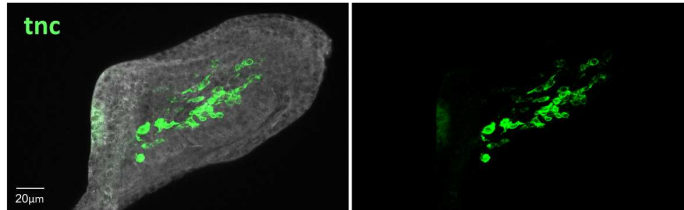




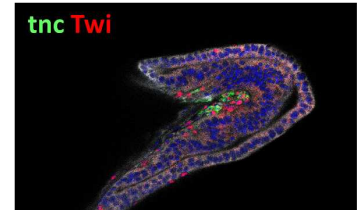
C1



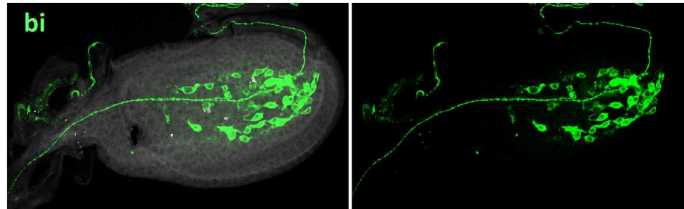
C2a



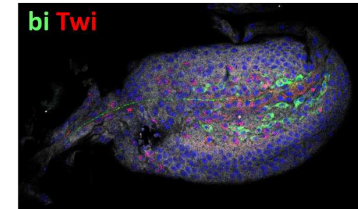
C2b



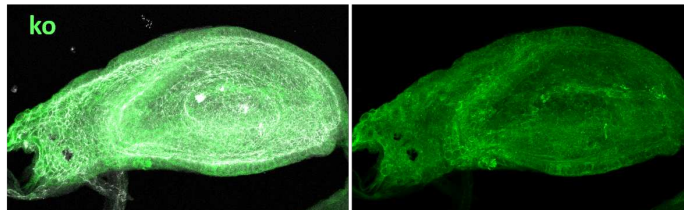
C3a



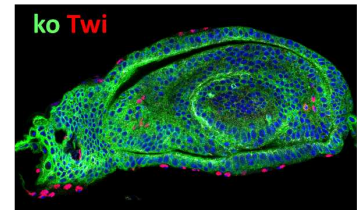
C3b



C4a



C4b



### Figure 3C – Single-cell sequencing at 93h AEL

#### Identification of cluster 0.

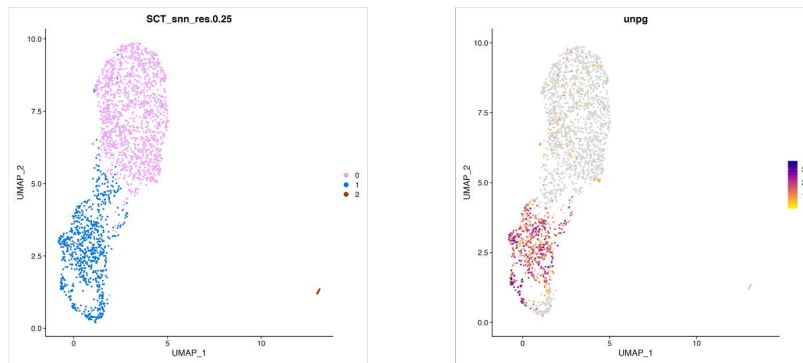
C1/ UMAP representation of the expression of *tnc*, *bi* and *ko* at 93h AEL. The three genes are markers of cluster 0.

C2 to C4/ Confocal pictures of leg discs at 93h AEL (z projection C2a, C3a, C4a, or z-stack C2b, C3b, C4b).

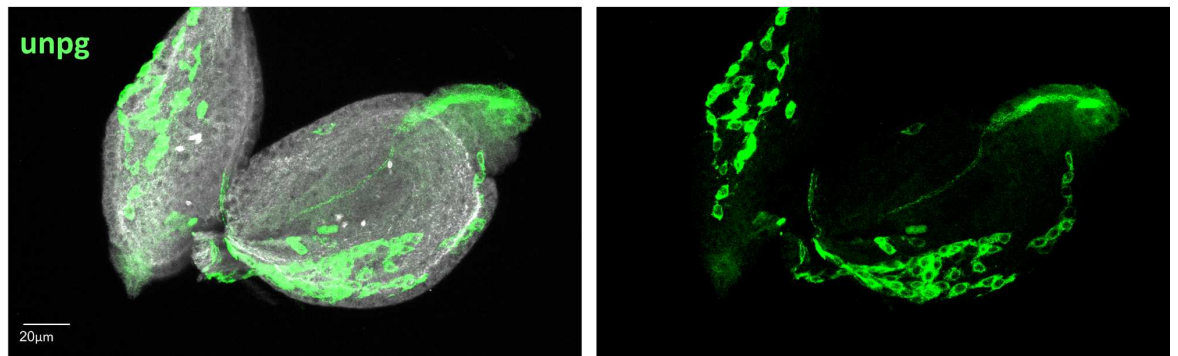
C2/ GFP expression of enhancer trap of *tnc*-Gal4 (green), epithelium in gray (Fas3), myoblasts in red (Twist+) and nuclei in blue (DAPI). C3/ GFP expression of enhancer trap of *bi*-Gal4 (green), epithelium in gray (Fas3), myoblasts in red (Twist+) and nuclei in blue (DAPI). C4/ GFP expression of enhancer trap of *ko*-Gal4 (green), epithelium in gray (Fas3), myoblasts in red (Twist+) and nuclei in blue (DAPI).



D1



D2



D3



### **Figure 3D – Single-cell sequencing at 93h AEL**

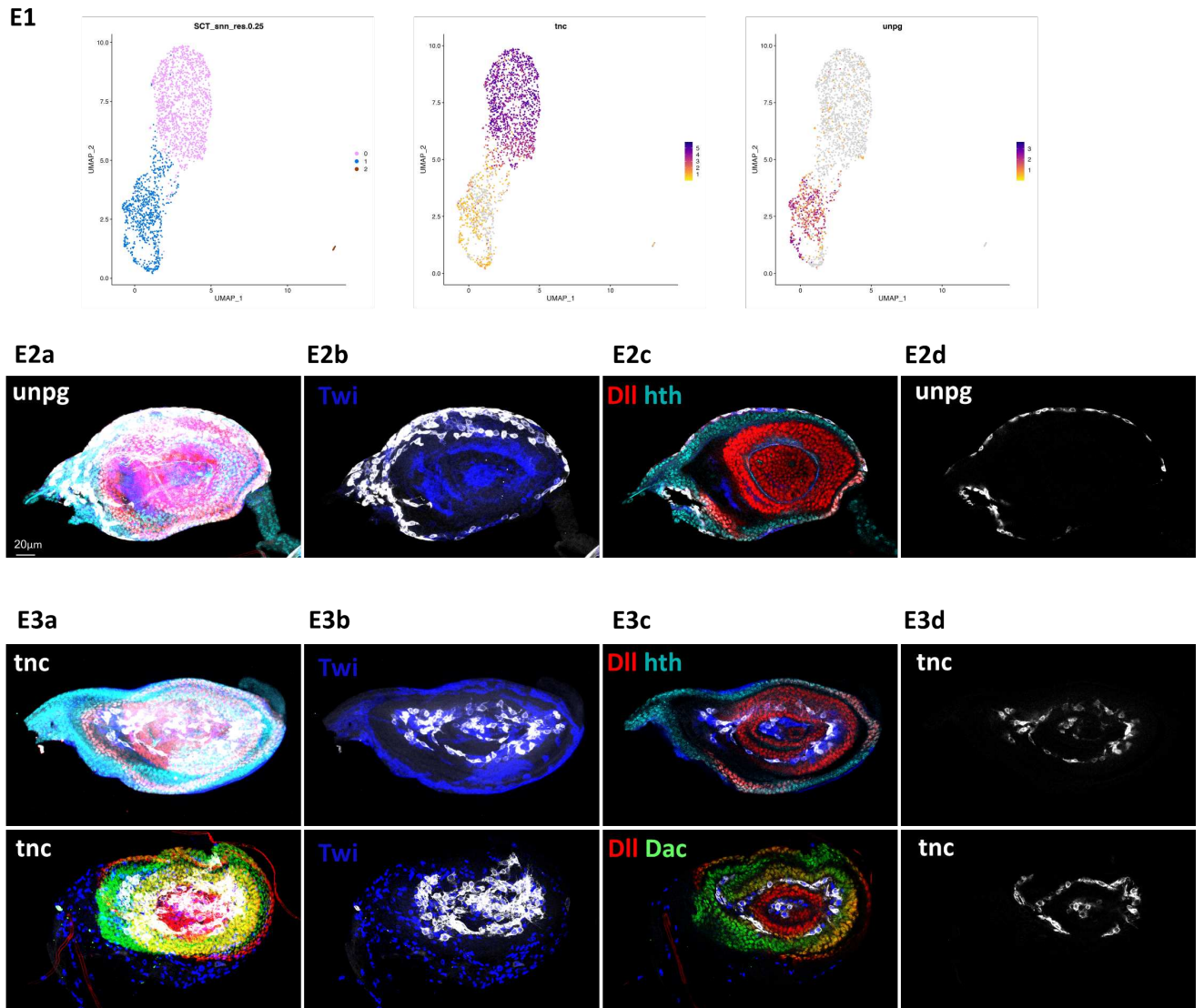
#### **Identification of cluster 1.**

D1/ UMAP representation of the expression of unpg, marker of cluster 1, at 93h AEL.

D2/ z projection of confocal pictures of two leg discs attached at 93h AEL. GFP expression of enhancer trap of unpg-Gal4 (green) with epithelium in gray (Fas3).

D3/ z-stack D3 of confocal pictures of two leg discs attached at 93h AEL. GFP expression of enhancer trap of unpg-Gal4 (green) with epithelium in gray (Fas3), myoblasts in red (Twist+) and nuclei in blue (DAPI).





**Figure 3E – Single-cell sequencing at 93h AEL**

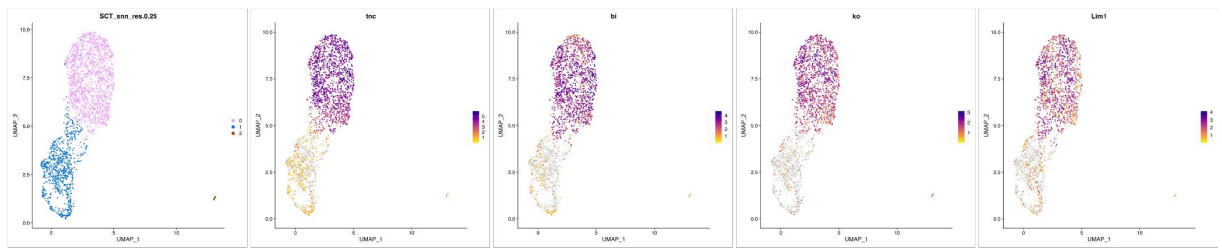
**Characterize the position of the two clusters.**

E/ UMAP representation of the expression at 93h AEL of *tnc* (marker cluster central, cluster 0) and *unpg* (marker cluster periphery, cluster 1).

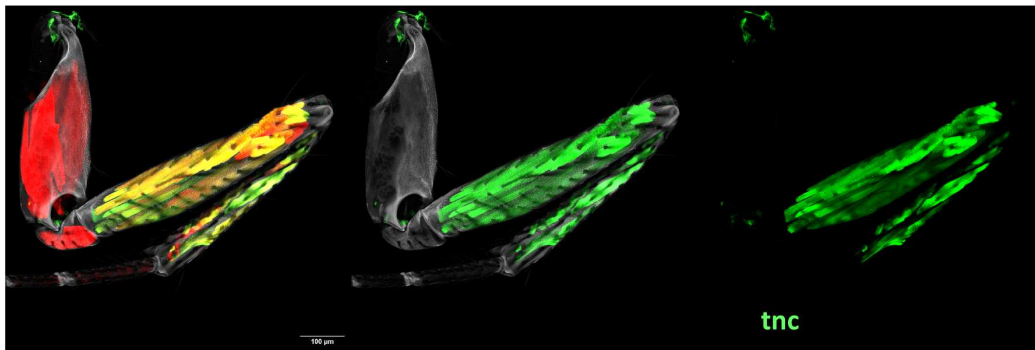
E2 and E3/ The myoblasts (*Twist*<sup>+</sup>, in blue) of central cluster localized in the *Dll*<sup>+</sup> (in red) and *Dac*<sup>+</sup> (in green) regions, while the peripheral cluster localized in the *Hth*<sup>+</sup> (cyan) regions. z projection (E2a and E2b) or z-stack (E2c and E2d) of confocal pictures of a leg disc at 93h AEL with in white GFP expression of enhancer trap of *unpg*-Gal4. z projection (E3a and E3b) or z-stack (E3c and E3d) of confocal pictures of a leg disc at 93h AEL with in white GFP expression of enhancer trap of *tnc*.



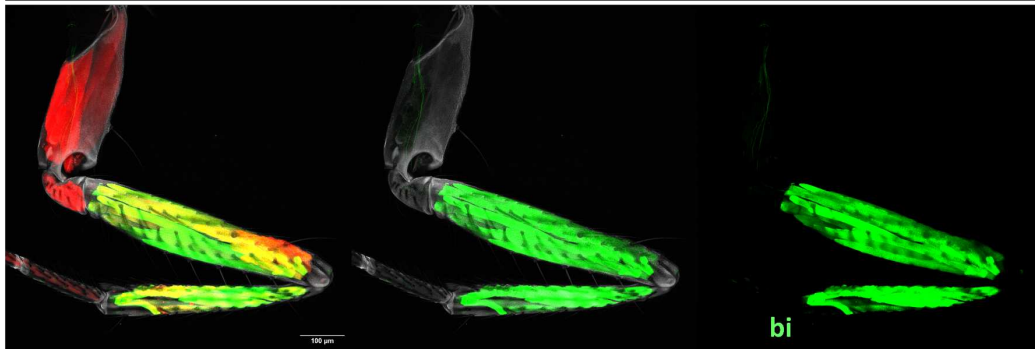
F1



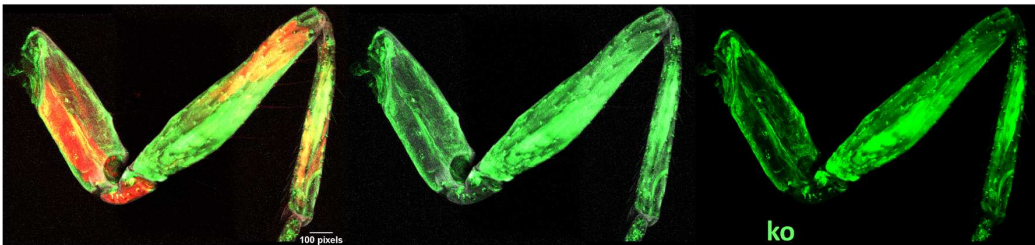
F2



F3



F4



### Figure 3F – Single-cell sequencing at 93h AEL

Lineage muscle tracing system to maintain GFP expression until the adult stage (memory system).

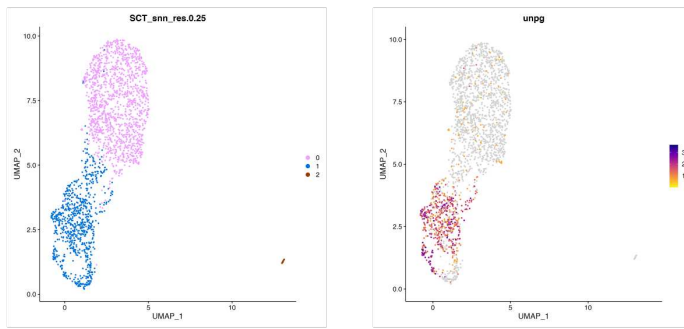
F1/ UMAP representation of the expression of *tnc*, *bi*, *ko* and *Lim1* at 93h AEL. The four genes are marker of cluster 0.

F2 to F4/ Confocal pictures of T1 leg using the memory system with enhancer trap of *tnc*-Gal4 (F2), *bi*-Gal4 (F3), or *ko*-Gal4 (F4). In red all the muscles labeled with Mhc-RFP, and in gray the cuticle.

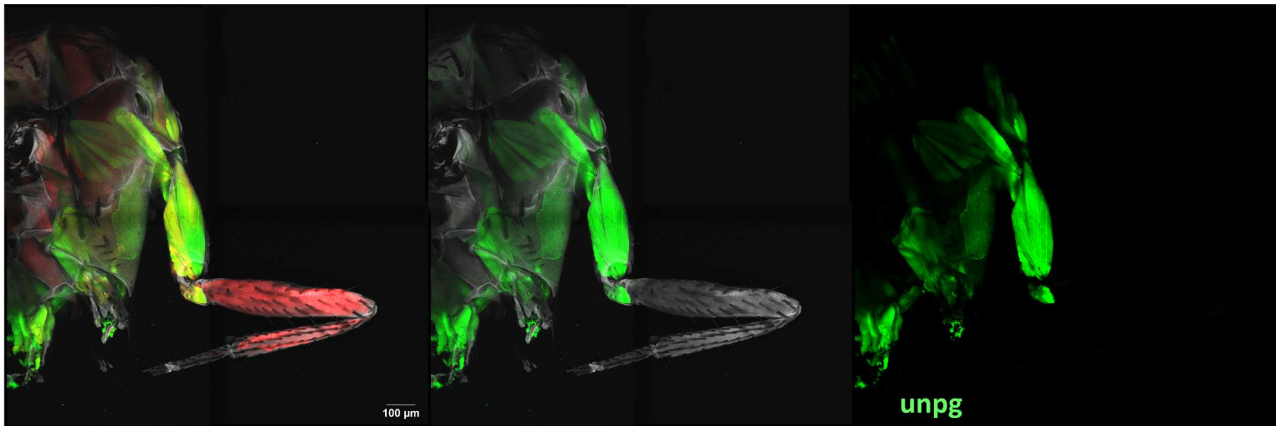




**G1**



**G2**



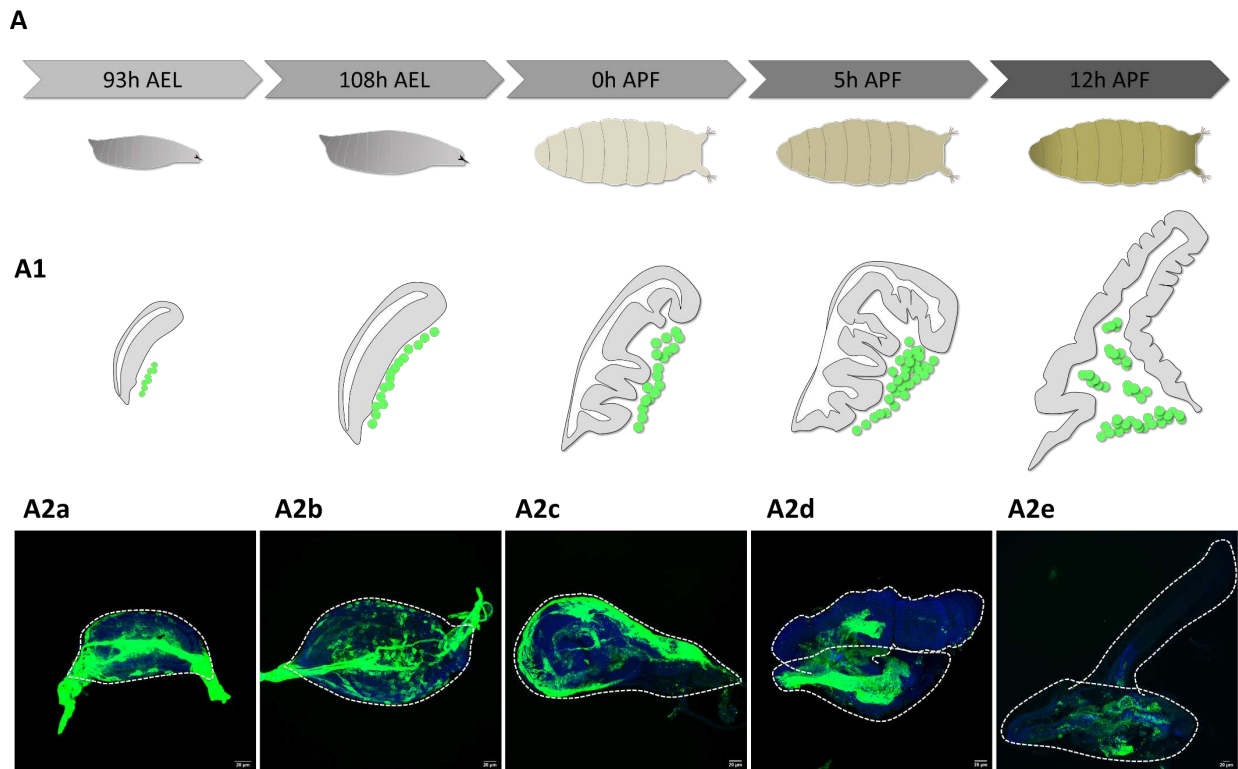
**Figure 3G – Single-cell sequencing at 93h AEL**

Lineage muscle tracing system to maintain GFP expression until the adult stage (memory system).

G1/ UMAP representation of the expression of unpg at 93h AEL, marker of cluster 1.

G2/ Confocal pictures of T1 leg using the memory system with enhancer trap of unpg-Gal4. In red all the muscles labeled with Mhc-RFP, and in gray the cuticle.





**Figure 4A – Single-cell sequencing through the development**

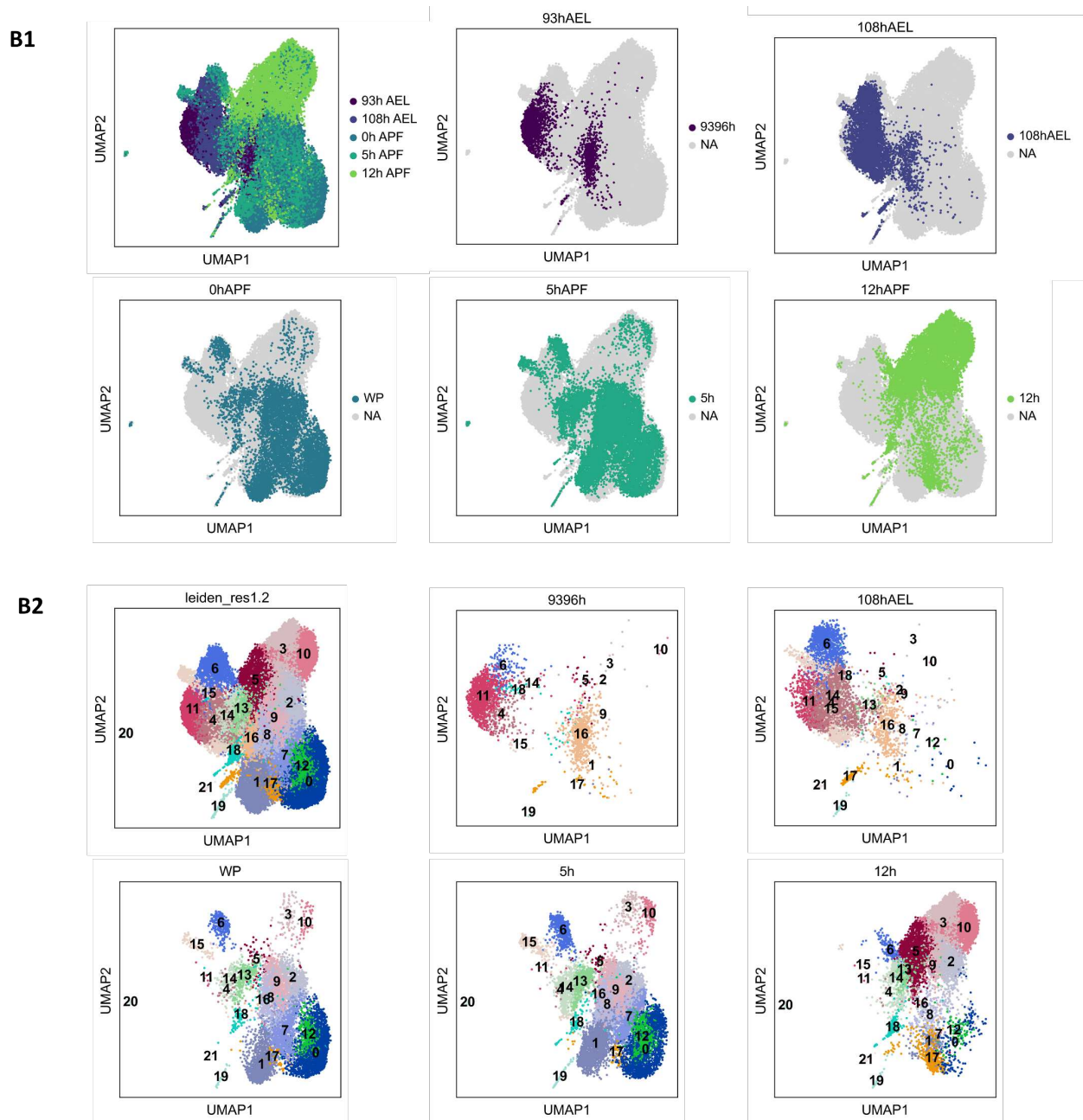
**Representation of the five single experiments.**

A/ Myoblasts were sequenced at five different time points during the development: 93-96h AEL (mid-L3 stage), 108-111h AEL, 0h APF (White Pupae), 5h APF and 12h APF.

A1/ schematic representation of cross section of the leg disc at the different stages with in green representation of myoblasts.

A2/ Confocal pictures of leg discs 93h AEL (A2a), 108h AEL (A2b), 0h APF (A2c), 5h APF (A2d) and 12h APF (A2e). Myoblasts are labeled with GFP under 24B-Gal4>UAS-mCD8::GFP (A2a and A2b) or Mef2-Gal4>UAS-mCD8::GFP (A2c, A2d and A2e), nuclei are in blue (DAPI).





**Figure 4B – Single-cell sequencing through the development**

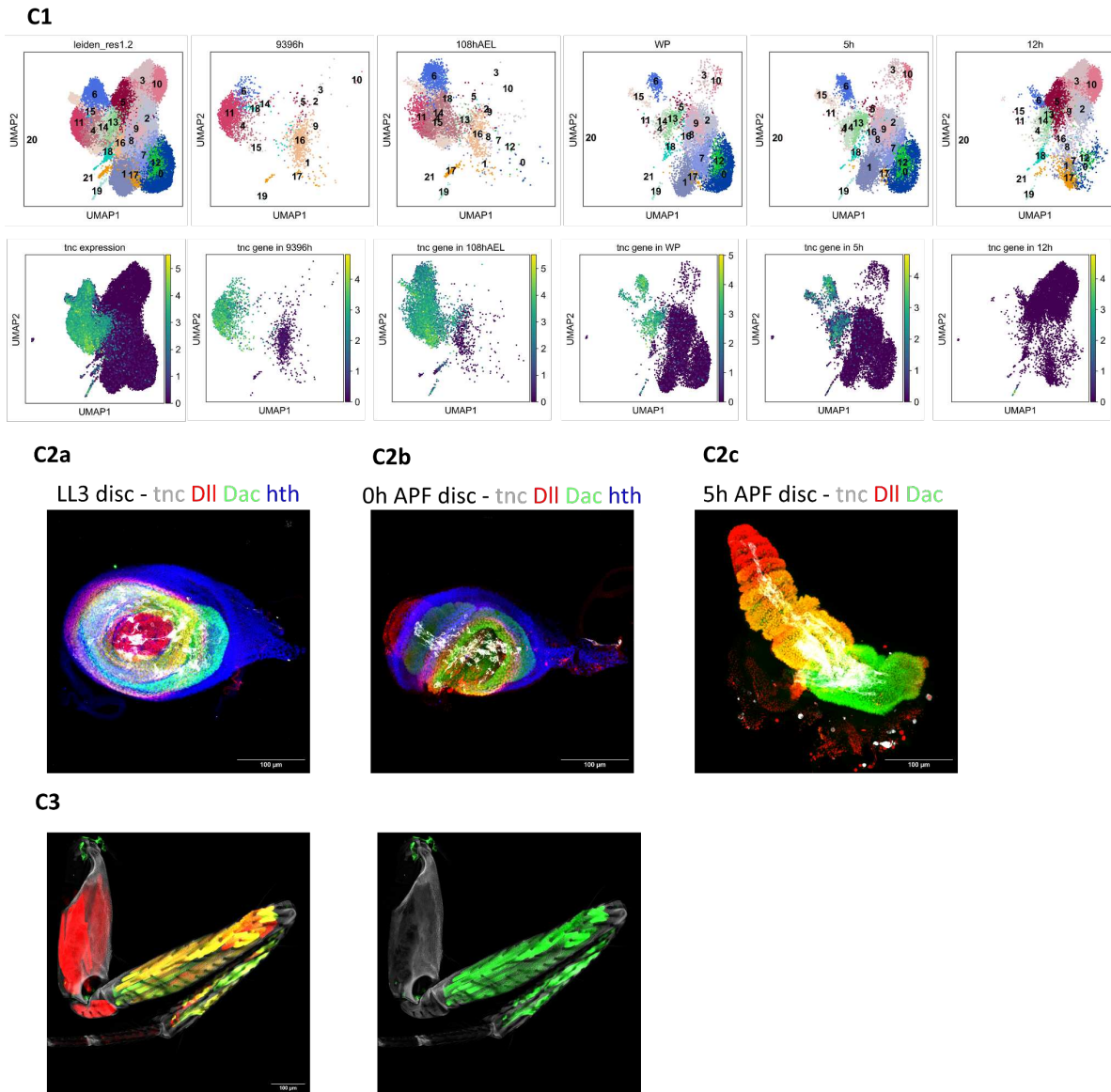
**Analyzing results with Scanpy.**

B/ The five time points have been merge and all the quality control, filtering and normalization have been done on this final dataset after removing the batch effect.

B1/ UMAP representations of each time points.

B2/ UMAP representation of the 22 clusters and UMAP of the cells of each time points belonging to each cluster.





**Figure 4C – Single-cell sequencing through the development**

**Identification of the central clusters with *tnc* marker.**

C1/ UMAP representation of the expression of *tnc* in the full single-cell dataset and for each time point. *Tnc* is highly expressed in the clusters 4, 6, 13, 14 and 15.

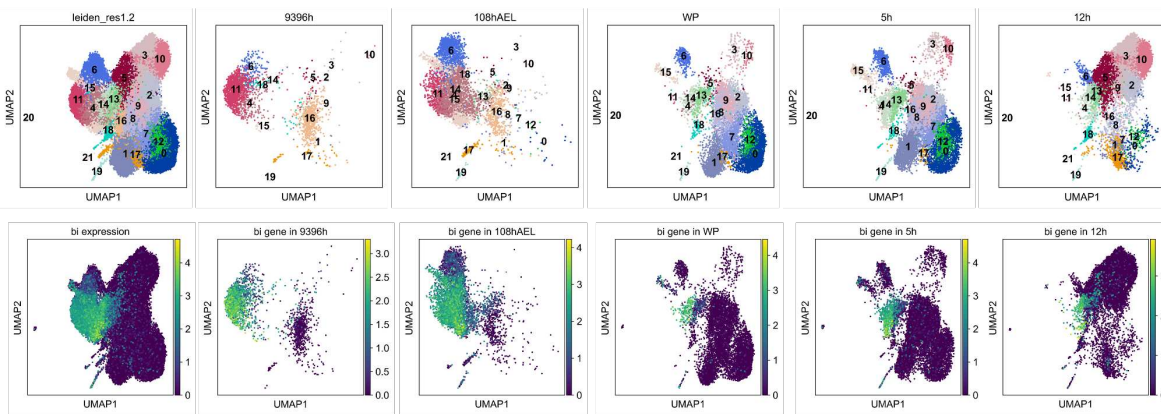
C2/ Confocal pictures of leg disc at LL3 (C2a), 0h APF (C2b), and 5h APF (C2c) stages. Immunostaining with the spatial markers *Dll*, *Dac* and *hth*, show that *tnc*<sup>+</sup> myoblasts are localize in the center of the disc, and later on in femur and tibia segments.

C3/ Confocal picture of T1 leg with the memory system to maintain the expression of GFP until the adult stage, under *tnc*-Gal4 control, resulting in labeling of muscles of the tibia and the femur.



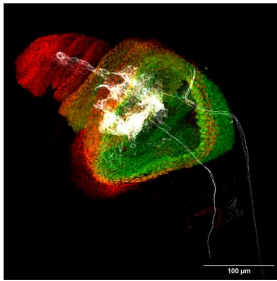


## D1



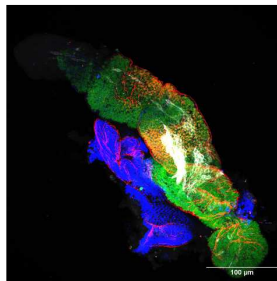
## D2a

0h APF disc - bi Dll Dac

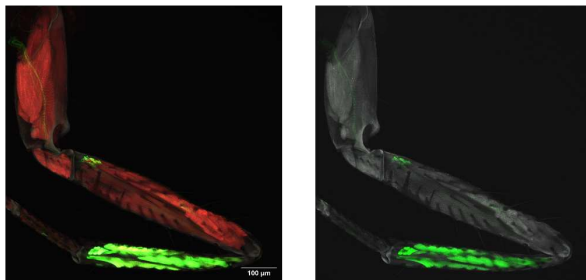


## D2b

5h APF disc - bi Dll Dac hth



## D3



## Figure 4D – Single-cell sequencing through the development

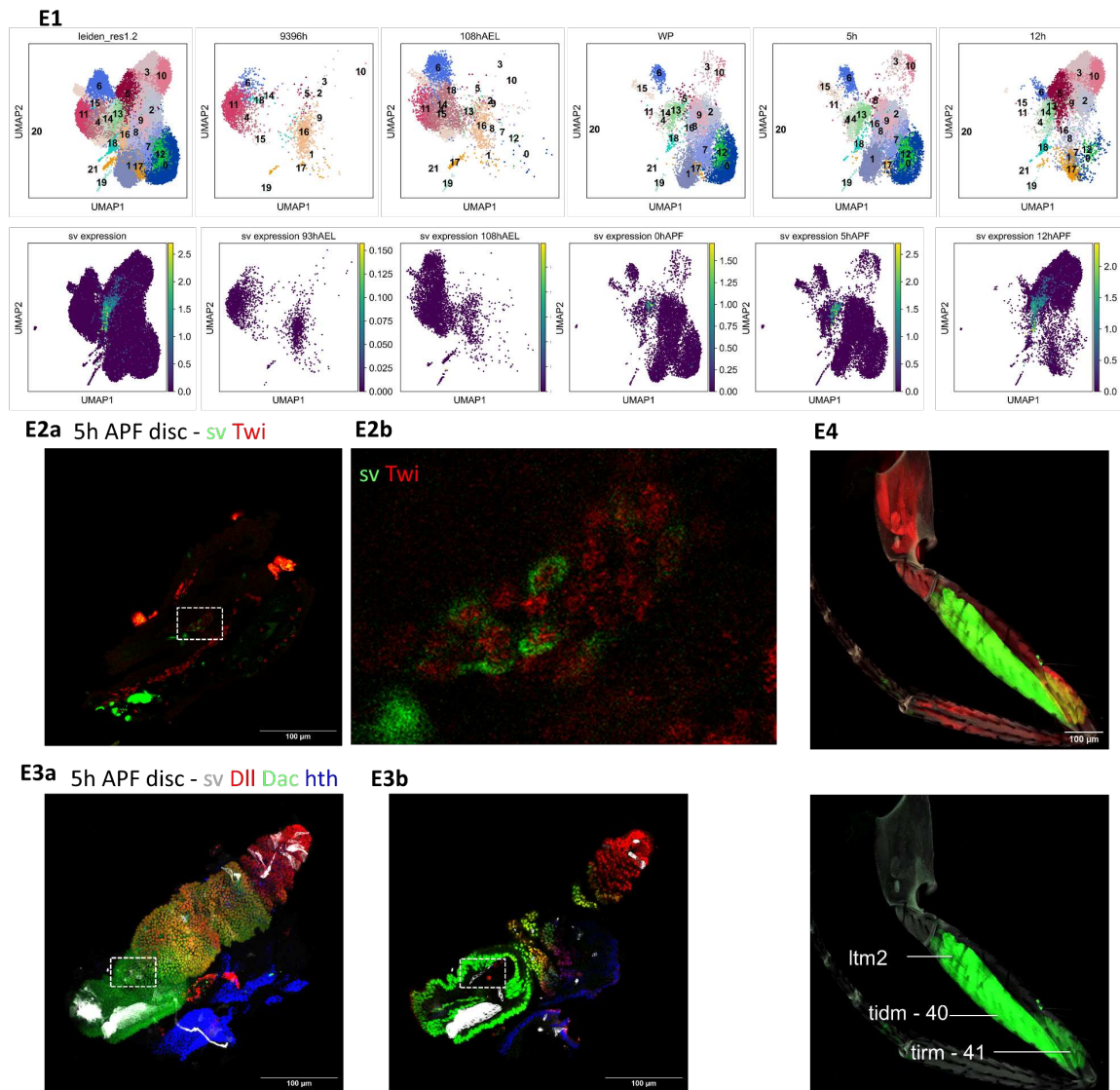
### Identification of the cluster 14 with bi marker.

D1/ UMAP representation of the expression of bi in the full single-cell dataset and for each time point. bi is highly expressed in the cluster 14.

D2/ Confocal pictures of leg disc at 0h APF (D2a), and 5h APF (D2b) stages. Immunostaining with the spatial markers Dll, Dac and hth, show that bi<sup>+</sup> myoblasts are localized in the center of the disc, and later on the tibia segments.

D3/ Confocal picture of T1 leg with lineage muscle tracing system to maintain GFP expression until the adult stage, activating it at early larval pupal using bi-gal4 with tub-gal80ts (0 hour APF), and labeled only tibial muscles.





## **Figure 4E – Single-cell sequencing through the development**

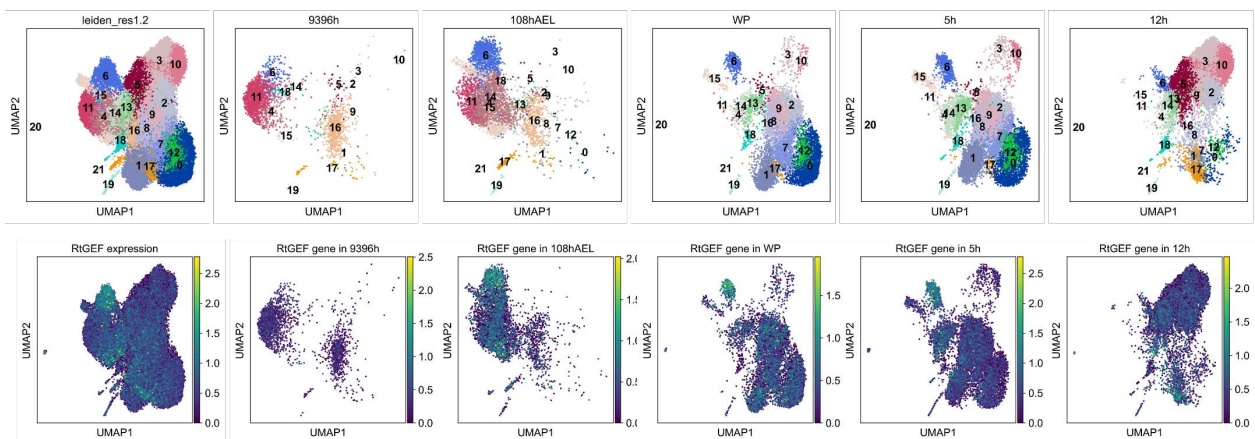
### **Identification of the cluster 13 with *sv* marker.**

E1/ UMAP representation of the expression of *sv* in the full single-cell dataset and for each time point. *sv* is highly expressed in the clusters 13.

E2/ Confocal pictures of leg disc at 5h APF stages (E2a, z-stack) with in green *sv*<sup>+</sup> cells and in red myoblasts (*Mef2*<sup>+</sup>). *sv* is expressed in a sub-population of myoblasts (zoom, E2b). E3/ Confocal pictures of leg disc at 5h APF stages (z-projection E3a and z-stack E3b). Immunostaining with the spatial markers *Dll*, *Dac* and *hth*, show that *sv*<sup>+</sup> myoblasts (dot points) are localized in the femur segment. E4/ Confocal picture of T1 leg with the memory system to maintain the expression of GFP until the adult stage, under *sv*-Gal4 control, resulting in labeling of three muscles of the femur (*ltm2*, *tidm* and *tirm*).

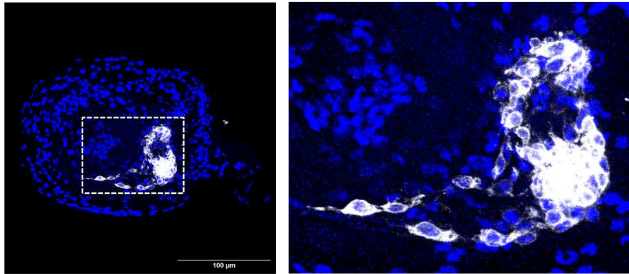


## F1

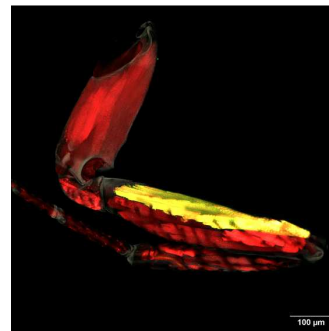


## F2

LL3 APF disc - *tilm* *Twi*

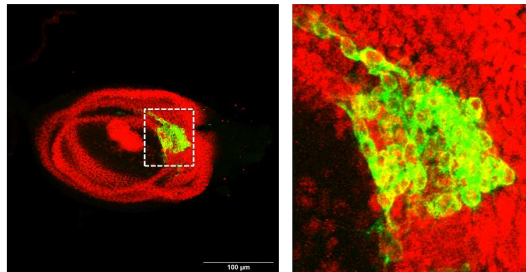


## F4



## F3

LL3 APF disc - *tilm* *Lim1*



## Figure 4F – Single-cell sequencing through the development

**Identification of the cluster 6 with RtGEF marker.**

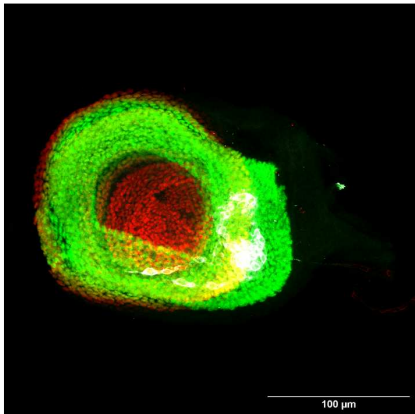
F1/ UMAP representation of the expression of RtGEF in the full single-cell dataset and for each time point. RtGEF is highly expressed in the cluster 6.

F2/ Confocal pictures of leg disc at LL3 stage, in gray the expression of RtGEF-Gal4 and in blue the myoblasts (Twi+). F3/ Confocal pictures of leg disc at LL3 stage, in green the expression of RtGEF-Gal4 and in red the Lim1+ cells. F4/ Confocal picture of T1 leg with the memory system to maintain the expression of GFP until the adult stage, under RtGEF-Gal4 control, resulting in labeling of the *tilm* muscle in the femur segment.



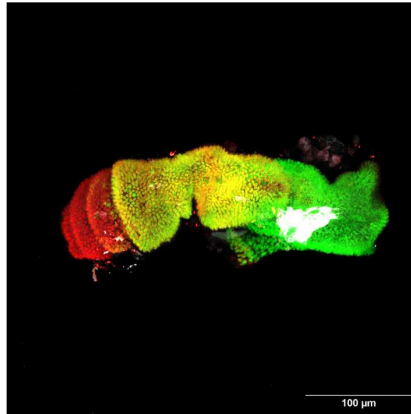
F5a

LL3 disc - tilm Dll Dac



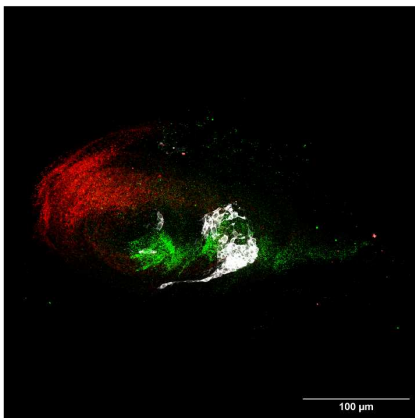
F5b

5h APF disc - tilm Dll Dac



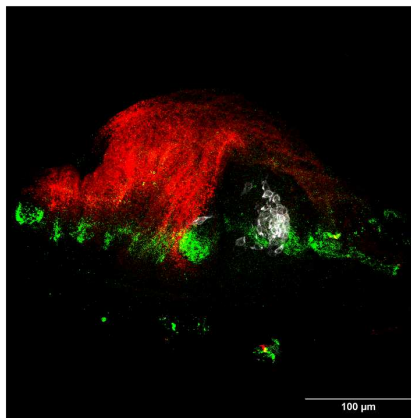
F6a

LL3 disc - tilm wg dpp



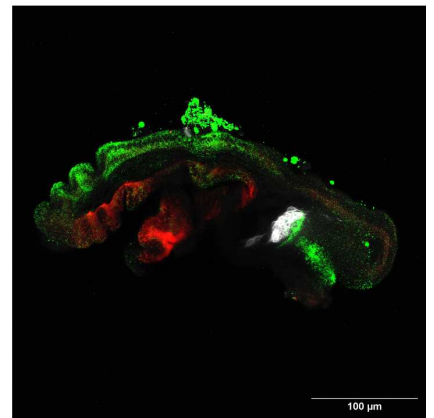
F6b

0h APF disc - tilm wg dpp



F6c

5h APF disc - tilm wg dpp



#### Figure 4F – Single-cell sequencing through the development

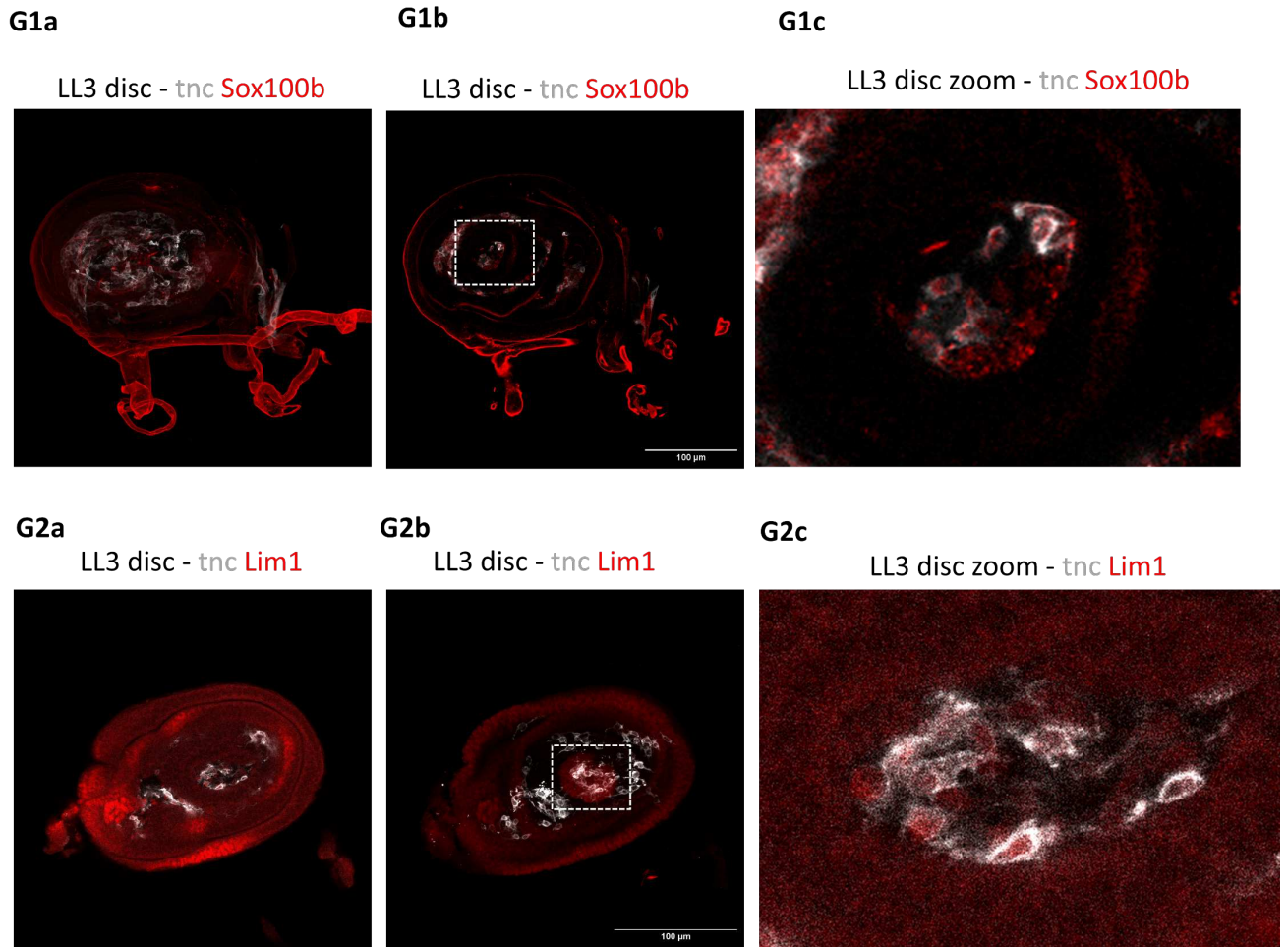
##### **Identification of the cluster 6 with RtGEF marker.**

F5/ Confocal pictures of leg disc at LL3 (F5a) and 5h APF (F5b) stages. Immunostaining with the spatial markers Dll, Dac and hth, show that RtGEF+ myoblasts are localize in the center of the disc, and later on in femur segment.

F6/ Confocal pictures of leg disc at LL3 (F6a), 0h APF (F6b), and 5h APF (F6c) stages. Immunostaining with the spatial markers wg (ventral marker) and dpp (dorsal marker), show that RtGEF+ myoblasts are localize dorsally in the disc.







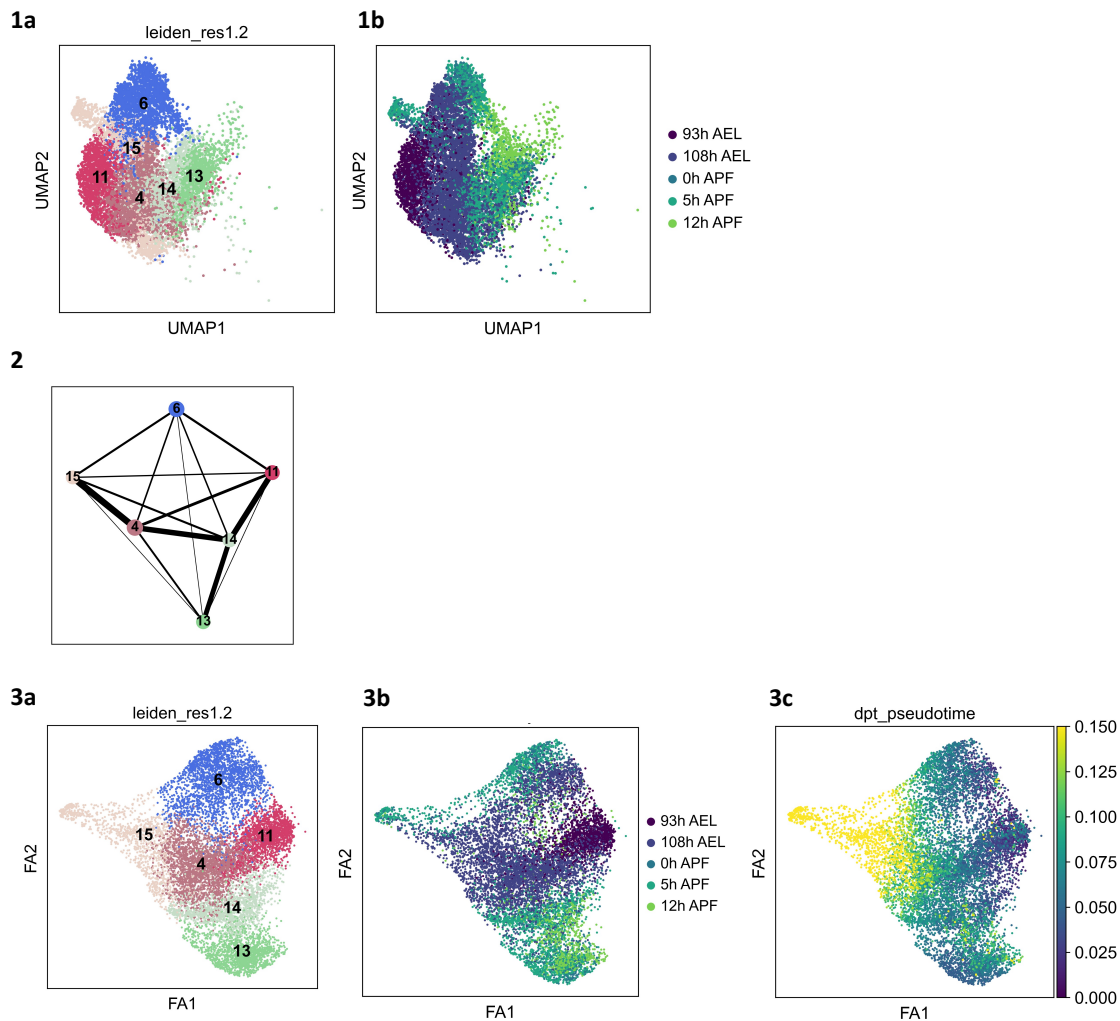
**Figure 4G – Single-cell sequencing through the development**

**Identification of a newly identified cell type: a neural lamella-associated cell type.**

G1/ Confocal pictures of leg disc at LL3 stage with in white tnc-Gal4 expression and in red Sox100b+ cells (G1a z-projection, G1b z-stack, G1b zoom of G2b). Tnc+ and Sox100B+ cells were detect in the presumptive region of the tarsus as well as in the femur and tibia.

G2/ Confocal pictures of leg disc at LL3 stage with in white tnc-Gal4 expression and in red Lim1+ cells (G2a z-projection, G2b z-stack, G2c zoom of G2b). Lim1 is express in the tnc+ cells in the tarsus, tibia and femur.





**Figure 5 – Single-cell sequencing trajectory analysis**

**Identification of a transitory cluster thanks to PAGA analysis.**

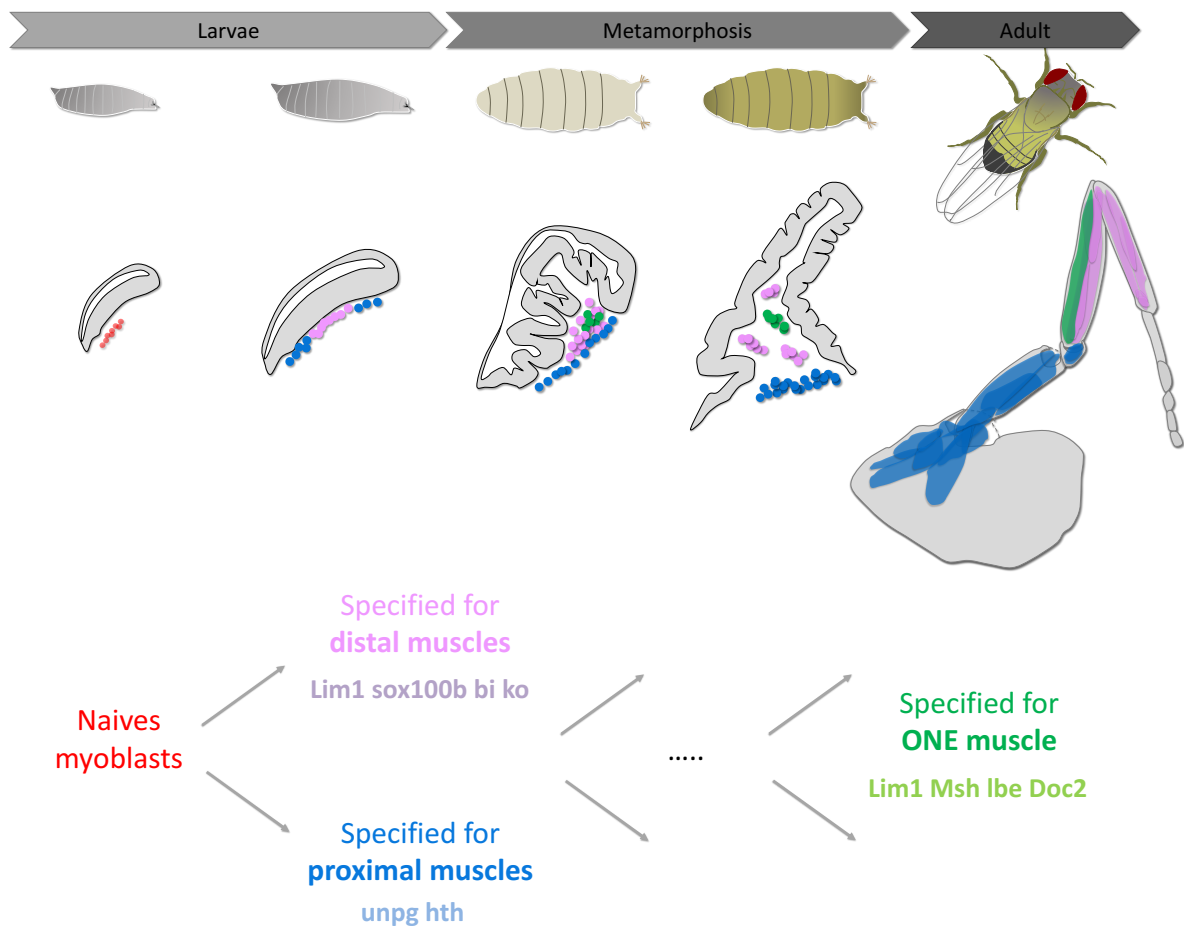
1a/ UMAP after sub-clustering of the “central” cluster forming by clusters 4, 6, 11, 13, 14, and 15.

1b/ UMAP of each time points in the “central” clustering.

2/ Pseudotime analysis (to link the different time points and obtain an evolution in time of the belonging of the clusters) using the Partition-based graph abstraction method (PAGA) design to provide insights into the relationships and connectivity between different cell types or clusters within a heterogeneous population (Wolf et al., 2019).

3/ The SCS trajectory analysis using PAGA strongly suggests that the cluster 4 is a transitory cluster between cluster 6, mostly found at 96 hours AEL, and clusters 13, 14, and 15 (3a and 3b). 3c is showing the pseudotime relationship after PAGA analysis.





**Figure 6 – Model of myoblasts specification during the leg development**

In summary, the two populations of myoblasts giving rise to proximal vs. distal muscles are progressively specify into subpopulations of myoblasts or single muscles expressing specific TFs codes. This data demonstrates that the specification of muscle morphology is establish very early, more than 24 hours before the fusion process. Myoblasts specified to produce individual muscles are position very early during development in the future location of the adult muscle, and myoblasts specified to give rise to the same muscle fuse only between themselves.



## MATERIALS & METHODS

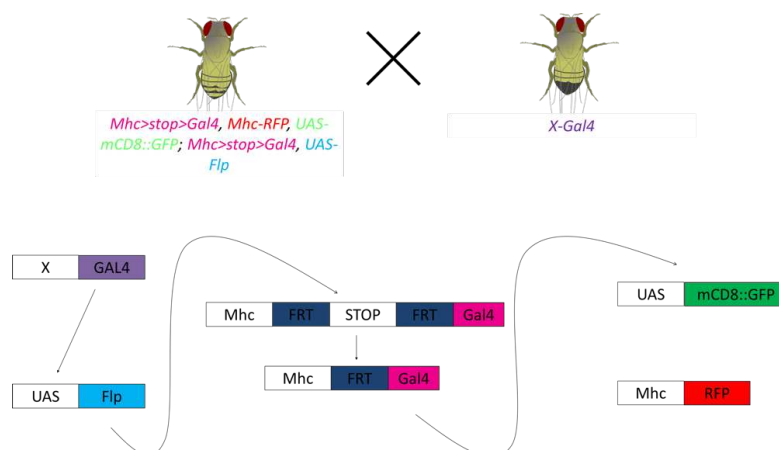
### Experimental model

The development of *Drosophila* is indirect. This genetic model follows a developmental cycle of ten days at 25°C (this cycle is thermos-dependent) composed of distinct stages: embryo, larvae, pupa, and adult.

### Genetics

**The UAS/Gal4 binary system:** The binary system UAS/Gal4 is used in *Drosophila* to ectopically express a gene of interest. The Gal4 protein is a transcriptional factor that binds Upstream Activation Sequences (UAS), activating downstream transgenes such as RNAs or reporters.

**The Memory System:** The Memory System give the possibility to identify specific enhancer-trap Gal4 lineage, even if they are transiently expressed during development. GFP expression will be possible by crossing a Mhc>stop>Gal4, Mhc-RFP, UAS-mCD8::GFP; Mhc>stop>Gal4, UAS-Flp line with a X-Gal4 line. The Gal4 enhancer trap collection activates the Flipase recombinase (Flp), which, in turn, deletes an FRT-flanked stop cassette from a transgene containing the Mhc enhancer (Myosin heavy chain), driving the expression of Gal4. The Gal4 TF exclusively binds a UAS sequence, which will drive expression of an mCD8::GFP reporter.



*Figure: The memory system: lineage muscle tracing system to maintain GFP expression until the adult stage.*





***Drosophila lines used:***

Figure 1A and 1B: line stock: *yw*; *Mhc-Gal4 (attp40)/UAS-mCD8::GFP*; *+ /TM6b*

Figure 1C: line stock: *yw*; *Dll-Gal4, UAS-mCD8::GFP/CyO*; *UAS-mCD8::GFP/MKRS*

Figure 1D and 1F: from cross, progeny: *yw*; *UAS-mCD8::GFP/+*; *24B-Gal4/TM6b*

Figure 1E: line stock: *yw*; *Twi-LacZ//*; *+*

Figure 1G, and 1H: line stock: *yw*; *UAS-mCD8::GFP, Mhc-RFP/CyO*; *Mef2-Gal4/TM6b*

Figure 2B: from cross: *hs-mFlp5/yw*; *UAS-FB1.1/+*; *Mef2-Gal4/+*

Figure 3C2: “*tnc*”: from cross: *yw*; *UAS-mCD8::GFP, Mhc-RFP/+*; *PBac{w[+mC]=IT.GAL4}2090-G4/TM6b*

Figure 3C3: “*bi*”: from cross: *y[1] w[\*] P{w[+m\*]=GAL4}MD735/yw* ; *UAS-mCD8::GFP, Mhc-RFP/+*; *+ /TM6b*

Figure 3C4: “*ko*”: from cross: *yw*; *UAS-mCD8::GFP, Mhc-RFP/+*; *Mi{Trojan-GAL4.0}ko[MI10038-TG4.0]/TM6b*

Figure 3D: “*unpg*”: from cross: *yw*; *Tl{GFP[3xP3.cLa]=CRIMIC.TG4.1}unpg[CR00787-TG4.1]/UAS-mCD8::GFP, Mhc-RFP*; *+ /TM6b*

Figure 3E2: “*unpg*”: from cross: *yw*; *Tl{GFP[3xP3.cLa]=CRIMIC.TG4.1}unpg[CR00787-TG4.1]/UAS-mCD8::GFP, Mhc-RFP*; *+ /TM6b*

Figure 3E3: “*tnc*”: from cross: *yw*; *UAS-mCD8::GFP, Mhc-RFP/+*; *PBac{w[+mC]=IT.GAL4}2090-G4/TM6b*

Figure 3F: cross of the lines *tnc-Gal4*, *bi-Gal4*, or *ko-Gal4* with the line “*memory system*”: *yw*; *Mhc>>Gal4, UAS-mCD8::GFP, Mhc-RFP/CyO*; *Mhc>>Gal4, UAS-Flp//*

Figure 3G: cross of the line *unpg-Gal4* with the line “*memory system*”: *yw*; *Mhc>>Gal4, UAS-mCD8::GFP, Mhc-RFP/CyO*; *Mhc>>Gal4, UAS-Flp//*

Figure 4A2a and 4A2b: from cross, progeny: *yw*; *UAS-mCD8::GFP/+*; *24B-Gal4/TM6b*

Figure 4A2c, 4A2d, and 4A2e: from cross, progeny: *yw*; *UAS-mCD8::GFP/CyO*; *Mef2-Gal4/TM6b*



Figure 4C2: “tnc”: from cross: yw; UAS-mCD8::GFP, Mhc-RFP/+; PBac{w[+mC]=IT.GAL4}2090-G4/TM6b

Figure 4C3: cross of the line tnc-Gal4 with the line “memory system”: yw; Mhc>>Gal4, UAS-mCD8::GFP, Mhc-RFP/CyO; Mhc>>Gal4, UAS-Flp//

Figure 4D2: “bi”: from cross: y[1] w[\*] P{w[+m\*]=GAL4}MD735/yw ; UAS-mCD8::GFP, Mhc-RFP/+; +/TM6b

Figure 4D3: from cross, progeny: y[1] w[\*] P{w[+m\*]=GAL4}MD735/Tub-Gal80ts ; UAS-mCD8::GFP, Mhc-RFP/+; +/TM6b

Figure 4E2 and 4E3: “sv”: from cross: yw ; UAS-mCD8::GFP, Mhc-RFP/+; +/TM6b; TI{GFP[3xP3.cLa]=CRIMIC.TG4.2}sv[CR00370-TG4.2]/+

Figure 4E4: cross of the line sv-Gal4 with the line “memory system”: yw; Mhc>>Gal4, UAS-mCD8::GFP, Mhc-RFP/CyO; Mhc>>Gal4, UAS-Flp//

Figure 4F2, 4F3, 4F5, and 4F6: “tilm”: from cross: yw ; PBac{w[+mC]=IT.GAL4}RtGEF[4046-G4]/UAS-mCD8::GFP, Mhc-RFP; +/TM6b

Figure 4E4: cross of the line tilm-Gal4 with the line “memory system”: yw; Mhc>>Gal4, UAS-mCD8::GFP, Mhc-RFP/CyO; Mhc>>Gal4, UAS-Flp//

Figure 4G: “tnc”: from cross: yw; UAS-mCD8::GFP, Mhc-RFP/+; PBac{w[+mC]=IT.GAL4}2090-G4/TM6b

**Flybow system:** Flybow system (UAS-FB1.1), lineage-tracing method that allows to permanently label cells with different fluorescent markers. The UAS-Flybow.1.1 enables stochastic expression of the fluorescent protein from a single transgene using excision mediated by the hs-Flp5 recombinase. It contains UAS regulatory sequences fused upstream of two adjacent invertible cassettes, each flanked by inward-facing FRT sites and each containing a pair of fluorescent proteins in opposing orientations. In the absence of hs-Flp5 activity, the GFP protein is adjacent to the UAS sequences and is correctly oriented for expression. Upon activation of hs-Flp5, the cassettes are excised/inverted, allowing the m-Cherry sequence to become juxtaposed with the UAS sequence, resulting in the expression of



that protein. The system is specifically activated in the myoblasts thanks to the driver Mef2-Gal4 that activates the UAS-Flybow1.1. Males *yw; UAS-FB1.1/CyO; Mef2-Gal4/TM6b* are crossed with virgin females *hs-mFlp5; +; +* (it is important to use virgin female less than 4 days old). Experiment is conducted on the progeny (male or female) *hs-mFlp5/yw; UAS-FB1.1; Mef2-Gal4*. Different heat-shock times at 37°C are used depending on the developmental time point: 40min for 24h AEL, 30min for 48h, 72h and 84h AEL, and 35min for 96h AEL.

***Origin of the transgenes (the one not in this table are from Dr Jonathan Enriquez lab):***

<b>Transgene</b>	<b>Source</b>
bi-Gal4	Bloomington Drosophila Stock Center number: 80153
ko-Gal4	Bloomington Drosophila Stock Center number: 76207
sv-Gal4	Bloomington Drosophila Stock Center number: 78901
tilm-Gal4	Bloomington Drosophila Stock Center number: 77656
tnc-Gal4	Bloomington Drosophila Stock Center number: 65718
unpg-Gal4	Bloomington Drosophila Stock Center number: 83193
24B-Gal4	Gift from Dr Samir Merabet
Twi-LacZ	Gift from Dr Kat Millen

### **Immunostaining**

***Protocol for larvae and pupae stages:*** During the development, leg discs are dissected and immunostained. To begin, the Drosophila larvae/pupae are placed on ice-cold PBS in a dissecting plate to reduce their movement. After a quick wash with PBS, the dissection is performed to obtain reverse larvae/open pupae without extra tissue (lipids, intestinal system...). Immediately after the dissection, reverse larvae/open pupae are placed in ice cold 4% PFA (freshly prepared, for 2ml of 4%PFA: 500µl 16% PFA, 200µl 10X PBS and 1.3ml H<sub>2</sub>O) to fix them at room temperature for 20min for larvae and 25min for pupae. Before the antibody incubation, five washes (20min each) with PBST-BSA (PBS with TritonX-100 and Bovine Serum Albumin, for 50ml of PBST-BSA: 5ml 10X PBS, 150µl TritonX-100, 0.5g BSA and 8.87ml H<sub>2</sub>O) are performed to permeabilize membranes (TritonX-100) and to avoid non-specific binding (BSA). After this step, the leg disc of the larvae/pupae is dissected and incubated with primary antibodies (dilute with PBST-BSA) at 4°C while shaking overnight.



Then, after five washes (20 minutes each) with PBST-BSA, leg discs are incubated with secondary antibodies (diluted with PBST-BSA) at 4°C shaking overnight, followed by five washes with PBST (20 minutes each). For the final step, Vectashield is used as a mounting medium.

**Protocol for embryos:** Embryos are collected on yeasted (mix yeast with water) apple juice plates the night before the experiment. The first step is to dechorionate the embryos. Embryos are placed in 70% bleach (diluted with warm water) directly on the plate for 4 min of incubation. Simultaneously, a fixative solution is prepared in a 1.5ml Eppendorf with 500ul PFA4% and 500ul Heptane. The plate with the embryos is then inverted over a sieve to drain excess solution. After thorough washing with tap water to remove all bleach, excess water is drained (not completely dried, as embryos stick to sieves). The embryos are transferred from the sieve into freshly prepared fix solution (the fix solution is vortexed before adding). Then, the 1.5ml Eppendorf with the sample is placed into a 50ml Eppendorf with warm water during the fixation period, which lasts between 20 and 25 minutes. The next step is the manual embryo devitellinisation. First, a plate is prepared with a double-sided sticky tape. Then, the fix solution (lower phase) is removed, and the embryos are resuspended in heptane. The embryos are taken in the Heptane solution and placed in the double-sided sticky tape, and left to dry for 2-5minutes under the hood. Then, PBST 0,1% is added to the plate, and the embryos are devitellinized with a tungsten needle (cut along the dorsal midline, starting from the posterior end of the embryo, poke the anterior end of the embryo and lift the embryos). Finally, the embryos are aspirated and placed in a 1.5ml PBST to undergo the immunostaining protocol, similar to the larvae or pupae stage.

**Protocol for Adult:** The adult flies are placed in 70% ethanol (to dehydrate the flies) and then in PBST (to permeabilize membranes). The dissection of these flies consists of taking their T1 legs and fixing them overnight with 4% PFA. After fixation, the legs are washed five times with PBST and placed in a mounting medium before the final mounting.





**Antibodies list:**

Secondary	Fluorochrome	Concentration	Source	Information
Donkey anti Chicken	Alexa 488	2 mg/ml	Jackson	703-145-155
Donkey anti Guinea Pig	DyLight 405	1,5 mg/ml	Jackson	706-475-148
Donkey anti Mouse	Cy3 ~ Alexa + 555	2 mg/ml	Jackson	715-165-151
Donkey anti Rat	Alexa 647	1,5 mg/ml	Jackson	712-605-153
Donkey anti Rat 647	Alexa 647	1,5 mg/ml	Jackson	712-605-153
Goat anti Chicken	Alexa 647	1/800	Abcam	ab150171
Goat anti Mouse	Alexa 405	2 mg/ml	Abcam	ab175661
Goat anti Mouse	Alexa 647	2 mg/ml	Invitrogen	A-21236
Goat anti Rabbit	Alexa 405	2 mg/ml	Abcam	ab175653
Goat anti Rabbit	Alexa Plus 555	2 mg/ml	Invitrogen	A-32732
Goat anti Rabbit	Alexa Plus 647	2 mg/ml	Invitrogen	A-32733
Goat anti Rat	Alexa 555	2 mg/ml	Abcam	ab150166
Phalloïdin	Alexa 546	1/500	Invitrogen	A22283

Primary	Host Species	Concentration	Source
bi = omb	Rabbit	1/1000	Gift from Dr Shinya Matsuda, <shinya.matsuda@unibas.ch>
Dac	Mouse	1/100	DSHB <a href="https://dshb.biology.uiowa.edu/mAbdac1-1">https://dshb.biology.uiowa.edu/mAbdac1-1</a>
Dachshund	Mouse	1/100	DSHB ab579773 (mAbdac1-1-s)
Dll	Guinea pig	1/1000	JE lab
Doc2	Rabbit	1/1000	Gift from Dr Ingolf Reim, <ingolf.reim@biologie.uni-marburg.de>
Dpp	Rabbit	1/100	Gift from Dr Akiyama Takuya, <taa@stowers.org>
en	Mouse	1/2	Gift from Dr Borges Pinto, <pedro.borges_pinto@ens-lyon.fr>



Fasciclin III	Mouse	1/100	DSHB ab528238 (7G10 anti-Fasciclin III)
Hth	Rabbit	1/5000	Gift from Dr Richard S. Mann, <rsm10@columbia.edu>
lbe	Mouse	1/500	Gift from Dr Cedric SOLER, <cedric.soler@udamail.fr>
Lim1	Rabbit	1/50	Gift from Dr Lalanti Venkatasubramanian, <lv347@cam.ac.uk>
Mef2	Rabbit	1/750	Gift from Dr Katrin Domsch, <katrin.domsch@fau.de>
Msh	Rabbit	1/500	Gift from Dr Chris Doe, <cdoe@uoregon.edu>
Sox100b	Rabbit	1/1000 (re-use)	Gift from Dr Xi Rongwen, <xirongwen@nibs.ac.cn>
Tenectin	Rabbit	1/3000	Gift from Dr Stephane Fraichard, <stephane.fraichard@u-bourgogne.fr>
Twist	Rabbit	1/500	Gift from Dr Cedric SOLER, cedric.soler@udamail.fr, and Dr Yad Ghavi-Helm, <ghavi@ens-lyon.fr>
wg	Mouse	1/50	JE lab
$\beta$ -Galactosidase	Mouse	1/500	Promega Z3781
$\beta$ -Galactosidase	Chicken	1/500	Abcam ab134435

### **Microscopy**

A fluorescent stereomicroscope Leica M205 FA is used to select GFP, or RFP positive flies at the different developmental time points.

Confocal pictures have been take with a confocal microscopy Zeiss LSM780 or a confocal microscopy Leica SP8.



## **Software**

All the confocal stacks were analyzed with Image J software, and the 3D pictures were generated using Imaris software.

## **Single-cell sequencing**

**Sample preparation:** To understand how muscles acquire their unique morphology, we conducted single-cell sequencing at different time points during development: 93-96h AEL (mid-L3 stage), 108-111h AEL, 0h APF (White Pupae), 5h APF and 12h APF.

**Dissection:** For each single-cell preparation, we collected a pool of leg discs from the different developmental time points. Myoblasts were labeled with a membrane GFP (UAS-mCD8::GFP) under the control of 24B-Gal4 enhancer trap transgene for earlier time points (93-96h AEL and 108-111h AEL) (24B is a transgene expressed mostly in embryonic mesodermal cells) (Brand and Perrimon, 1993; Luo et al., 1994), or Mef2-Gal4 enhancer trap transgene for the three other time points (0h APF, 5h APF and 12h APF) (mef2 is a transgene containing the regulatory region of the myocyte enhancer factor 2 gene expressed in all myoblasts). Crosses were performed on day 1, and the next day, an egg collection of 3 hours was made, and the flies were raised at 25°C. Subsequently, larvae or pupae GFP positive are dissected at the appropriate developmental time point. Leg discs were dissected in less than 1h to ensure tissue integrity, in M3 medium. Then, they were transferred, using a glass Pasteur pipette, into a 1.5 ml tube with 450µL of M3 medium.

**Dissociation:** (Protocol adapted from: Harzer et al., 2013, “FACS purification of *Drosophila* larval neuroblasts or next generation sequencing”): For the dissociation, 25µL of collagenase I and 25µL of papain are added to the tube (check if the discs have sunken to the bottom of the tube). Then, the tube is placed in a Thermomixer for 1h at 30°C, 300rpm, with the content mixed gently twice during incubation by taking up 200 µL of dissociation solution and expelling it with force. Afterward, the dissociation solution is slowly removed without disrupting the discs at the bottom of the tube, ensuring all the discs have sunken. Next, the solution is removed, and 500µL of PBS1X + 5%SBF is added, and the tissue is disrupted using a 200µL



pipette tip by pipetting the solution up and down between 10 and 20 times until the solution appears homogenous. Finally, the tubes are kept on ice until the FACS step.

	93-96h AEL	108-111h AEL	0h APF	5h APF	12h APF
Number of discs dissected	48 discs	96 discs	82 discs	80 discs	80 discs

**FACS:** The FACS step is performed on a BD FACS ARIA II, (Plateau Technique AniRA-Cytométrie, SFR Biosciences). Dead cells are identified and removed by the addition of DAPI to the sample, and high-quality single cells are sorted into cold PBS1X + 5%SBF. The analysis of the cytometer data is done with the BD FACS Diva 9.0.1 software.

	93-96h AEL	108-111h AEL	0h APF	5h APF	12h APF
Number of cells isolated by FACS	5979 cells	15543 cells	25 872 cells	32 934 cells	19 000 cells

After the isolation of GFP+ cells by FACS, the sample is centrifuged for 5min at 500g at 4°C (2017 rpm= 500rcf), and the volume is calculated depending on the number of cells needed for the sequencing experiment. The appropriate supernatant volume is removed, and the pellet is resuspended in the adequate volume. A control of cell concentration using a Malassez chamber is done before the next step."

**Preparation of the library, 10X Genomics protocol:** Single-cell RNA sequencing was performed using the 10x Genomics single-cell capturing system. Cells suspensions were loaded onto the 10x Genomics Chromium Controller following the manufacturer's instructions. Then, single cell cDNA libraries were prepared using the Chromium Single Cell 3' Reagent Kit version v3.1. The Chromium 10x Genomics Single Cell Capturing system delivers a pool of approximately 3 500 000 10x barcodes, allowing the separate indexing of each cell's transcriptome by partitioning thousands of nuclei into nanoliter-scale Gel Beads-in-emulsion (GEMs), where all generated cDNA share a common 10x Barcode. GEMs are generated by combining barcoded Single Cell 3' v3.1 Gel Beads, Master Mix containing cells and reverse





transcription reagents, and Partitioning Oil onto Chromium Next GEM Chip. The Chip is then placed into the 10x Genomics Chromium controller, and GEMs are generated. Immediately following GEM generation, the Gel Bead is dissolved, primers are released, and mixed with the nucleus lysate. The sample is collected from the Chip and placed into a thermal cycler according to 10x Genomics guidelines. This incubation produces a barcoded cDNA. Then, GEMs are broken, and pool fractions are recovered. Silane magnetic beads are used to remove the RT reagents leftovers and primers. The resulting pure barcoded cDNA is amplified via PCR to increase cDNA to a sufficient mass for library construction (the number of cycles is different depending on the sample). After subsequent rounds of silane magnetic beads-based purification and size selection, enzymatic fragmentation of the full-length cDNA is employed. P5 and P7 sample index and TrueSeq Read 2 are added to the barcoded cDNA via End Repair, A-tailing, Adaptor Ligation, and PCR (number of cycles is different depending on the sample). For all the time points, a simple index sequence is used, except for time point 2 (108h AEL) where a dual index is employed. Prior to the sequencing step, the amount, size, and quality of the cDNA were assessed using the TapeStation 2200 (Agilent DNA ScreenTape High Sensitivity System) and the Qubit system.

**Sequencing:** The resulting libraries were sequenced by the IGFL's sequencing platform (PSI, Lyon, France) using an Illumina NextSeq500 sequencer (28 bp for R1 and 132 bp for R2).

**Data pre-processing:** The BCL read files were converted into FASTQ using the command 'makefastq' from Cell Ranger Software version 5.0.0 (10x Genomics; Zheng et al., 2017). The sequencer generates raw data in the base call (BCL) format, which contains sequencing data for all the libraries in the sequencing run. The Cell Ranger pipeline allows for the demultiplexing of BCL files into FASTQ files for individual libraries. Once the FASTQ files for each sample are generated, data analysis begins, and the Cell Ranger pipeline performs read alignment to the reference genome and UMI counting for a single sample. Reads were mapped to the *Drosophila melanogaster* genome assembly, in which the UAS-mCD8::GFP sequence was added.



**Analysis with Seurat v5:** The Seurat v5 R package is employed for the quality-control, analysis, and exploration of single-cell RNA-seq data (Hao, et al., bioRxiv 2022). Seurat is designed to identify and interpret sources of heterogeneity in single-cell transcriptomics and can integrate diverse types of single-cell data. Upon loading the data, a Seurat object is created, and quality control metrics are explored. Cells are then filtered based on different criteria.

Filters used in the Seurat analysis include:

- Remove genes expressed in less than 3 cells (min.cells): The number of unique genes detected in each cell
- Remove cells with less than 200 genes expressed (min.features): Low-quality cells or empty droplets will often have very few genes
- Keep cells that have unique feature counts over 150,000 or less than 0
- Keep cells that have >5% mitochondrial counts (low-quality / dying cells often exhibit extensive mitochondrial contamination, the set of all genes starting with MT- as a set of mitochondrial genes is used)
- For dataset 93-96h AEL and 108-111h AEL: keep only cells that are GFP+, and Twist+
- For time 0h APF, 5h APF and 12h APF: keep only cells that are GFP+ Twi or/and Mef2+

Then, after removing unwanted cells from the dataset, the subsequent steps involve normalizing the data. Following normalization, the analysis focuses on genes displaying high cell-to-cell variation in the dataset, indicating varying expression levels across cells. A linear transformation, particularly the PCA method, is applied before utilizing dimensional reduction techniques. Next, clustering of the dataset is performed, and non-linear dimensional reduction techniques, such as UMAP, are employed to visualize and explore the datasets. These algorithms aim to learn the underlying manifold of the data, placing similar cells together in low-dimensional space. Cells within the same clusters determined before should co-localize on these dimensional reduction plots. Finally, Seurat is used to identify markers that define clusters through differential expression analysis. By default, Seurat identifies positive and negative markers for a single cluster compared to all other cells. This process can be extended to test differential expression for all clusters, test groups of clusters against each other, or against all cells.

Code availability Seurat: [http://gitbio.ens-lyon.fr/igfl/enriquez/myo\\_d/clustering\\_analysis.git](http://gitbio.ens-lyon.fr/igfl/enriquez/myo_d/clustering_analysis.git)



**Analysis with Scanpy:** In parallel with the Seurat analysis, another strategy was employed by integrating all the samples and conducting downstream analysis using the SCANPY toolkit (Wolf et al., 2018). The initial steps involved filtering cells based on the exclusion of ribosomal genes, expressed mitochondrial genes, and total number of counts. Additionally, cells expressing more than 600 and less than 5000 genes, with a total number of counts lower than 150,000, were retained. The filter also included genes expressed in fewer than three cells. Specifically for time steps 0h APF, 5h APF, and 12h APF, interest was placed in cells expressing at least one among the genes GFP, Mef2, and Twi. For time steps 93h AEL and 108h AEL, interest was in cells expressing at least one among the genes GFP and Twi. Once the data had been preprocessed, and cells and genes had been filtered out, normalization and scaling were performed. Subsequently, batch effects were removed using the BBKNN algorithm (Polanski et al., 2019). Finally, clustering of the dataset was carried out for each time point, and differential gene expression was highlighted, akin to the Seurat analysis.

Code availability Scanpy: [https://github.com/simcada/Thesis\\_single\\_cell.git](https://github.com/simcada/Thesis_single_cell.git)

**Pseudotime analysis:** Pseudotime analysis to link different time points and observe the temporal evolution of clusters was conducted using the Partition-based graph abstraction method (PAGA), with Scanpy, designed to offer insights into the relationships and connectivity between different cell types or clusters within a heterogeneous population (Wolf et al., 2019). This approach allows for a comprehensive understanding of the progression and interrelationships within the identified clusters across various time points.

Sub-clustering of cluster 4, 6, 11, 13, 14 and 15 was performed, and a PAGA analysis was conducted on this subset. The objective was to visualize pseudotime dynamics, particularly focusing on the "central" cluster. This sub-clustering and PAGA analysis allow for a more detailed exploration of the temporal transitions within the specified clusters, shedding light on the pseudotime progression in the central cluster development.

Code availability Scanpy: [https://github.com/simcada/Thesis\\_single\\_cell.git](https://github.com/simcada/Thesis_single_cell.git)



## BIBLIOGRAPHY

1. Christ B, Ordahl CP (1995). Early stages of chick somite development. *Anat Embryol (Berl)*. 10.1007/BF00304424. PMID: 7625610.
2. Amthor H, Christ B, Patel K. (1999). A molecular mechanism enabling continuous embryonic muscle growth - a balance between proliferation and differentiation. *Development* 126(5):1041-53. 10.1242/dev.126.5.1041. PMID: 9927604.
3. Jaglaa, K., Doll, P., Jaglaa, T., Schuhbaw, B., Dretzena, G., and Bellarda, M. (1995). Mouse *Lbx1* and human *LBX1* define a novel mammalian homeobox gene family related to the *Drosophila* lady bird genes'.
4. Mankoo, B.S., Collins, N.S., Ashby, P., Grigorieva, E., Pevny, L.H., Candia, A., Wright, C.V.E., Rigby, P.W.J., and Pachnis, V. (1999). *Mox2* is a component of the genetic hierarchy controlling limb muscle development. *Nature* 400, 69–73. 10.1038/21892.
5. Pownall ME, Gustafsson MK, Emerson CP Jr. (2002). Myogenic regulatory factors and the specification of muscle progenitors in vertebrate embryos. *Annu Rev Cell Dev Biol*. 10.1146/annurev.cellbio.18.012502.105758.
6. Black, B.L., and Olson, E.N. (1998). TRANSCRIPTIONAL CONTROL OF MUSCLE DEVELOPMENT BY MYOCYTE ENHANCER FACTOR-2 (MEF2) PROTEINS. *Annu. Rev. Cell Dev. Biol.* 14, 167–196. 10.1146/annurev.cellbio.14.1.167.
7. Soler, C. (2004). Coordinated development of muscles and tendons of the *Drosophila* leg. *Development* 131, 6041–6051. 10.1242/dev.01527.
8. Carmena, A., Bate, M., and Jiménez, F. (1995). *Lethal of scute*, a proneural gene, participates in the specification of muscle progenitors during *Drosophila* embryogenesis. *Genes Dev.* 9, 2373–2383. 10.1101/gad.9.19.2373.
9. Condic, M.L., Fristrom, D., and Fristrom, J.W. (1991). Apical cell shape changes during *Drosophila* imaginal leg disc elongation: a novel morphogenetic mechanism. *Development* 111, 23–33. 10.1242/dev.111.1.23.





10. Baylies, M.K., and Bate, M. (1996). *twist* : A Myogenic Switch in *Drosophila* . Science 272, 1481–1484. 10.1126/science.272.5267.1481.
11. Figeac, N., Jagla, T., Aradhya, R., Da Ponte, J.P., and Jagla, K. (2010). *Drosophila* adult muscle precursors form a network of interconnected cells and are specified by the *rhomboid* -triggered EGF pathway. Development 137, 1965–1973. 10.1242/dev.049080.
12. Cripps, R.M., Black, B.L., Zhao, B., Lien, C.-L., Schulz, R.A., and Olson, E.N. (1998). The myogenic regulatory gene Mef2 is a direct target for transcriptional activation by Twist during *Drosophila* myogenesis. Genes Dev. 12, 422–434. 10.1101/gad.12.3.422.
13. Dobi, K.C., Schulman, V.K., and Baylies, M.K. (2015). Specification of the somatic musculature in *Drosophila*: Specification of the somatic musculature in *Drosophila*. Wiley Interdiscip. Rev. Dev. Biol. 4, 357–375. 10.1002/wdev.182.
14. Spletter, M.L., and Schnorrer, F. (2014). Transcriptional regulation and alternative splicing cooperate in muscle fiber-type specification in flies and mammals. Exp. Cell Res. 321, 90–98. 10.1016/j.yexcr.2013.10.007.
15. Baylies, M.K., Bate, M., and Gomez, M.R. (1998). Myogenesis: A View from *Drosophila*. Cell 93, 921–927. 10.1016/S0092-8674(00)81198-8.
16. Bate, M. The embryonic development of larval muscles in *Drosophila*. 14.
17. Miller, A. (1950). The internal anatomy and histology of the imago of *Drosophila melanogaster*. Biology of *Drosophila*, 420–531.
18. Lawrence, P.A. (1982). Cell lineage of the thoracic muscles of *drosophila*. Cell 29, 493–503. 10.1016/0092-8674(82)90166-0.
19. González-Crespo, S., and Morata, G. (1996). Genetic evidence for the subdivision of the arthropod limb into coxopodite and telopodite. Development 122, 3921–3928. 10.1242/dev.122.12.3921.
20. Brand, A.H., and Perrimon, N. (1993). Targeted gene expression as a means of altering cell fates and generating dominant phenotypes. Development 118, 401–415. 10.1242/dev.118.2.401.
21. Luo, L., Liao, Y.J., Jan, L.Y., and Jan, Y.N. Distinct morphogenetic functions of similar small GTPases: *Drosophila* Drac1 is involved in axonal outgrowth and myoblast fusion.



## DISCUSSION

### A. Control of muscle morphologies development

This project aims to demonstrate, for the first time, that muscles constitute heterogeneous populations established during development. We bring new observations on the diversity of morphologies of adult leg muscle and its development. Through precise analysis of adult leg muscle architecture, we have identified the complete prothoracic musculature, including both leg and thorax muscles. Our work complements the nomenclature established by Miller (1950) and by Soler et al. (2004), as we have identified two previously undocumented muscles: one in the prothorax (designated as the coxa levator muscle 2) and one in the coxa (named the trochanter reductor muscle 2). This new analysis will facilitate further phenotypical studies on muscle shape, fiber count, and nuclear count following genetic modifications. Thanks to multiple approaches integrating genetics, confocal microscopy, single-cell sequencing, image post-processing, and bioinformatics, we have elucidated the cellular and molecular mechanisms underlying the development of the unique and stereotyped morphological diversity of adult muscles.

#### A.1. Early specification of myoblasts

To understand the establishment of muscle diversity during development, we investigate myoblast proliferation at the cellular and molecular levels within the leg disc. At the embryo stage, myoblasts are already positioned near the future location of the leg disc and will ingress from its dorsal side. Upon entering the disc, they begin proliferation and progressively undergo specification before the fusion process. Myoblasts specified into two sub-populations expressing distinct TFs codes: a peripheral population (*unpg* and *hth*) giving rise to proximal muscles, and a central population generating distal muscles (*lim1*, *sox100b*, *bi*, and *ko*). Concurrently, myoblasts possess their own TF code, operating parallel to the general myogenesis program regulated by *Mef2* and *Twist*. Subsequently, myoblasts further



specify to the level of individual muscles, as exemplified by tilm development, defined by specific gene expression such as *lim1*, *Msh*, *lbe*, and *Doc2*. Furthermore, our single-cell data suggest the existence of a sub-population at the disc center, characterized by *Mef2+* expression, yet not contributing to adult muscle formation. This population exhibits high expression levels of the TFs *Sox100b* and *Lim1*. In a 2022 study, authors sequencing the first tarsal segment identified a novel cell type potentially involved in neural lamella construction, expressing *Sox100B+* and *Lim1+* (Hopkins et al., 2022). Our single-cell RNA-seq results confirm the presence of this population early in leg development, situated at the disc center where the tarsal segment will evaginate.

Moreover, we validate that myoblast populations are specified and spatially organized to prefigure the development and positioning of future specific muscles. . In the future, our laboratory intends to employ MERFISH (Multiplexed Error-Robust Fluorescence In Situ Hybridization) experiments, and computational approaches, to recover all the spatial organization lost during our single-cell experiments (Chen et al., 2015). These data will be integrated with the single myoblast transcriptional trajectory analysis employed in this study to elucidate the spatio-temporal dynamics of gene networks expressed in each myoblast from development through adulthood. Presently, our laboratory is conducting single-nuclei sequencing of the different segments of the adult leg. These investigations will enable us to construct a comprehensive transcriptomic atlas of muscle development, spanning from early stages to adulthood.

Simultaneously, with the TF codes identified in this study, we aim to elucidate the role of candidate genes in establishing unique muscle morphologies. The selection of these candidate genes will involve two filters: genes that are differentially expressed across our single-myoblast RNA-seq clusters, and genes encoding TFs or membrane/secreted proteins that exhibit the highest expression levels within each single-cell cluster. We will abolish the function of these genes using genetic tools such as RNAi to assess their impact on muscle morphologies, including volume, shape, orientation, and number of nuclei. This investigation will underscore the significance of these specific TF codes, in conjunction with the *Mef2*/*Twist* myogenic program, in specifying muscle morphology.



## **A.2. Control of the spatio-temporal organization of myoblasts**

As demonstrated previously, the sub-populations of myoblasts prefigure the position of the future adult muscles. What mechanisms control, firstly, the positioning of these myoblasts, and secondly, the spatial stability of the myoblasts to their final position?

Several studies indicate that arthropod limb development follows a proximal-distal fate, where the proximal part generates the body wall, and the coxa and trochanter together form the coxopodia, while the distal part forms the femur, tibia, and tarsus constituting the telopodia (González-Crespo and Morata, 1996). Our findings offer genetic evidence that muscles also adhere to this proximo-distal developmental pattern, facilitating spatial organization into coxopodia vs telopodia as observed. We raise the hypothesis that muscle development may be regulated by the genetic pattern of the epithelium, which sends spatial signals to myoblasts expressing the necessary receptors, thereby specifying them into specific muscles. Studies indicate that Wg and Dpp are key molecules in establishing the proximal-distal axis of *Drosophila* legs during early leg development. We aim to investigate the roles of Dpp and Wg in defining myoblast proximo-distal specification. We will examine the expression of myoblast marker genes, such as *tnc* for the center, *unpg* for the periphery, and *Tw* for all myoblasts, under Dpp or Wg mutant conditions. By early removal of Dpp or Wg during development, we will assess whether they influence the spatial organization of myoblasts. These experiments will elucidate the impact of modified Dpp and Wg signaling on myoblast proximo-distal specification. In parallel, we will manipulate the direct receptors of Dpp/Wg expressed in myoblasts. Through RNAi targeting *Frazzled/Mad* in myoblasts, we will investigate its effects on myoblast development. It is plausible that without these signals, myoblasts may fail to develop properly, positioning themselves correctly within the leg disc?

Simultaneously, we aim to investigate the role of the epithelium in ventro-dorsal specification of myoblasts. It is established that, after 84h AEL, Dpp facilitates dorsal patterning of the disc, while Wg is involved in ventral domain specification through mutual repression of both genes. Mutations in Dpp result in an expansion of the Wg domain, leading to ventralization of the original dorsal region. We will employ temperature-sensitive mutants of Wg and Dpp to modulate their expression at different time points following their proximo-distal segmental roles (post-84h AEL). We will assess the development of both a dorsal muscle,





the tilm, and a ventral muscle, the tidm, within the femur under these conditions. In the case of the tilm, which normally localizes in the dorsal position, we anticipate impaired development and potential duplication of the ventral muscle, the tidm, in Dpp mutants. Additionally, we will conduct RNAi experiments targeting the receptor of the Dpp/Wg pathway after 84h AEL to directly examine the impact of ventro-dorsal patterning absence.

Moreover, once positioned correctly, these sub-populations of myoblasts must remain clustered and within the appropriate region to facilitate the development of the leg disc. The extracellular matrix (ECM) constitutes a complex network surrounding cells, offering structural support, influencing cell behavior, and regulating various cellular processes. It is well-established that the ECM plays a crucial role in neuromuscular development and the formation of the myotendinous junction (Gullberg and Ekblom, 1995; Borchiellini et al., 1996). Our single-cell results reveal the expression of numerous ECM receptors, such as nidogen and members of the integrin family, in myoblasts during development, suggesting a role for ECM signals in adult myogenesis. In addition to investigating epithelial function, we will explore the possibility of modifying either the ligand or the receptor involved in ECM signaling to assess their impact on muscle development within the leg disc.

## **B. Development of the muscle innervation**

While the current project focuses on the muscular aspect of the neuromuscular system, it's important to acknowledge the role of motoneurons (MNs) within this system. Another ongoing project in our laboratory aims to unravel the molecular mechanisms governing axon-muscle communication during development, which facilitates specific innervations. Throughout development, a precise and stereotyped connection between axons and their corresponding muscle partners is established. As for muscles, we hypothesize that each MN expresses a unique gene network enabling them to communicate effectively and establish the axon-muscle connectome.

In the vertebrate spinal cord, the production of MNs along the dorso-ventral axis is mediated by graded extrinsic signals, notably the protein Shh. These signals enable the



expression of specific TFs within the MN progenitor domain, including Olig2, Nkx6.1/6.2, and Pax6. Subsequently, these cells transition to a post-mitotic state and express a new set of TFs, including Hb9, Islet1/2, and Lhx3. Along the rostro-caudal axis, morphogens such as members of the FGF family regulate the expression of different Hox genes, which in turn govern the antero-posterior identity of MNs (Dasen and Jessell, 2009). These Hox genes exhibit differential expression along the antero-posterior axis, thereby specifying MNs for fore limb and hind limb development. Subsequently, in each region, MNs are subdivided into group of MNs called pools which are compact anatomical structures that connect the same muscle target and express a given TF code (Stifani, 2014).

In the *Drosophila* model, the network of a single MN follows specific and stereotyped muscle connectivity. Studies of the larval stage neuromuscular system have highlighted specific molecules and their roles in synaptic formation (Snow and Keshishian, 1995; Nose A., et al. 1997; Winberg et al., 1998). These investigations have characterized two classes of molecules: attractive molecules (such as Connectin, Capricious, and Netrin B) and repulsive molecules (for example, Wnt4, Semaphorin II, and Toll) (Sanes and Yamagata, 2009). Among the attractive molecules, homophilic cell adhesion molecules (Connectin, Fasciclin III, and Capricious) are present in both muscles and MNs, playing roles in axon-muscle contact. Functional studies have demonstrated that mis-expression of Fasciclin III in one muscle group leads to defects in MN innervation. Similarly, ectopic expression of Connectin in muscles redirects MNs that normally project onto Connectin-expressing muscles (Nose et al., 1997). Another instance is observed with Capricious, where ectopic expression alters the target specificity of MNs (Shishido et al., 1998). Conversely, repulsive molecules such as Semaphorin II, when ectopically expressed in muscles, cause mistargeting of MNs (Masuko et al., 1999). Toll, another repulsive molecule, leads to loss of innervation by the appropriate motor axon upon ectopic expression (Rose et al., 1997). These findings collectively suggest a delicate balance between attractive and repulsive forces for growth cones, with the selection of correct targets relying on the combinatorial and simultaneous input of multiple cues.

Although significant progress has been made in understanding the mechanisms underlying axonal guidance, the molecular cues involved in the specific recognition between axons and their muscle targets remain largely elusive (Sanes and Yamagata, 2009). Notably, all mechanisms described in the *Drosophila* model regarding axon-muscle recognition have



been observed in the larval stage, where axon guidance occurs after muscle formation: muscles first emerge, followed by axon growth towards their targets (Ruiz-Cañada and Budnik, 2006). Conversely, in *Drosophila* adult legs, myogenesis of leg muscles occurs concurrently with innervation, akin to vertebrate myogenesis (Stifani, 2014). Consequently, this system offers observations in a model analogous to vertebrate limbs in terms of structure and development. Furthermore, previous studies from our laboratory have revealed that a combinatorial code of TFs expressed in each MN during development controls the unique morphology of MNs, termed morphological TFs (mTFs) (Guan et al., 2022). Indeed, reprogramming the mTF code in early developmental stages induces predictable morphological switches (Enriquez et al., 2015, 2018). However, the effectors of these codes (such as surface molecules like adhesion molecules and/or receptors) remain unidentified. Recent studies have shed light on the roles of two novel proteins: Defective proboscis extension response 10 (Dpr10), expressed in leg muscles, and Dpr-interacting protein alpha (DIP-alpha), expressed in MNs. It has been demonstrated that these proteins interact *in vitro* and stabilize stereotyped terminal axonal branches *in vivo* (Venkatasubramanian et al., 2019). Although DIP-alpha and Dpr10 are not essential for early recognition between muscle and motor axon, they are necessary to stabilize established muscle-axon connections. This example serves as the foundation of my hypothesis that muscles (and MNs) differentially express molecules ensuring specific innervation.

To elucidate the molecules underlying communication between a leg MN and its target, we propose the hypothesis that each MN and muscle expresses a unique gene network facilitating communication to establish the axon-muscle connectome. This project is conducted in collaboration with the laboratory of Pr RS Mann at Columbia University (New York), who has established single-cell sequencing of MNs. Leveraging our single-cell dataset on myoblasts during development, we aim to compare it with MN datasets to identify relationships such as ligand-receptor pairs between both datasets. Using the single myoblasts transcriptomic data we obtained, we selected candidate genes based on their differential expression to test the hypothesis that each myoblast expresses specific proteins. Initially, we are focusing on membrane proteins and secreted molecules. We will analyze their function through RNAi knockdown and overexpression to assess the impact of these modifications in muscle targeting *in vivo*. These experiments will be complemented with genetic tools enabling



labeling of both MNs and muscles in fixed tissues, as well as in live imaging experiments to observe the dynamics of recognition.

Through these experiments, we aim to unveil the dynamics of axon-muscle recognition in appendages and decipher the molecular pathways controlling unique axon-muscle connections. Understanding the general mechanisms governing muscle innervation development and identifying the machinery controlling these specificities will aid in the development of efficient therapies for repairing muscle innervation.

### **C. Maintenance of the adult leg architecture**

Through previous studies, we endeavor to comprehend how the architecture of the neuromuscular system is established during development. Each muscle or MN exhibits a distinct architecture. This unique wiring and architecture of MN axon terminals and muscles are crucial for ensuring proper muscle contraction, thereby facilitating correct movements. In adulthood, maintaining the neuromuscular architecture is imperative for sustaining locomotor function. Muscles and MN axon terminals are dynamic structures. For instance, MNs innervating *Drosophila* flight muscles undergo dismantling and reestablishment on a daily basis. Synaptic buttons decrease in number during the night, only to be restored to the correct number during the day (Mehnert et al., 2007). Additionally, peripheral nerves have the capacity to regenerate following damage from mechanical, chemical, or biological stresses. For example, ablation of wing sensory neurons in *Drosophila* leads to axon regeneration (Fang and Bonini, 2015). These examples imply the existence of a gene network that regulates the maintenance of muscle innervation across various physiological conditions or after injury, preserving the unique wiring and architecture of MN axon terminals and muscle morphologies to sustain locomotor function. Currently, the genes responsible for maintaining the specificity of this system remain elusive. We hypothesize that TFs used during development, are maintained in adulthood, play a role in the specific maintenance of muscle architecture.

In our laboratory, we have demonstrated that individual MNs express a unique code of mTFs during development, which govern the muscle fibers targeted by axons (Guan et al., 2022). Recent findings indicate that these mTF codes expressed during development are





maintained in adult flies, exemplified by genes such as *Oli* (Olig family) and *tj* (traffic jam), both of which encode mTFs. Furthermore, ongoing single nuclei sequencing of different adult leg segments in our laboratory will confirm additional instances of TF maintenance within adult muscles. Currently, we have successfully assigned a code to each muscle of the femur. To elucidate the function of genes in maintaining muscle innervation, we will disrupt their function in adult flies. We will analyze phenotypes of MN axon terminal wiring and muscle morphology at various time points during adulthood after abolishing TF function in adult MNs or muscles. Our laboratory has developed genetic tools enabling manipulation of gene expression exclusively in adults to ensure proper fly development. Subsequently, we will challenge the system by inducing gene mutations and subjecting flies to walking tests or performing laser ablation of axon branches using a UV pulsed laser microscope, monitoring axon regeneration, among other assays. We will also assess the impact of these modifications on locomotion and muscle morphology using technologies such as the Flywalker, a high-resolution video-based assay developed in our laboratory, to evaluate various parameters of fly walking behavior. This project aims to demonstrate, for the first time, how gene networks utilized during development are maintained in adulthood to actively preserve the architecture of the neuromuscular system.

#### **D. CONCLUSION**

This thesis brings new insights into the mechanisms governing the establishment of muscular architecture. We elucidate the dynamics of muscle development in appendages and the molecular pathways regulating unique morphologies. Myoblasts undergo early specification during fly development to give rise to specific muscles. This specificity is orchestrated by combination of TF codes expressed by myoblasts that are spatially organized within the disc where future muscles will form. Preliminary studies suggest that this organization is influenced by signals from the epidermis. Our further investigations will confirm the functional role of molecules expressed in both the epidermis and the subpopulations of myoblasts. Understanding the general mechanisms of muscle development and identifying the unique machinery controlling this specificity are steps toward comprehending complex mechanisms that could lead to the development of efficient therapies for diseases such as muscular dystrophies.



## BIBLIOGRAPHY

- Abu-Shaar, M., Mann, R.S., 1998. Generation of multiple antagonistic domains along the proximodistal axis during *Drosophila* leg development. *Development* 125, 3821–3830. <https://doi.org/10.1242/dev.125.19.3821>
- Alvares, L.E., Schubert, F.R., Thorpe, C., Mootoosamy, R.C., Cheng, L., Parkyn, G., Lumsden, A., Dietrich, S., 2003. Intrinsic, Hox-Dependent Cues Determine the Fate of Skeletal Muscle Precursors. *Dev. Cell* 5, 379–390. [https://doi.org/10.1016/S1534-5807\(03\)00263-6](https://doi.org/10.1016/S1534-5807(03)00263-6)
- Amthor H, Christ B, Patel K., 1999. A molecular mechanism enabling continuous embryonic muscle growth - a balance between proliferation and differentiation. *Development* 126(5):1041–53. <https://doi.org/10.1242/dev.126.5.1041>. PMID: 9927604
- Arnold, S.J., Robertson, E.J., 2009. Making a commitment: cell lineage allocation and axis patterning in the early mouse embryo. *Nat. Rev. Mol. Cell Biol.* 10, 91–103. <https://doi.org/10.1038/nrm2618>
- Azpiazu, N., Lawrence, P.A., Vincent, J.P., Frasch, M., 1996. Segmentation and specification of the *Drosophila* mesoderm. *Genes Dev.* 10, 3183–3194. <https://doi.org/10.1101/gad.10.24.3183>
- Baek, M., Mann, R.S., 2009. Lineage and Birth Date Specify Motor Neuron Targeting and Dendritic Architecture in Adult *Drosophila*. *J. Neurosci.* 29, 6904–6916. <https://doi.org/10.1523/JNEUROSCI.1585-09.2009>
- Basler, K., 1994. Compartment boundaries and the control of *Drosophila* limb pattern by hedgehog protein 368.
- Bate, M., n.d. The embryonic development of larval muscles in *Drosophila* 14.
- Bate, M. (1993). The mesoderm and its derivatives. In *The Development of Drosophila melanogaster*. Vol. 2 (ed. M. Bate and A. Martinez-Arias), pp. 10131090. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press., n.d.
- Baylies, M.K., Bate, M., 1996. **twist** : A Myogenic Switch in *Drosophila* . *Science* 272, 1481–1484. <https://doi.org/10.1126/science.272.5267.1481>



Baylies, M.K., Bate, M., Gomez, M.R., 1998. Myogenesis: A View from *Drosophila*. *Cell* 93, 921–927. [https://doi.org/10.1016/S0092-8674\(00\)81198-8](https://doi.org/10.1016/S0092-8674(00)81198-8)

Bentzinger, Cf., Von Maltzahn, J., Rudnicki, M.A., 2010. Extrinsic regulation of satellite cell specification. *Stem Cell Res. Ther.* 1, 27. <https://doi.org/10.1186/scrt27>

Bentzinger, C.F., Wang, Y.X., Rudnicki, M.A., 2012. Building Muscle: Molecular Regulation of Myogenesis. *Cold Spring Harb. Perspect. Biol.* 4, a008342–a008342. <https://doi.org/10.1101/cshperspect.a008342>

Besse, L., Sheeba, C.J., Holt, M., Labuhn, M., Wilde, S., Feneck, E., Bell, D., Kucharska, A., Logan, M.P.O., 2020. Individual Limb Muscle Bundles Are Formed through Progressive Steps Orchestrated by Adjacent Connective Tissue Cells during Primary Myogenesis. *Cell Rep.* 30, 3552-3565.e6. <https://doi.org/10.1016/j.celrep.2020.02.037>

Biressi, S., Molinaro, M., Cossu, G., 2007. Cellular heterogeneity during vertebrate skeletal muscle development. *Dev. Biol.* 308, 281–293. <https://doi.org/10.1016/j.ydbio.2007.06.006>

Bladt, F., D. Riethmacher, et al., 1995. Essential role for the c-met receptor in the migration of myogenic precursor cells into the limb bud. *Nature*.

Borchiellini, C., Coulon, J., Le Parco, Y., 1996. The function of type IV collagen during *Drosophila* muscle development. *Mech. Dev.* 58, 179–191. [https://doi.org/10.1016/S0925-4773\(96\)00574-6](https://doi.org/10.1016/S0925-4773(96)00574-6)

Borycki, A.-G., Brunk, B., Tajbakhsh, S., Buckingham, M., Chiang, C., Emerson, C.P., 1999. Sonic hedgehog controls epaxial muscle determination through *Myf5* activation. *Development* 126, 4053–4063. <https://doi.org/10.1242/dev.126.18.4053>

Bour, B.A., Chakravarti, M., West, J.M., Abmayr, S.M., 2000. *Drosophila* SNS, a member of the immunoglobulin superfamily that is essential for myoblast fusion. *Genes Dev.* 14, 1498–1511. <https://doi.org/10.1101/gad.14.12.1498>

Bourgouin, C., Lundgren, S.E., Thomas, J.B., 1992. *apterous* is a *drosophila* LIM domain gene required for the development of a subset of embryonic muscles. *Neuron* 9, 549–561. [https://doi.org/10.1016/0896-6273\(92\)90192-G](https://doi.org/10.1016/0896-6273(92)90192-G)



Brierley, D.J., Rathore, K., VijayRaghavan, K., Williams, D.W., 2012. Developmental origins and architecture of *Drosophila* leg motoneurons. *J. Comp. Neurol.* 520, 1629–1649.

<https://doi.org/10.1002/cne.23003>

Broadie, K.S., Bate, M., n.d. The development of adult muscles in *Drosophila*: ablation of identified muscle precursor cells.

Brook, W.J., Cohen, S.M., 1996. Antagonistic Interactions Between Wingless and Decapentaplegic Responsible for Dorsal-Ventral Pattern in the *Drosophila* Leg. *Science* 273, 1373–1377. <https://doi.org/10.1126/science.273.5280.1373>

Bryson-Richardson, R.J., Currie, P.D., 2008. The genetics of vertebrate myogenesis. *Nat. Rev. Genet.* 9, 632–646. <https://doi.org/10.1038/nrg2369>

Buckingham, M., 2001. Skeletal muscle formation in vertebrates. *Curr. Opin. Genet. Dev.* 11, 440–448. [https://doi.org/10.1016/S0959-437X\(00\)00215-X](https://doi.org/10.1016/S0959-437X(00)00215-X)

Buckingham, M., Bajard, L., Chang, T., Daubas, P., Hadchouel, J., Meilhac, S., Montarras, D., Rocancourt, D., Relaix, F., 2003. The formation of skeletal muscle: from somite to limb. *J. Anat.* 202, 59–68. <https://doi.org/10.1046/j.1469-7580.2003.00139.x>

Buckingham M, Relaix F., 2007. The role of Pax genes in the development of tissues and organs: Pax3 and Pax7 regulate muscle progenitor cell functions. *Annu Rev Cell Dev Biol.* <https://doi.org/10.1146/annurev.cellbio.23.090506.123438>. PMID: 17506689

Buckingham, M., Rigby, P.W.J., 2014. Gene Regulatory Networks and Transcriptional Mechanisms that Control Myogenesis. *Dev. Cell* 28, 225–238.

<https://doi.org/10.1016/j.devcel.2013.12.020>

Bumcrot, D.A., McMahon, A.P., 1995. Somite Differentiation: Sonic signals somites. *Curr. Biol.* 5, 612–614. [https://doi.org/10.1016/S0960-9822\(95\)00123-0](https://doi.org/10.1016/S0960-9822(95)00123-0)

Cairns, D.M., Sato, M.E., Lee, P.G., Lassar, A.B., Zeng, L., 2008. A gradient of Shh establishes mutually repressing somitic cell fates induced by Nkx3.2 and Pax3. *Dev. Biol.* 323, 152–165.

<https://doi.org/10.1016/j.ydbio.2008.08.024>

Camille Viaut, Andrea Münsterberg, 2020. Fine-tuning of the PAX-SIX-EYA-DACH network by multiple microRNAs controls embryo myogenesis. *bioRxiv*.

<https://doi.org/10.1101/2020.02.18.954446>





Campbell G, Weaver T, Tomlinson A., 1993. Axis specification in the developing *Drosophila* appendage: the role of wingless, decapentaplegic, and the homeobox gene aristaless. *Cell* 74(6):1113–23. [https://doi.org/10.1016/0092-8674\(93\)90732-6](https://doi.org/10.1016/0092-8674(93)90732-6)

Carmena, A., Bate, M., Jiménez, F., 1995. Lethal of scute, a proneural gene, participates in the specification of muscle progenitors during *Drosophila* embryogenesis. *Genes Dev.* 9, 2373–2383. <https://doi.org/10.1101/gad.9.19.2373>

Chaturvedi, V., Dye, D.E., Kinnear, B.F., Van Kuppevelt, T.H., Grounds, M.D., Coombe, D.R., 2015. Interactions between Skeletal Muscle Myoblasts and their Extracellular Matrix Revealed by a Serum Free Culture System. *PLOS ONE* 10, e0127675. <https://doi.org/10.1371/journal.pone.0127675>

Chen, K.H., Boettiger, A.N., Moffitt, J.R., Wang, S., Zhuang, X., 2015. Spatially resolved, highly multiplexed RNA profiling in single cells. *Science* 348, aaa6090. <https://doi.org/10.1126/science.aaa6090>

Chevallier, A., M. Kieny, et al., 1977. Limb-somite relationship: origin of the limb musculature. *J Embryol Exp Morphol.*

Christ B, Ordahl CP, 1995. Early stages of chick somite development. *Anat Embryol (Berl).* <https://doi.org/10.1007/BF00304424>. PMID: 7625610

Collins, C. A., Olsen, I., Zammit, P. S., Heslop, L., Petrie, A., Partridge, T. A. and Morgan, J. E., 2005. Stem cell function, self-renewal, and behavioral heterogeneity of cells from the adult muscle satellite cell niche.

Condic, M.L., Fristrom, D., Fristrom, J.W., 1991. Apical cell shape changes during *Drosophila* imaginal leg disc elongation: a novel morphogenetic mechanism. *Development* 111, 23–33. <https://doi.org/10.1242/dev.111.1.23>

Dasen, J.S., Jessell, T.M., 2009. Chapter Six Hox Networks and the Origins of Motor Neuron Diversity, in: *Current Topics in Developmental Biology*. Elsevier, pp. 169–200. [https://doi.org/10.1016/S0070-2153\(09\)88006-X](https://doi.org/10.1016/S0070-2153(09)88006-X)

Davis, R.L., Weintraub, H., Lassar, A.B., n.d. Expression of a Single Transfected cDNA Converts Fibroblasts to Myoblasts.



Diaz-Benjumea, F.J., Cohen, B., Cohen, S.M., 1994. Cell interaction between compartments establishes the proximal-distal axis of *Drosophila* legs. *Nature* 372, 175–179.

<https://doi.org/10.1038/372175a0>

Dietrich, S., Abou-Rebyeh, F., Brohmann, H., Bladt, F., Sonnenberg-Riethmacher, E., Yamaai, T., Lumsden, A., Brand-Saber, B., Birchmeier, C., 1999. The role of SF/HGF and c-Met in the development of skeletal muscle. *Development* 126, 1621–1629.

<https://doi.org/10.1242/dev.126.8.1621>

Doberstein, S.K., Fetter, R.D., Mehta, A.Y., Goodman, C.S., 1997. Genetic Analysis of Myoblast Fusion: *blown fuse* Is Required for Progression Beyond the Prefusion Complex. *J. Cell Biol.* 136, 1249–1261. <https://doi.org/10.1083/jcb.136.6.1249>

Dobi, K.C., Schulman, V.K., Baylies, M.K., 2015. Specification of the somatic musculature in *Drosophila*: Specification of the somatic musculature in *Drosophila*. *Wiley Interdiscip. Rev. Dev. Biol.* 4, 357–375. <https://doi.org/10.1002/wdev.182>

Duan, H., Zhang, C., Chen, J., Sink, H., Frei, E., Noll, M., 2007. A key role of *Pox meso* in somatic myogenesis of *Drosophila*. *Development* 134, 3985–3997.

<https://doi.org/10.1242/dev.008821>

Dunglison, G.F., Scotting, P.J., Wigmore, P.M., 1999. Rat embryonic myoblasts are restricted to forming primary myofibers while later myogenic populations are pluripotent. *Mech. Dev.*

Duxson, M.J., Usson, Y., Harris, A.J., 1989. The origin of secondary myotubes in mammalian skeletal muscles: ultrastructural studies. *Development* 107, 743–750.

<https://doi.org/10.1242/dev.107.4.743>

Dworak, H.A., Charles, M.A., Pellerano, L.B., Sink, H., 2001. Characterization of *Drosophila hibris*, a gene related to human nephrin. *Development* 128, 4265–4276.

<https://doi.org/10.1242/dev.128.21.4265>

E. B. Lewis, 1978. A gene complex controlling segmentation in *Drosophila*.

Enriquez, J., De Taffin, M., Crozatier, M., Vincent, A., Dubois, L., 2012. Combinatorial coding of *Drosophila* muscle shape by Collier and Nautilus. *Dev. Biol.* 363, 27–39.

<https://doi.org/10.1016/j.ydbio.2011.12.018>



- Enriquez, J., Rio, L.Q., Blazeski, R., Bellemin, S., Godement, P., Mason, C., Mann, R.S., 2018. Differing Strategies Despite Shared Lineages of Motor Neurons and Glia to Achieve Robust Development of an Adult Neuropil in *Drosophila*. *Neuron* 97, 538-554.e5.  
<https://doi.org/10.1016/j.neuron.2018.01.007>
- Enriquez, J., Venkatasubramanian, L., Baek, M., Peterson, M., Aghayeva, U., Mann, R.S., 2015. Specification of Individual Adult Motor Neuron Morphologies by Combinatorial Transcription Factor Codes. *Neuron* 86, 955–970.  
<https://doi.org/10.1016/j.neuron.2015.04.011>
- Estella, C., Voutev, R., Mann, R.S., 2012. A Dynamic Network of Morphogens and Transcription Factors Patterns the Fly Leg, in: *Current Topics in Developmental Biology*. Elsevier, pp. 173–198. <https://doi.org/10.1016/B978-0-12-386499-4.00007-0>
- Fan, C.-M., Tessier-Lavigne, M., 1994. Patterning of mammalian somites by surface ectoderm and notochord: Evidence for sclerotome induction by a hedgehog homolog. *Cell* 79, 1175–1186. [https://doi.org/10.1016/0092-8674\(94\)90009-4](https://doi.org/10.1016/0092-8674(94)90009-4)
- Fang, Y., Bonini, N., 2015. Hope on the (fruit) fly: the *Drosophila* wing paradigm of axon injury. *Neural Regen. Res.* 10, 173. <https://doi.org/10.4103/1673-5374.152359>
- Ferguson, E.L., Anderson, K.V., 1992. decapentaplegic acts as a morphogen to organize dorsal-ventral pattern in the *Drosophila* embryo. *Cell* 71, 451–461.  
[https://doi.org/10.1016/0092-8674\(92\)90514-D](https://doi.org/10.1016/0092-8674(92)90514-D)
- Fernandes J, Bate M, Vijayraghavan K., 1991. Development of the indirect flight muscles of *Drosophila*. *Development*.
- Ferretti, E., Hadjantonakis, A.-K., 2019. Mesoderm specification and diversification: from single cells to emergent tissues. *Curr. Opin. Cell Biol.* 61, 110–116.  
<https://doi.org/10.1016/j.ceb.2019.07.012>
- Galindo, M.I., Bishop, S.A., Greig, S., Couso, J.P., 2002. Leg Patterning Driven by Proximal-Distal Interactions and EGFR Signaling. *Science* 297, 256–259.  
<https://doi.org/10.1126/science.1072311>



Gal-Levi, R., Leshem, Y., Aoki, S., Nakamura, T., Halevy, O., 1998. Hepatocyte growth factor plays a dual role in regulating skeletal muscle satellite cell proliferation and differentiation. *Biochim. Biophys. Acta BBA - Mol. Cell Res.* 1402, 39–51. [https://doi.org/10.1016/S0167-4889\(97\)00124-9](https://doi.org/10.1016/S0167-4889(97)00124-9)

Ghazi, A., Vijayraghavan, K., n.d. *Muscle Development in Drosophila*.

González-Crespo, S., Morata, G., 1996. Genetic evidence for the subdivision of the arthropod limb into coxopodite and telopodite. *Development* 122, 3921–3928.

<https://doi.org/10.1242/dev.122.12.3921>

Gotti, C., Sensini, A., Zucchelli, A., Carloni, R., Focarete, M.L., 2020. Hierarchical fibrous structures for muscle-inspired soft-actuators: A review. *Appl. Mater. Today* 20, 100772.

<https://doi.org/10.1016/j.apmt.2020.100772>

Goulding, M.D., Chalepakis, G., Deutsch, U., Erselius, J.R., Gruss, P., 1991. Pax-3, a novel murine DNA binding protein expressed during early neurogenesis. *EMBO J.* 10, 1135–1147.

<https://doi.org/10.1002/j.1460-2075.1991.tb08054.x>

Guan, W., Bellemin, S., Bouchet, M., Venkatasubramanian, L., Guillermin, C., Laurençon, A., Kabir, C., Darnas, A., Godin, C., Urdy, S., Mann, R.S., Enriquez, J., 2022. Post-transcriptional regulation of transcription factor codes in immature neurons drives neuronal diversity. *Cell Rep.* 39, 110992.

<https://doi.org/10.1016/j.celrep.2022.110992>

Gullberg D, Ekblom P., 1995. Extracellular matrix and its receptors during development. *Int J Dev Biol.*

Gunage, R.D., Dhanyasi, N., Reichert, H., VijayRaghavan, K., 2017. *Drosophila* adult muscle development and regeneration. *Semin. Cell Dev. Biol.* 72, 56–66.

<https://doi.org/10.1016/j.semcdb.2017.11.017>

Guo, M., Jan, L.Y., Jan, Y.N., n.d. Control of Daughter Cell Fates during Asymmetric Division: Interaction of Numb and Notch.

Hirsinger, E., Duprez, D., Jouve, C., Malapert, P., Cooke, J., Pourquié, O., 1997. Noggin acts downstream of Wnt and Sonic Hedgehog to antagonize BMP4 in avian somite patterning. *Development* 124, 4605–4614.

<https://doi.org/10.1242/dev.124.22.4605>





- Hopkins, B.R., Barmina, O., Kopp, A., 2022. Charting the development of *Drosophila* leg sensory organs at single-cell resolution (preprint). *Developmental Biology*.  
<https://doi.org/10.1101/2022.10.07.511357>
- Houzelstein, D., Auda-Boucher, G., Chéraud, Y., Rouaud, T., Blanc, I., Tajbakhsh, S., Buckingham, M.E., Fontaine-Péru, J., Robert, B., 1999. The homeobox gene *Msx1* is expressed in a subset of somites, and in muscle progenitor cells migrating into the forelimb. *Development* 126, 2689–2701. <https://doi.org/10.1242/dev.126.12.2689>
- Ikeya, M., Takada, S., 1998. Wnt signaling from the dorsal neural tube is required for the formation of the medial dermomyotome. *Development* 125, 4969–4976.  
<https://doi.org/10.1242/dev.125.24.4969>
- Jostes, B., Walther, C., Gruss, P., 1990. The murine paired box gene, *Pax7*, is expressed specifically during the development of the nervous and muscular system. *Mech. Dev.* 33, 27–37. [https://doi.org/10.1016/0925-4773\(90\)90132-6](https://doi.org/10.1016/0925-4773(90)90132-6)
- Kahane, N., Cinnamon, Y., Bachelet, I., Kalcheim, C., 2001. The third wave of myotome colonization by mitotically competent progenitors: regulating the balance between differentiation and proliferation during muscle development. *Development* 128, 2187–2198.  
<https://doi.org/10.1242/dev.128.12.2187>
- Kassar-Duchossoy, L., Giacone, E., Gayraud-Morel, B., Jory, A., Gomès, D., Tajbakhsh, S., 2005. *Pax3/Pax7* mark a novel population of primitive myogenic cells during development. *Genes Dev.* 19, 1426–1431. <https://doi.org/10.1101/gad.345505>
- Kelly, R.G., Jerome-Majewska, L.A., Papaioannou, V.E., 2004. The del22q11.2 candidate gene *Tbx1* regulates branchiomic myogenesis. *Hum. Mol. Genet.* 13, 2829–2840.  
<https://doi.org/10.1093/hmg/ddh304>
- Kojima, T., 2004. The mechanism of *Drosophila* leg development along the proximodistal axis. *Dev. Growth Differ.* 46, 115–129. <https://doi.org/10.1111/j.1440-169X.2004.00735.x>
- Kopan, R., Nye, J.S., Weintraub, H., 1994. The intracellular domain of mouse Notch: a constitutively activated repressor of myogenesis directed at the basic helix-loop-helix region of MyoD. *Development* 120, 2385–2396. <https://doi.org/10.1242/dev.120.9.2385>



- Krumlauf, R., 1994. Hox genes in vertebrate development. *Cell* 78, 191–201.  
[https://doi.org/10.1016/0092-8674\(94\)90290-9](https://doi.org/10.1016/0092-8674(94)90290-9)
- LaBeau, E.M., Trujillo, D.L., Cripps, R.M., 2009. Bithorax Complex genes control alary muscle patterning along the cardiac tube of *Drosophila*. *Mech. Dev.* 126, 478–486.  
<https://doi.org/10.1016/j.mod.2009.01.001>
- Laurichesse, Q., Soler, C., 2020. Muscle development : a view from adult myogenesis in *Drosophila*. *Semin. Cell Dev. Biol.* 104, 39–50. <https://doi.org/10.1016/j.semcd.2020.02.009>
- Lecuit, T., Cohen, S.M., 1997. Proximal–distal axis formation in the *Drosophila* leg. *Nature* 388, 139–145. <https://doi.org/10.1038/40563>
- Lehmacher, C., Abeln, B., Paululat, A., 2012. The ultrastructure of *Drosophila* heart cells. *Arthropod Struct. Dev.* 41, 459–474. <https://doi.org/10.1016/j.asd.2012.02.002>
- Lilly, B., Galewsky, S., Firulli, A.B., Schulz, R.A., Olson, E.N., 1994. D-MEF2: a MADS box transcription factor expressed in differentiating mesoderm and muscle cell lineages during *Drosophila* embryogenesis. *Proc. Natl. Acad. Sci.* 91, 5662–5666.  
<https://doi.org/10.1073/pnas.91.12.5662>
- MacKrell, A.J., Blumberg, B., Haynes, S.R., Fessler, J.H., 1988. The lethal myospheroid gene of *Drosophila* encodes a membrane protein homologous to vertebrate integrin beta subunits. *Proc. Natl. Acad. Sci.* 85, 2633–2637. <https://doi.org/10.1073/pnas.85.8.2633>
- Mallo, M., 2018. Reassessing the Role of Hox Genes during Vertebrate Development and Evolution. *Trends Genet.* 34, 209–217. <https://doi.org/10.1016/j.tig.2017.11.007>
- Marcelle, C., Stark, M.R., Bronner-Fraser, M., 1997. Coordinate actions of BMPs, Wnts, Shh and Noggin mediate patterning of the dorsal somite. *Development* 124, 3955–3963.  
<https://doi.org/10.1242/dev.124.20.3955>
- Masuko, N., Makino, K., Kuwahara, H., Fukunaga, K., Sudo, T., Araki, N., Yamamoto, H., Yamada, Y., Miyamoto, E., Saya, H., 1999. Interaction of NE-dlg/SAP102, a Neuronal and Endocrine Tissue-specific Membrane-associated Guanylate Kinase Protein, with Calmodulin and PSD-95/SAP90. *J. Biol. Chem.* 274, 5782–5790. <https://doi.org/10.1074/jbc.274.9.5782>
- Mauro, A., 1961. Satellite cell of skeletal muscle fibers. *J Biophys Biochem Cytol.*



Mehnert, K.I., Beramendi, A., Elghazali, F., Negro, P., Kyriacou, C.P., Cantera, R., 2007.

Circadian changes in *Drosophila* motor terminals. *Dev. Neurobiol.* 67, 415–421.

<https://doi.org/10.1002/dneu.20332>

Michelson, A.M., 1994. Muscle pattern diversification in *Drosophila* is determined by the autonomous function of homeotic genes in the embryonic mesoderm. *Development* 120,

755–768. <https://doi.org/10.1242/dev.120.4.755>

Miller, A. (1950). The internal anatomy and histology of the imago of *Drosophila melanogaster*. In *Biology of Drosophila* (ed. M. Demerec), pp. 420-531. New York, NY: John Wiley & Sons.

Miyazono, K., Maeda, S., Imamura, T., 2005. BMP receptor signaling: Transcriptional targets, regulation of signals, and signaling cross-talk. *Cytokine Growth Factor Rev.* 16, 251–263.

<https://doi.org/10.1016/j.cytogfr.2005.01.009>

Molkentin, J.D., Black, B.L., Martin, J.F., Olson, E.N., 1995. Cooperative activation of muscle gene expression by MEF2 and myogenic bHLH proteins. *Cell* 83, 1125–1136.

[https://doi.org/10.1016/0092-8674\(95\)90139-6](https://doi.org/10.1016/0092-8674(95)90139-6)

Montarras D, Chelly J, Bober E, Arnold H, Ott MO, Gros F, Pinset C., 1991. Developmental patterns in the expression of Myf5, MyoD, myogenin, and MRF4 during myogenesis. *New Biol* 3(6):592-600. <https://doi.org/1911647>

<https://doi.org/1911647>

Musumeci, G., Castrogiovanni, P., Coleman, R., Szychlinska, M.A., Salvatorelli, L., Parenti, R., Magro, G., Imbesi, R., 2015. Somitogenesis: From somite to skeletal muscle. *Acta Histochem.* 117, 313–328. <https://doi.org/10.1016/j.acthis.2015.02.011>

<https://doi.org/10.1016/j.acthis.2015.02.011>

Neumann, C.J., Cohen, S.M., 1996. A hierarchy of cross-regulation involving *Notch*, *wingless*, *vestigial* and *cut* organizes the dorsal/ventral axis of the *Drosophila* wing. *Development* 122, 3477–3485. <https://doi.org/10.1242/dev.122.11.3477>

<https://doi.org/10.1242/dev.122.11.3477>

Noden, D.M., 1986. Patterning of avian craniofacial muscles. *Dev Biol.*

Noden, D.M., 1983a. The embryonic origins of avian cephalic and cervical muscles and associated connective tissues. *Am J Anat.*

Noden, D.M., 1983b. The role of the neural crest in patterning of avian cranial skeletal, connective, and muscle tissues. *Dev Biol.*



Nose, A., Umeda, T., and Takeichi, M., 1997. Neuromuscular target recognition by a homophilic interaction of connectin cell adhesion molecules in *Drosophila*. *Development* Volume 124, Issue 8.

Nowicki, J.L., Burke, A.C., 2000. *Hox* genes and morphological identity: axial versus lateral patterning in the vertebrate mesoderm. *Development* 127, 4265–4275.

<https://doi.org/10.1242/dev.127.19.4265>

Pearson, J.C., Lemons, D., McGinnis, W., 2005. Modulating *Hox* gene functions during animal body patterning. *Nat. Rev. Genet.* 6, 893–904. <https://doi.org/10.1038/nrg1726>

Penton, A., Hoffmann, F.M., 1996. Decapentaplegic restricts the domain of wingless during *Drosophila* limb patterning. *Nature* 382, 162–165. <https://doi.org/10.1038/382162a0>

Potthoff, M.J., Olson, E.N., 2007. MEF2: a central regulator of diverse developmental programs. *Development* 134, 4131–4140. <https://doi.org/10.1242/dev.008367>

Pownall ME, Gustafsson MK, Emerson CP Jr., 2002. Myogenic regulatory factors and the specification of muscle progenitors in vertebrate embryos. *Annu Rev Cell Dev Biol.*

<https://doi.org/10.1146/annurev.cellbio.18.012502.105758>

Relaix, F., Montarras, D., Zaffran, S., Gayraud-Morel, B., Rocancourt, D., Tajbakhsh, S., Mansouri, A., Cumano, A., Buckingham, M., 2006. Pax3 and Pax7 have distinct and overlapping functions in adult muscle progenitor cells. *J. Cell Biol.* 172, 91–102.

<https://doi.org/10.1083/jcb.200508044>

Relaix, F., Rocancourt, D., Mansouri, A., Buckingham, M., 2004. Divergent functions of murine Pax3 and Pax7 in limb muscle development. *Genes Dev.* 18, 1088–1105.

<https://doi.org/10.1101/gad.301004>

Riechmann, V., Irion, U., Wilson, R., Grosskortenhaus, R., Leptin, M., 1997. Control of cell fates and segmentation in the *Drosophila* mesoderm. *Development* 124, 2915–2922.

<https://doi.org/10.1242/dev.124.15.2915>

Rochlin, K., Yu, S., Roy, S., Baylies, M.K., 2010. Myoblast fusion: When it takes more to make one. *Dev. Biol.* 341, 66–83. <https://doi.org/10.1016/j.ydbio.2009.10.024>





- Rose, D., Zhu, X., Kose, H., Hoang, B., Cho, J., Chiba, A., 1997. Toll, a muscle cell surface molecule, locally inhibits synaptic initiation of the RP3 motoneuron growth cone in *Drosophila*. *Development* 124, 1561–1571. <https://doi.org/10.1242/dev.124.8.1561>
- Rout, P., Preußner, M., Önel, S.F., 2022. *Drosophila melanogaster*: A Model System to Study Distinct Genetic Programs in Myoblast Fusion. *Cells* 11, 321. <https://doi.org/10.3390/cells11030321>
- Roy, S., Shashidhara, L.S., VijayRaghavan, K., 1997. Muscles in the *Drosophila* second thoracic segment are patterned independently of autonomous homeotic gene function. *Curr. Biol.* 7, 222–227. [https://doi.org/10.1016/S0960-9822\(06\)00117-5](https://doi.org/10.1016/S0960-9822(06)00117-5)
- Roy, S., VijayRaghavan, K., 1999. Muscle pattern diversification in *Drosophila*: the story of imaginal myogenesis. *BioEssays* 21, 486–498. [https://doi.org/10.1002/\(SICI\)1521-1878\(199906\)21:6<486::AID-BIES5>3.0.CO;2-M](https://doi.org/10.1002/(SICI)1521-1878(199906)21:6<486::AID-BIES5>3.0.CO;2-M)
- Ruiz-Cañada, C., and Budnik, V., 2006. Introduction on the use of the *Drosophila* embryonic/larval neuromuscular junction as a model system to study synapse development and function, and a brief summary of pathfinding and target recognition. *Int. Rev. Neurobiol* Volume 75, 1-31. [https://doi.org/10.1016/S0074-7742\(06\)75001-2](https://doi.org/10.1016/S0074-7742(06)75001-2)
- Ruiz-Gomez, M., Coutts, N., Price, A., Taylor, M. V. and Bate, M., 2000. *Drosophila* dumbfounded: a myoblast attractant essential for fusion.
- Ruiz-Gómez, M., Romani, S., Hartmann, C., Jäckle, H., Bate, M., 1997. Specific muscle identities are regulated by *Krüppel* during *Drosophila* embryogenesis. *Development* 124, 3407–3414. <https://doi.org/10.1242/dev.124.17.3407>
- Sanes, J.R., Yamagata, M., 2009. Many Paths to Synaptic Specificity. *Annu. Rev. Cell Dev. Biol.* 25, 161–195. <https://doi.org/10.1146/annurev.cellbio.24.110707.175402>
- Schiaffino, S., Reggiani, C., 2011. Fiber Types in Mammalian Skeletal Muscles. *Physiol. Rev.* 91, 1447–1531. <https://doi.org/10.1152/physrev.00031.2010>
- Scime, A., 2009. Advances in myogenic cell transplantation and skeletal muscle tissue engineering. *Front. Biosci.* Volume, 3012. <https://doi.org/10.2741/3431>



Shih, H.P., Gross, M.K., Kioussi, C., 2007. Cranial muscle defects of Pitx2 mutants result from specification defects in the first branchial arch. *Proc. Natl. Acad. Sci.* 104, 5907–5912. <https://doi.org/10.1073/pnas.0701122104>

Shishido, E., Takeichi, M., Nose, A., 1998. *Drosophila* Synapse Formation: Regulation by Transmembrane Protein with Leu-Rich Repeats, CAPRICIOUS. *Science* 280, 2118–2121. <https://doi.org/10.1126/science.280.5372.2118>

Snow, P., Keshishian, H., 1995. Fasciclin III as a synaptic target recognition molecule in *Drosophila* 374.

Soler, C., 2004. Coordinated development of muscles and tendons of the *Drosophila* leg. *Development* 131, 6041–6051. <https://doi.org/10.1242/dev.01527>

Spletter, M.L., Schnorrer, F., 2014. Transcriptional regulation and alternative splicing cooperate in muscle fiber-type specification in flies and mammals. *Exp. Cell Res.* 321, 90–98. <https://doi.org/10.1016/j.yexcr.2013.10.007>

Squire, J.M., 1997. Architecture and function in the muscle sarcomere. *Curr Opin Struct Biol.* [https://doi.org/10.1016/s0959-440x\(97\)80033-4](https://doi.org/10.1016/s0959-440x(97)80033-4). PMID: 9094325

Staebling-Hampton, K., Hoffmann, F.M., Baylies, M.K., Rushton, E., Bate, M., 1994. dpp induces mesodermal gene expression in *Drosophila*. *Nature* 372, 783–786. <https://doi.org/10.1038/372783a0>

Stifani, N., 2014. Motor neurons and the generation of spinal motor neuron diversity. *Front. Cell. Neurosci.* 8. <https://doi.org/10.3389/fncel.2014.00293>

Strükelberg, M., n.d. Rst in muscle development.

Tajbakhsh, S., Rocancourt, D., Cossu, G., Buckingham, M., 1997. Redefining the Genetic Hierarchies Controlling Skeletal Myogenesis: Pax-3 and Myf-5 Act Upstream of MyoD. *Cell* 89, 127–138. [https://doi.org/10.1016/S0092-8674\(00\)80189-0](https://doi.org/10.1016/S0092-8674(00)80189-0)

Tixier, V., Bataillé, L., Jagla, K., 2010. Diversification of muscle types: Recent insights from *Drosophila*. *Exp. Cell Res.* 316, 3019–3027. <https://doi.org/10.1016/j.yexcr.2010.07.013>

Vasyutina, E., Birchmeier, C., 2006. The development of migrating muscle precursor cells. *Brain Struct. Funct.* 211, 37–41. <https://doi.org/10.1007/s00429-006-0118-9>



- Venkatasubramanian, L., Guo, Z., Xu, S., Tan, L., Xiao, Q., Nagarkar-Jaiswal, S., Mann, R.S., n.d. Stereotyped terminal axon branching of leg motor neurons mediated by IgSF proteins DIP-a and Dpr10 27.
- Venuti, J.M., Morris, J.H., Vivian, J.L., Olson, E.N., Klein, W.H., 1995. Myogenin is required for late but not early aspects of myogenesis during mouse development. *J. Cell Biol.* 128, 563–576. <https://doi.org/10.1083/jcb.128.4.563>
- Welcome Bender et al., 1983. Molecular Genetics of the Bithorax Complex in *Drosophila melanogaster*. *Science* 221,23–29. <https://doi.org/10.1126/science.221.4605.23>
- Wigmore, P.M., Evans, D.J.R., 2002. Molecular and cellular mechanisms involved in the generation of fiber diversity during myogenesis, in: *International Review of Cytology*. Academic Press, pp. 175–232. [https://doi.org/10.1016/S0074-7696\(02\)16006-2](https://doi.org/10.1016/S0074-7696(02)16006-2)
- Winberg, M.L., Mitchell, K.J., Goodman, C.S., 1998. Genetic Analysis of the Mechanisms Controlling Target Selection: Complementary and Combinatorial Functions of Netrins, Semaphorins, and IgCAMs. *Cell* 93, 581–591. [https://doi.org/10.1016/S0092-8674\(00\)81187-3](https://doi.org/10.1016/S0092-8674(00)81187-3)
- Wu, J., Cohen, S.M., 1999. Proximodistal axis formation in the *Drosophila* leg: subdivision into proximal and distal domains by Homothorax and Distal-less. *Development* 126, 109–117. <https://doi.org/10.1242/dev.126.1.109>
- Yin, Z., Xu, X.-L., Frasch, M., 1997. Regulation of the Twist target gene *tinman* by modular *cis*-regulatory elements during early mesoderm development. *Development* 124, 4971–4982. <https://doi.org/10.1242/dev.124.24.4971>
- Yu, M., Wang, H., Xu, Y., Yu, D., Li, D., Liu, X., Du, W., 2015. Insulin-like growth factor-1 (IGF-1) promotes myoblast proliferation and skeletal muscle growth of embryonic chickens via the PI3K/Akt signalling pathway: IGF-1 stimulates myoblast proliferation. *Cell Biol. Int.* 39, 910–922. <https://doi.org/10.1002/cbin.10466>
- Zaffran, S., Reim, I., Qian, L., Lo, P.C., Bodmer, R., Frasch, M., 2006. Cardioblast-intrinsic Tinman activity controls proper diversification and differentiation of myocardial cells in *Drosophila*. *Development* 133, 4073–4083. <https://doi.org/10.1242/dev.02586>

