



**HAL**  
open science

# Analyses moléculaires et fonctionnelles de la diversification musculaire par de nouvelles approches génomiques cellules-spécifiques chez la drosophile.

Benjamin Bertin

► **To cite this version:**

Benjamin Bertin. Analyses moléculaires et fonctionnelles de la diversification musculaire par de nouvelles approches génomiques cellules-spécifiques chez la drosophile.. Médecine humaine et pathologie. Université Clermont Auvergne [2017-2020], 2017. Français. NNT : 2017CLFAS002 . tel-01894672

**HAL Id: tel-01894672**

**<https://theses.hal.science/tel-01894672>**

Submitted on 12 Oct 2018

**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.

UNIVERSITE BLAISE PASCAL

UNIVERSITE D'AUVERGNE

Année 2017

N° d'ordre

*ECOLE DOCTORALE  
DES SCIENCES DE LA VIE ET DE LA SANTE*

N° d'ordre :

*Thèse*

Présentée à l'Université d'Auvergne

pour l'obtention du grade de DOCTEUR  
(Décret du 5 juillet 1984)

Spécialité Génétique et développement

Soutenue publiquement le Vendredi 29 Septembre 2017

**BERTIN Benjamin**

---

**Analyses moléculaires et fonctionnelles de la diversification musculaire  
par des approches génomiques cellules-spécifiques chez la Drosophile**

---

Membres du JURY

Président du jury	DR. CROZATIER Michèle
Rapporteur	DR. MAIRE Pascal
Examinatrice	CR. GHAVI-HELM Yad
Membre invité	CR. JUNION Guillaume
Directeur de thèse	DR. JAGLA Krzysztof

Laboratoire GReD (Génétique, Reproduction et Développement)  
CNRS (UMR6293) – INSERM (U1103)  
Université Clermont Auvergne



## Remerciements

Mes premiers remerciements sont destinés à Krzysztof Jagla, mon directeur de thèse qui m'a accueilli dans son équipe et qui m'a permis de m'épanouir au sein du laboratoire. Merci pour la confiance que tu m'as accordé ainsi que pour ton aide et tes encouragements.

J'aimerais tout particulièrement remercier Guillaume Junion avec qui j'ai commencé cette aventure. Merci pour ton soutien, tes conseils avertis, ta bonne humeur ainsi que pour tout ce que tu m'as appris. Tu as toujours été là, disponible, à la fois dans les bons moments et ceux un peu plus compliqués. Ce fut un réel plaisir de passer toutes ces années avec toi.

J'aimerais aussi remercier Pascal Maire, Michèle Crozatier et Yad Ghavi-Helm d'avoir accepté de faire partie de ce jury de thèse et d'évaluer mon travail.

Un merci aussi à Serge Plaza, Denis Thieffry et Phillipe Arnaud d'avoir pris part à mes des comités de thèse. Merci pour vos conseils et vos critiques positives.

Je tiens à remercier chaleureusement les membres du GReD, tout le service commun et en particulier Françoise, Nadège, Eric et Christian.

Je souhaite aussi dire merci aux personnes des autres équipes pour les bons moments partagés avec eux: Catherine, Isabelle, Franck, Anne, Stéphanie MM, Stéphanie Hermitte, Sabine, Marie-Elisa, Jill, Elisa. Un merci au jeune padawan Nico qui a été un soutien technique sans faille pendant quelques années et avec qui j'ai bien rigolé. Merci aussi à Bertille ou l'as des paniers à la Stephen Curry, arrivée en même temps que moi au Gred et avec qui je partage de très bons souvenirs.

Merci à tous les membres de l'équipe JAGLA, avec qui j'ai travaillé pendant ces années : Teresa, Ouarda, Monika (merci pour tes conseils), Michaela dit wild wild west, Cédric, Lucie et Thomas, tu m'as vendu du rêve.

Je voudrais remercier plus particulièrement Jean Philippe qui a été d'une aide précieuse dès le départ. Merci aussi de m'avoir fait voir les endives sous un autre angle et pour toute ta bonne humeur.

Merci à Yoan dit père Yoyo, notre bio-informaticien passionné de camping, qui s'est cassé la tête plus d'une fois sur mes données mais qui a géré du début à la fin, tu as assuré !!

J'aimerais aussi remercier les thésards de l'équipe pour la bonne ambiance général et pour l'ensemble des interactions que l'on a eu ensemble: Lilia dit MaK, la spécialiste de la vanne malgré elle, pour ta gentillesse au quotidien, Quentin ou comment purifier tellement un plasmide qu'il en disparaît, pour ta jeunesse et tes amorties au bad; Merci à Guigui, Roi des phoques, pour ton intérêt sans limites pour la science, notre side project, nos chambrages et nos discussions parfois constructives (bon j'avoue souvent). Merci à mon italienne préférée Cristiana « fan de Ed Sheeran et oui ca existe !! » pour ta joie de vivre, pour les moments partagés en salle de flip et au labo, ton soutien et l'aide que tu m'as apporté quand j'en avais besoin.



Je souhaite remercier une ancienne thésarde et surtout amie d'enfance, je parle bien sûr de toi Emilie. Le Dieu de la drosophile nous a réuni au sein de cette équipe pour mon plus grand plaisir. J'ai apprécié tous les moments partagés avec toi pendant ces années, c'était vraiment cool.

Mes derniers remerciements et non des moindres vont à ma famille et surtout à ma femme Mélanie et mon fils Jules qui m'ont soutenu toutes ces années. Vous avez toujours été la pour m'encourager même si je sais que l'investissement que j'ai mis dans ma thèse a parfois été lourd à gérer pour vous. Je ne trouve pas de mots assez forts pour vous dire à quel point votre soutien a été important pour moi. Pour cela je vous remercie de tout mon cœur.



Les muscles squelettiques présentent des propriétés individuelles de taille, forme, orientation, site d'attachement, nombre de noyaux déterminées par l'expression d'une combinaison de facteurs d'identité. Les processus par lesquels ces gènes déterminent les caractéristiques finales du muscle ne sont pas pleinement compris. Découvrir comment ces régulateurs de l'identité sont connectés entre eux, à travers un réseau de gènes spécifiant le destin cellulaire de chaque type de muscle, reste un enjeu majeur.

Pour identifier de nouveaux acteurs qui déterminent les caractéristiques individuelles du muscle au cours de l'embryogenèse chez la drosophile, nous avons optimisé l'approche cellule spécifique TRAP (Translating Ribosome Affinity Purification) permettant l'isolation des ARNs messagers en cours de traduction. Cette méthode a été utilisée sur deux populations restreintes de muscles (Lms- et Slou-positives) et sur la population musculaire globale (Duf-positive) à trois fenêtres de temps ciblant des processus clés de la myogenèse tels que la spécification de cellules fondatrices, la fusion des myoblastes, l'attachement aux cellules de tendon et l'innervation par les motoneurones.

Pour avoir une vue d'ensemble de l'expression des gènes dans ces différentes conditions des analyses transcriptomiques à l'aide de puces ADNc ont été réalisées. Une étude spatiale nous a permis d'identifier de nouveaux acteurs potentiels du code d'identité nécessaires à la diversification de la population Slou ainsi que l'implication de la protéine de liaison à l'ARN, Boule, dans l'organisation des filaments d'actine au sein des fibres musculaires. Dans un second temps une analyse temporelle des données a révélé un ensemble de gènes codant pour des protéines de liaison à l'actine avec un enrichissement spécifique dans la population Lms positive. L'étude plus approfondie des gènes *gel* et *dCryAB* a permis de déterminer leur rôle au cours de la myogenèse. Ces deux protéines participent de manière importante à la régulation de la croissance des muscles Lms positifs en régulant le nombre d'évènement de fusion, l'élongation et l'attachement des myotubes.

A travers ce projet nous avons identifié de nouveaux acteurs potentiels du code d'identité, impliqués dans l'acquisition des propriétés distinctes des sous populations musculaires Slou et Lms au cours de l'embryogenèse en participant soit au contrôle de l'expression des gènes soit à l'acquisition des caractéristiques morphologiques finales. Nous espérons que les données générées permettront dans le futur d'acquérir de nouvelles connaissances sur les fonctions des orthologues de ces gènes au cours de la myogenèse des vertébrés et que cela aura un impact sur la compréhension de certaines myopathies.





Skeletal muscles display specific features (size, shape, orientation, attachment site, number of nuclei), which are tightly controlled by the combinatorial expression of identity factors. Currently, process by which identity genes determine muscle characteristics is not fully understood. Uncovering how identity factors are interconnected and how they control network of genes specifying muscle cell fate remains a major challenge.

To discover new actors of the identity code allowing the acquisition of individual muscle characteristics during *Drosophila* embryogenesis we optimised a genome wide cell specific approach called TRAP (Translating Ribosome Affinity Purification), which allows the isolation of mRNA engaged in translation. This method has been used on two restricted populations of muscle cells (Slou and Lms positive) and on the global musculature (Duf-positive) at three time windows covering key muscle developmental steps such as founder cell specification, myoblast fusion, attachment to the epidermis via tendon cells and the innervation by motoneurons.

To assess gene expression at a global level in these muscle subsets, we performed transcriptomic analyses at targeted developmental windows by using microarray. First, a comparison of spatial gene expression allowed us to identify new potential actors of the identity code for Slou-positive muscles and the involvement of the RNA binding protein Bol in the structural organisation of muscle fibers. We also performed analyses of temporal transition profiles of genes differentially expressed in Slou- and Lms-positive muscles and identified cluster of genes enriched preferentially in the Lms population and encoding actin-interacting proteins. We focused on two genes from this cluster, *gel* and *dCryAB* and determined their functions during myogenesis. These two proteins control a coordinated growth of the Lms-positive muscles by fine-tuning of both the number of fusion events and the elongation of myotubes to tendon cells.

With this project, we have identified news potential actors of the identity code that govern the acquisition of individual properties of Slou and Lms muscle subsets during *Drosophila* embryogenesis. We hope that the newly generated data will enable to gain further knowledge on the corresponding orthologues during vertebrate myogenesis and will help to understand why particular muscles are affected in some myopathies.



Synthèse bibliographique.....	1
<b>A - Généralités sur le muscle .....</b>	<b>2</b>
<b>I- Les différents types de muscles .....</b>	<b>2</b>
<b>II- Le muscle squelettique.....</b>	<b>3</b>
1. Rôle .....	3
2. Structure et composition .....	3
<b>B - Myogenèse embryonnaire: Origine des muscles squelettiques et régulation transcriptionnelle .....</b>	<b>4</b>
<b>I-Somitogenèse .....</b>	<b>4</b>
<b>II-Formation des structures dérivées des somites et du mésoderme crânial .....</b>	<b>5</b>
<b>III- Rôle des morphogènes au cours de la myogenèse embryonnaire .....</b>	<b>7</b>
1. La voie Wnt .....	7
2. La voie Sonic Hedgehog (shh) .....	8
3. La voie FGF .....	8
4. La voie BMP.....	9
5. La voie Notch .....	9
6. Morphogènes et muscles craniofaciaux .....	10
<b>IV- Contrôle génétique de la diversité musculaire.....</b>	<b>10</b>
1. Les facteurs clés de la spécification des muscles du tronc et des membres.....	10
1.1. Les MRFs (Myogenic Regulatory Factors) .....	10
1.2. Les Pax.....	12
1.3. Les Six.....	14
1.4. Les facteurs impliqués dans la formation des muscles des membres .....	15
1.5. Les facteurs répresseurs des MRFs .....	16
2. Les facteurs clés de la spécification des muscles craniofaciaux.....	16
2.1 Pitx2 .....	16
2.2 Tbx1.....	17
2.3 Islet1.....	18
2.4 Les facteurs répresseurs de la myogenèse crâniofaciale. ....	18
<b>V- La différenciation terminale des muscles squelettiques .....</b>	<b>18</b>
1. Les facteurs : Mef2, Myog et Nfix .....	18
2. Fusion des myoblastes : Fibres primaires et secondaires .....	20
3. Attachement et jonction myotendineuse .....	22
4. Innervation et jonction neuromusculaire .....	23
<b>C – La drosophile comme modèle d'étude de la diversification musculaire.....</b>	<b>25</b>
<b>I. La drosophile.....</b>	<b>25</b>
<b>II. Généralités sur le muscle somatique .....</b>	<b>25</b>
<b>III. Le mésoderme.....</b>	<b>26</b>
1. Origine et détermination du mésoderme .....	26

2. Subdivision du mésoderme.....	26
<b>IV. Etapes de formation du muscle .....</b>	<b>28</b>
1. Origine des cellules fondatrices et des FCMs.....	28
2. Spécification des cellules fondatrices : division asymétrique .....	29
3. La fusion des myoblastes et le positionnement des noyaux.....	31
3.1 Reconnaissance et adhésion FC/FCM .....	32
3.2 Transduction du signal de fusion .....	32
3.3 Remodelage du cytosquelette d'actine et formation de pores de fusion .....	33
3.4 Positionnement des noyaux au sein du muscle .....	34
3.5 Régulation muscle spécifique de la fusion .....	36
4. Elongation et attachement aux cellules tendineuses .....	36
4.1 Détermination des cellules de tendon .....	37
4.2 Guidance et attachement .....	37
4.3 Jonction myotendineuse.....	39
5. Innervation.....	40
5.1 Formation des motoneurones .....	40
5.2 Migration et reconnaissance muscle-motoneurone.....	40
5.3 Jonction neuromusculaire.....	42
6. Sarcomères .....	43
<b>V. identité et diversification musculaire .....</b>	<b>44</b>
1. Les facteurs de régulation transcriptionnelle généraux .....	44
1.1. Twist.....	44
1.2 Mef2.....	46
2. Les facteurs d'identités .....	47
2.1 Le facteur dorsal Tailup/islet1 .....	47
2.2 Les facteurs latéraux.....	48
2.2.1 Lms .....	48
2.2.2 Araucan/Caupolican.....	50
2.3 Le facteur Org-1 .....	51
2.4. Le facteur Slouch /S59 .....	52
2.5 Le facteur dSix4.....	53
<b>Conclusion .....</b>	<b>55</b>
<b>Objectifs de travail .....</b>	<b>58</b>
<b>Résultats.....</b>	<b>60</b>
<b>A – Mise en place et validation de la méthode TRAP .....</b>	<b>61</b>
<b>I. La méthode TRAP .....</b>	<b>61</b>
1.1 Création des lignées TRAP.....	62

1.2 Optimisation de la technique.....	63
1.3 Qualité, quantité et spécificité des ARNs isolés.....	64
<b>II. Validation des résultats obtenus par puce ADNc .....</b>	<b>65</b>
<b>III. Publication 1 « TRAP-rc, Translating Ribosome Affinity Purification from Rare Cell Populations of Drosophila Embryos » .....</b>	<b>66</b>
<b>B – Identification de nouveaux gènes impliqués dans l’identité musculaire par la méthode TRAP .....</b>	<b>67</b>
<b>I. Publication 2 « Translational profiling of muscle subsets in <i>Drosophila</i> identifies muscle type specific growth tuners controlling multiple muscle properties » .....</b>	<b>67</b>
<b>II. Résultats complémentaires .....</b>	<b>69</b>
2.1 Identification d’autres acteurs de la diversification musculaire .....	69
2.2 Profil d’expression et régulation des gènes <i>Kah</i> , <i>CG18136</i> et <i>Boule</i> . .....	70
2.3 Analyses fonctionnelles des gènes candidats par ARN interférent.....	71
2.4 Caractérisation fonctionnelle du gène <i>bol</i> .....	71
2.5 Analyses fonctionnelles du gène <i>bol</i> en larve.....	72
<u>Discussion.....</u>	<u>75</u>
<b>A- La méthode TRAP : une technique efficace pour l’identification du translatome muscle spécifique .....</b>	<b>76</b>
<b>B- Identification de nouveaux acteurs potentiels du code d’identité et de l’acquisition des caractéristiques finales des sous populations musculaires Lms et Slou.....</b>	<b>79</b>
<b>I. De nouveaux acteurs potentiels du code d’identité des muscles Slou positifs.....</b>	<b>79</b>
<b>II. De nouveaux gènes réalisateurs du code d’identité des muscles Lms positifs, <i>Gel</i> et <i>dCryAB</i>.....</b>	<b>81</b>
<b>C- D’autres facteurs potentiels de la diversification musculaire.....</b>	<b>86</b>
<b>I. Le facteur de transcription <i>Kah</i> et l’histone déacétylase <i>CG18136</i> .....</b>	<b>86</b>
<b>II. <i>Bol</i>, une nouvelle protéine de liaison à l’ARN nécessaire au cours de la myogenèse .</b>	<b>88</b>
<u>Conclusion .....</u>	<u>92</u>
<u>Références bibliographiques.....</u>	<u>95</u>

## **FIGURES**

### **Synthèse Bibliographique**

Figure 1: Les différents types de muscles qui composent le corps humain

Figure 2: Structure et composition du muscle strié squelettique

Figure 3: Formation des somites selon le modèle de clock and wavefront

Figure 4: Structures dérivées des somites et du mésoderme crânial

Figure 5: Rôle des morphogènes dans la formation des muscles embryonnaires

Figure 6: Profil expression des MRFs au cours du développement embryonnaire chez la souris

Figure 7: Cascade myogénique à l'origine des muscles du tronc et des membres

Figure 8: Cascade myogénique à l'origine des muscles craniofaciaux

Figure 9: Fusion des myoblastes

Figure 10: Formation de la jonction myotendineuse

Figure 11: Innervation du muscle squelettique par les motoneurones

Figure 12: Développement des muscles somatiques chez la Drosophile

Figure 13: Le patron musculaire de l'embryon et de la larve

Figure 14: Détermination et migration du mésoderme

Figure 15: Subdivision du mésoderme

Figure 16: Origine et détermination des myoblastes

Figure 17: Spécification des cellules fondatrices

Figure 18: La fusion des myoblastes

Figure 19: Attachement des myotubes aux tendons

Figure 20: La jonction neuromusculaire

Figure 21: Structure des sarcomères

Figure 22: Les facteurs d'identité

Figure 23: Spécification de certains muscles dorsaux par Tailup

Figure 24: Spécification et sécurisation du programme d'identité des muscles latéraux par Lms

Figure 25: Régulation de l'identité des muscles LT3-4 par les facteurs iroquois Araucan/Caupolican

Figure 26: Org-1 est requis pour la spécification d'une sous population de muscles et pour la régulation de deux iTFs

Figure 27: L'identité musculaire fournie par le facteur Slou à l'origine d'une population hétérogène de muscles

### **Résultats**

Figure 28: Représentation schématique du TRAP

Figure 29: Patron musculaire des trois populations ciblées

Figure 30: Validation de la qualité, la quantité et la spécificité des ARN isolés après TRAP dans les trois populations au stade 13-15 (T2)

Figure 31: Validation de la qualité des échantillons de TRAP

Figure 32: Ontologies des gènes enrichis dans les muscles Slou aux trois stades de développement.

Figure 33: Comportement des gènes sélectionnés dans la population Slou et Lms au cours du temps

Figure 34: Profil d'expression des gènes candidats *kah*, *CG18136* et *bol*.

Figure 35: Régulation potentielle des gènes *kah*, *CG18136* et *bol* par le facteur myogénique Mef2

Figure 36: La perte de fonction de *bol* entraîne des défauts musculaires en embryon

Figure 37: La perte de *bol* induit une désorganisation des fibres musculaires chez la larve

Figure 38: Expression de *bol* chez les vertébrés

## Discussion

Figure 39: La méthode INTACT permet d'isoler les noyaux de la population Lms

Figure 40: Génération de mutant à l'aide du système CRISPR-Cas9.

## TABLEAUX

### Résultats

Tableau 1: Récapitulatif de l'enrichissement des gènes connus pour être exprimés et non exprimés dans la population Slou aux trois stades de développement

Tableau 2: Gènes codant pour des facteurs de régulation retrouvés comme enrichis dans les populations Lms et Slou





## Liste des abréviations

ADN	Acide désoxyribonucléique	LT	Latero-transverse
ADNc	ADN complémentaire	MAPK	Mitogen Activated Protein Kinase
AEL	After Egg Laying	Mef2	Myogenic enhancer factor 2
AMP	Adult Muscle Precursor	MRF	Myogenic Regulatory Factor
Ap	Apterous	MTJ	Myotendinous junction
ARN	Acide ribonucléique	MyHC	Myosin Heavy Chain
RNAi	ARN interférent	NMJ	Neuromuscular junction
bHLH	basic Helix-Loop-Helix	PC	Progenitor Cell
BMP	Bone Morphogenetic Protein	Ko	Knock out
Bol	Boule	RIP	RNA immunopurification
ChiP	Chromatin immunopurification	RRM	RNA recognition motif
CRM	Cis-regulatory module	RT-qPCR	Reverse transcriptase quantitative polymerase chain reaction
Dpp	Decapentaplegic	SBM	Segment border muscle
Duf	Dumbfounded	Shh	Sonic hedgehog
EGF	Epidermal Growth Factor	Slou	Slouch
ECM	Extracellular matrix	TF	Transcription factor
EOM	Extraocular muscles	Twist	Twist
GFP	enhance Green fluorescent protein	TRAP	Translating Ribosome Affinity Purification
GO	Gene Ontology	UAS	Upstream activating sequence
FC	Founder Cell	Wg	Wingless
FCM	Fusion Competent Myoblast	Wnt	Wint
FGF	Fibroblast Growth Factor	WT	Wild Type (sauvage)
Gel	Gelsolin		
iTF	identity Transcription Factor		
Hh	Hedgehog		
INTACT	Isolation of nuclei tagged in specific cell types		
Kah	Kahuli		
Lms	Lateral muscles scarcer		
LRR	Leucine-Rich Repeat		



?

?

?

?

?

?

?

?

?

?

?

?

?

?

?

?

?

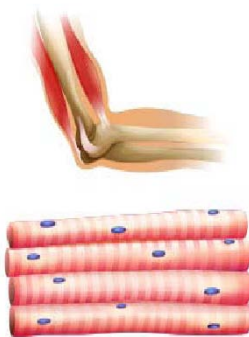
?

?

?

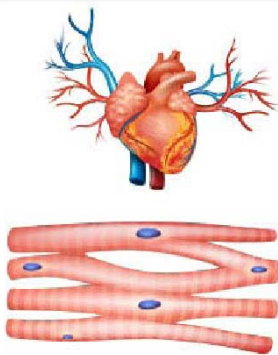
# 2nuAs Gr t | A i e

22



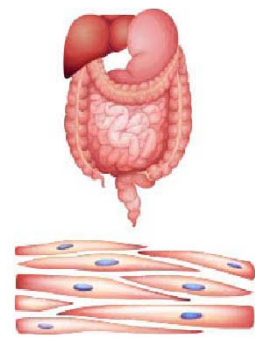
QUESTION BANK FOR THE EXAMINATION OF THE B.Sc. DEGREE COURSE IN PHYSIOLOGY  
QUESTION BANK FOR THE EXAMINATION OF THE B.Sc. DEGREE COURSE IN PHYSIOLOGY  
QUESTION BANK FOR THE EXAMINATION OF THE B.Sc. DEGREE COURSE IN PHYSIOLOGY

22



QUESTION BANK FOR THE EXAMINATION OF THE B.Sc. DEGREE COURSE IN PHYSIOLOGY  
QUESTION BANK FOR THE EXAMINATION OF THE B.Sc. DEGREE COURSE IN PHYSIOLOGY  
QUESTION BANK FOR THE EXAMINATION OF THE B.Sc. DEGREE COURSE IN PHYSIOLOGY

22



QUESTION BANK FOR THE EXAMINATION OF THE B.Sc. DEGREE COURSE IN PHYSIOLOGY  
QUESTION BANK FOR THE EXAMINATION OF THE B.Sc. DEGREE COURSE IN PHYSIOLOGY  
QUESTION BANK FOR THE EXAMINATION OF THE B.Sc. DEGREE COURSE IN PHYSIOLOGY

## A - Généralités sur le muscle

### I- Les différents types de muscles

Le corps humain est constitué d'environ 650 à 700 muscles qui remplissent des fonctions précises. Notamment, ils sont impliqués dans le mouvement à travers l'espace comme la marche qui mobilise l'organisme entier. D'autres mouvements plus fins nécessitent l'action mécanique de muscles plus spécialisés comme lorsque l'on met en mouvement une partie de l'organisme comme le bras ou bien encore lors du passage de la nourriture dans l'estomac. Le mouvement se produit donc à travers une grande variété de contextes et permet l'accomplissement d'une multitude de tâches. Chacun de ces mouvements a pour origine le muscle qui existe sous différents types, strié ou lisse, selon son aspect histologique. A l'intérieur de ces deux catégories on retrouve trois grands types de muscles qui se différencient par leurs caractéristiques structurelles et fonctionnelles.

Les muscles squelettiques (qui sont attachés au squelette) permettent d'assurer les mouvements et le maintien du corps. Ils ont un aspect strié du fait de leur composition en sarcomères, unités contractiles du muscle. De plus ils présentent une structure multi-fibrillaire ou chaque fibre est une cellule syncytiale multi-nucléée. L'activité du muscle squelettique est contrôlée par le système nerveux central (Figure 1A).

Le muscle cardiaque forme la partie contractile du cœur, il a aussi un aspect strié mais de nombreuses différences structurelles existent avec le muscle squelettique. Les cellules qui le composent sont mononucléées, ramifiées et contiennent des régions spécialisées appelées disques intercalés permettant un transport régulé des molécules et des signaux électriques d'une cellule à une autre (Figure 1B).

Enfin, le muscle lisse n'a pas d'aspect strié comme le muscle cardiaque ou squelettique car les filaments d'actine et de myosine qui le composent ne sont pas arrangés de la même manière au sein de la cellule. De plus, les cellules du muscle lisse sont fusiformes, leur conférant une résistance importante aux fortes tensions. Le muscle lisse est trouvé principalement dans les organes creux tels que les vaisseaux sanguins, les organes du tractus digestif ou encore l'utérus et la vessie. Il est contrôlé par le système nerveux autonome et contrôle les activités involontaires du corps (fonction végétative) comme la constriction des artères ou les mouvements des organes digestifs (Figure 1C).



Dans un souci d'efficacité et de clarté du manuscrit je développerai plus en détail seulement la partie associée au muscle squelettique, sujet de cette étude.

## **II- Le muscle squelettique**

### *1. Rôle*

Le muscle squelettique est le plus dynamique et plastique des tissus composant le corps humain. Chez l'homme le muscle squelettique représente approximativement 40 % du poids total du corps. Le muscle squelettique contribue de manière significative à une multitude de fonctions. La fonction principale est de convertir l'énergie mécanique pour générer de la force et du mouvement. Le muscle squelettique contribue également à la gestion d'énergie métabolique, notamment en stockant des acides aminés (AA) et en permettant la production de chaleur pour maintenir la température corporelle. Plus particulièrement, le muscle est un réservoir d'AA nécessaires aux autres tissus tels que la peau, le cerveau et le cœur pour la synthèse de protéines spécifiques (Wolfe, 2006).

### *2. Structure et composition*

L'architecture du muscle squelettique est caractérisée par un arrangement très précis de fibres musculaires aussi appelées myofibrilles associées à du tissu connectif (Frontera and Ochala, 2015).

Un muscle individuel est entouré d'une couche de tissu connectif appelé épimysium. A l'intérieur du muscle un groupe de fibres est arrangé en faisceau et est recouvert de tissu connectif, le péri-mysium. Chaque fibre individuelle est entourée d'une membrane plasmique, le sarcolemme. On retrouve associé au sarcolemme un complexe de plusieurs protéines qui sont physiquement connectées à la structure interne des myofilaments. Une fibre musculaire est composée à 80% de protéines (contractiles, régulatrices ou du cytosquelette) et de sarcoplasme (cytoplasme des cellules musculaires) (Hoppeler et al., 1973). Un autre type de cellules situées entre le sarcolemme et la lame basale constitue les cellules satellites qui sont les cellules souches adultes du muscle squelettique. Elles contribuent à la croissance musculaire, la réparation et la régénération du muscle (Relaix et al., 2005; Hikida et al, 2011).

La taille du muscle est déterminée par le nombre et la taille de chaque fibre individuelle qui le compose. Les fibres musculaires sont multinucléées et post-mitotiques. Pour la plupart, chaque noyau à l'intérieur d'une myofibrille contrôle le type de protéines



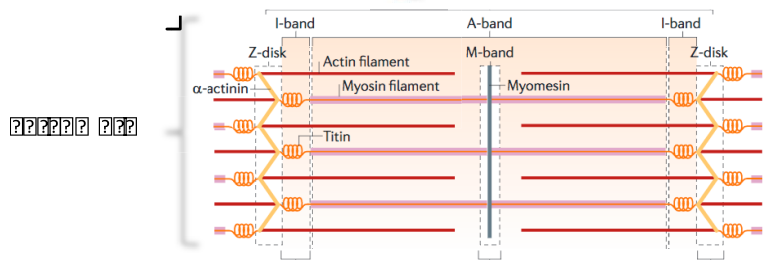
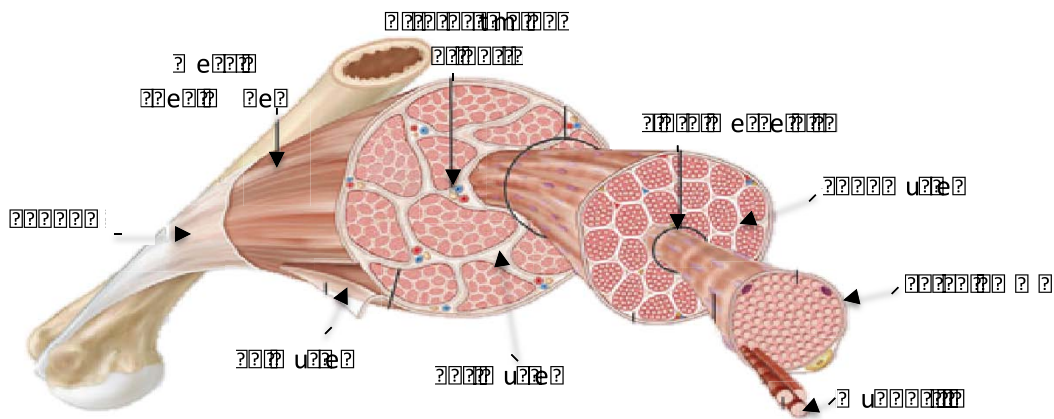


Diagram illustrating the structure and components of a sarcomere, showing the arrangement of actin filaments, myosin filaments, and titin proteins.

The diagram illustrates the structure and components of a sarcomere, showing the arrangement of actin filaments, myosin filaments, and titin proteins. The sarcomere is the basic contractile unit of muscle, and its structure is shown in the diagram. The actin filaments are attached to the Z-disks, and the myosin filaments are attached to the actin filaments. The titin proteins connect the actin filaments to the Z-disks. The diagram also shows the arrangement of the filaments in the I-band, A-band, and M-band regions.

synthétisées dans cette partie spécifique de la cellule. Ces régions sont connues comme étant des domaines nucléaires et ont une taille variable hautement régulée (Hikida et al., 2011). Suite à de nombreuses études protéomiques sur les cellules musculaires on estime que chaque fibre est composée de milliers de myofibrilles et contient des milliards de myofilaments (Gelfi et al., 2011). Lorsque les myofilaments sont assemblés et ordonnés en un patron très précis cela forme le sarcomère, l'unité contractile basique du muscle squelettique (Figure 2).

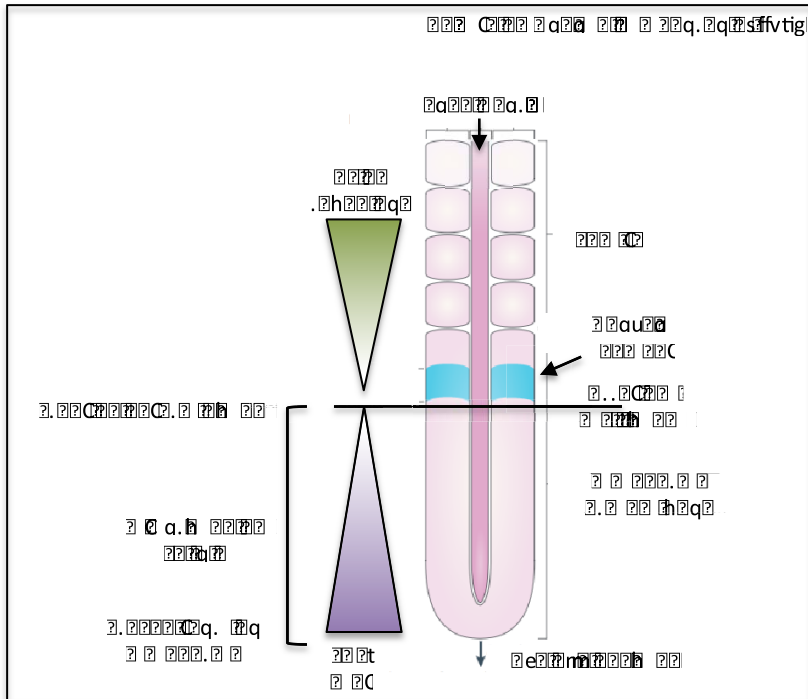
Les sarcomères sont composés de filaments fins d'actine et de filaments épais de myosine qui sont arrangés de manière hautement ordonnée. Ils contiennent plus d'une centaine de protéines régulatrices telles que le complexe troponine-tropomyosine, des protéines de liaison au cytosquelette comme l' $\alpha$ -actinine, la myomesine ou encore la titine responsable de l'élasticité de ces structures (Lange et al., 2006).

Les cellules musculaires à l'origine des différents muscles du corps acquièrent leurs caractéristiques fonctionnelles individuelles au cours de la myogenèse embryonnaire.

## **B - Myogenèse embryonnaire: Origine des muscles squelettiques et régulation transcriptionnelle**

### **I-Somitogenèse**

L'identité des cellules qui vont former les trois couches germinales, dont le mésoderme, est déterminée précocement au cours de la gestation (Arnold et Robertson, 2009). Le mésoderme est sous-divisé en trois parties, le mésoderme paraxial, intermédiaire et latéral en fonction de sa position par rapport à la ligne médiane du corps. Au cours du développement l'expression oscillatoire de gènes ainsi que des gradients de morphogènes induit la condensation du mésoderme paraxial en somite (Aulehla and Pourquié, 2006). Les muscles squelettiques trouvent leur origine au niveau des somites qui sont des structures épithéliales transitoires (Christ and Ordhal, 1995)(Huang et al., 1994). La formation des somites est un processus qui se produit de part et d'autre de la notochorde et du tube neural suivant un axe antéro-postérieur ou rostro-caudal de l'embryon. Dans un intervalle régulier une paire de somites est formée à partir du mésoderme pré-somitique (MPS) non segmenté situé dans la partie caudale. Ce modèle d'intervalle périodique appelé « clock and wavefront » propose une régulation spatio-temporelle de la segmentation de l'embryon



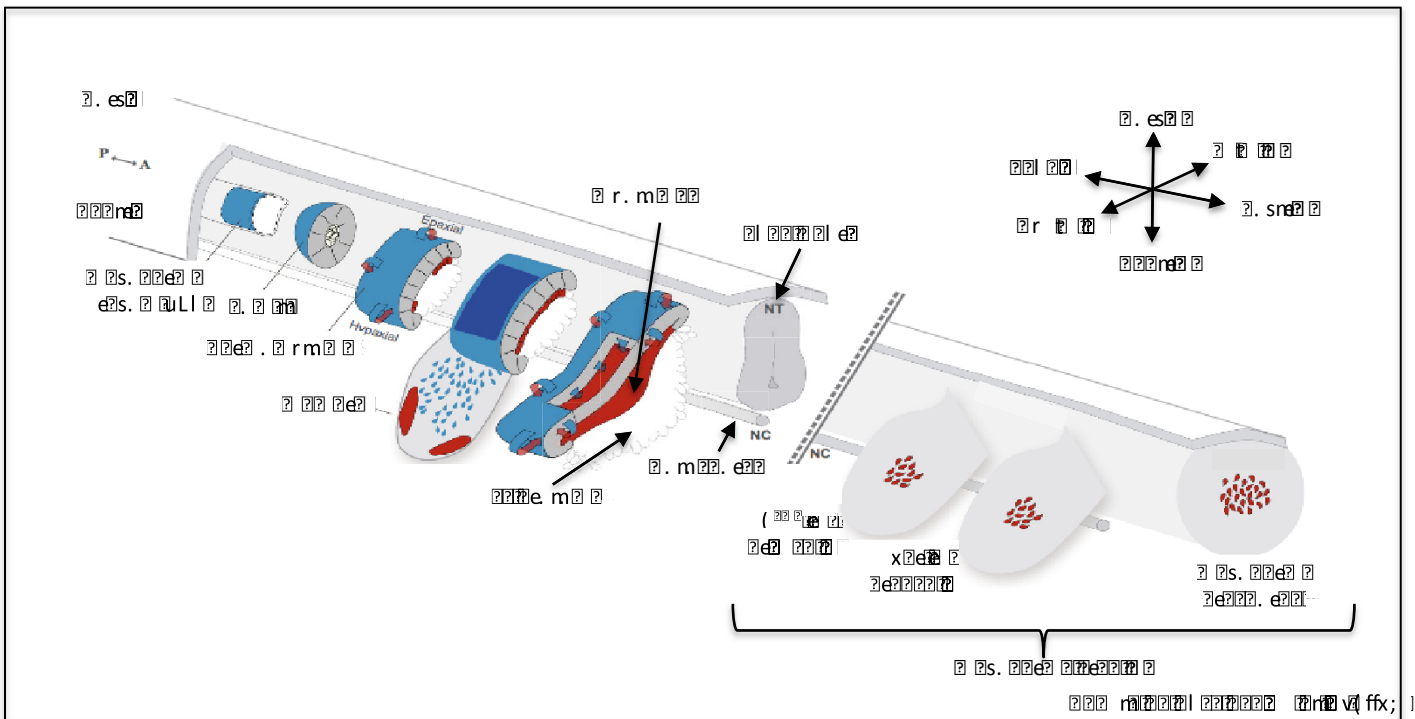
The diagram illustrates a sedimentation apparatus. On the left, a green funnel-shaped container is positioned above a purple inverted funnel-shaped container. A central vertical tube is inserted into the purple container, containing a blue liquid. A horizontal line is drawn across the tube. To the right of the tube, there is a vertical column of rectangular blocks. An arrow points from the blue liquid in the tube to the text "a. b. c." on the right side. Below the purple container, there is a vertical arrow pointing upwards. At the bottom right of the tube, there is a vertical arrow pointing downwards.

(Cooke and Zeeman, 1976). Il est basé sur l'existence d'une horloge (clock) caractérisée par l'expression oscillatoire de gènes tels que *lunatic fringe (Lnfg)*, *her1*, *Hes1* régulés par la voie Notch (Pourquié and Tam, 2001) et *axin2* (Aulehla et al., 2003) par la voie Wnt/B-catenine dont l'intensité diminue en se rapprochant de la partie rostrale de l'embryon (tête). Ces oscillations sont stoppées dans la zone appelée « front de détermination » où une concentration décroissante en morphogènes (Fgf8, Wnt3a) va amener les cellules à changer de propriétés afin de donner naissance à une future paire de somites (Aulehla et al., 2003; Dubrulle et al., 2001). L'acide rétinoïque a aussi un rôle important dans la segmentation du mésoderme car celui-ci est exprimé selon un gradient inverse à celui de Fgf8 et joue donc un rôle d'antagoniste (del Corral et al., 2003) (Figure 3).

Ces différents processus d'ondes et d'oscillations permettent de limiter la zone de segmentation du mésoderme ainsi que de définir l'emplacement, l'orientation et la taille des somites au sein du mésoderme pré-somitique. Ces deux mécanismes sont reliés entre eux car les oscillations normalement induites par les voies Notch-Wnt sont absentes lorsque le récepteur à Fgf8, Fgfr1 est muté chez l'embryon de souris (Wahl et al., 2007).

## II-Formation des structures dérivées des somites et du mésoderme crânial

Les somites sont formés le long de l'axe rostro-caudal de l'embryon et sont compartimentés selon l'axe dorso-ventral grâce à des signaux positifs et négatifs provenant des tissus environnements, l'ectoderme dorsal, le tube neural et la notochorde. Ces derniers induisent l'expression de gènes de la famille des Wnts, FGFs, Shh et BMPs. La partie ventrale du somite forme le sclérotome à l'origine du squelette axial, des côtes et du syndétome adjacent, source de tendons des muscles du tronc (Brent et al., 2003). Il est constitué de cellules mésenchymateuses et est identifiable par la détection de certains marqueurs tels que Pax1-9 ainsi que Scleraxis (Scx) (Cossu and Borello, 1999). La partie la plus dorsale du somite, le dermomyotome, reste une structure épithéliale plus longtemps et est à l'origine du derme, de tous les muscles squelettiques du tronc, des membres mais aussi à l'origine des cellules musculaires des vaisseaux sanguins et du corps gras brun (Ben-Yair and Kalcheim, 2008; Seale et al., 2008). Le dermomyotome est aussi compartimenté de façon médio-latérale. Les cellules du dermomyotome sont marquées par l'expression des



The diagram illustrates the development of the eye from the optic vesicle to the optic cup. The optic vesicle is formed by the folding of the neural tube. The epaxial and hypaxial folds are shown, along with the retinal margin (r. m) and margin ectoderm (m. e). The optic cup is formed by the folding of the optic vesicle, and the neural crest (NC) and axons (x) are shown. The diagram is divided into several stages, showing the progression from the flat optic vesicle to the folded optic cup. The labels 'E. es', 'P', 'A', 'Epaxial', 'Hypaxial', 'r. m', 'm. e', 'NT', 'NC', and 'x' are used to identify the various structures and tissues involved in the process.

facteurs de transcription Pax3 et Pax7 (Kiefer and Hauschka, 2001). Les cellules du bord du dermomyotome subissent tout d'abord une transition épithélio-mésenchymateuse pour former le myotome, qui contient les premières fibres différenciées. Ensuite la partie la plus proche du tube neural, le domaine épaxial (dorso-médiale) du dermomyotome donnera les muscles profonds du dos et la partie hypaxiale (ventro-latérale) la plus éloignée du tube neural une partie des muscles du tronc, ceux des membres, du diaphragme et de la langue (Parker et al., 2003). Les cellules provenant de la partie ventro-latérale (hypaxial) du dermomyotome subissent une transition épithélio-mésenchymateuse et migrent vers les territoires des membres (bourgeons) (Figure 4). Les molécules clés impliquées dans cette transition sont nombreuses (Vasyutina and Birchmeier, 2006; Yokoyama and Asahara, 2011), je reviendrai sur cet aspect un peu plus tard dans le manuscrit. Pour finir, la partie centrale du dermomyotome perd sa structure épithéliale et les cellules myogéniques progénitrices entrent dans le myotome sous-adjacent. Ces cellules se différencient ou prolifèrent, procurant ainsi un groupe de cellules appelées cellules satellites (Pax3/Pax7 positives) nécessaires pour la croissance musculaire, la myogenèse post-natale et la régénération (Gros et al., 2004; Kassar-Duchossoy, 2005; Relaix et al., 2005)

Comme décrit précédemment, la plupart des progéniteurs myogéniques embryonnaires sont formés à partir des somites. Cependant cela n'est pas le cas pour une partie des muscles craniofaciaux. Hormis les muscles du larynx et de la langue qui dérivent des somites les autres muscles de la tête dérivent du mésoderme crânial (muscles branchiomériques) et du mésoderme crânial préchordal (muscles extraoculaire) (Figure 4). A eux deux ils donnent naissance à approximativement 60 muscles dans la tête qui sont utilisés pour faciliter la prise de nourriture, le mouvement des yeux ou encore pour fournir les expressions de la face. Les précurseurs musculaires migrent en partie vers les trois arcs branchiaux alors que les précurseurs des muscles oculaires migrent vers le premier arc branchial et le territoire oculaire (Noden and Francis-West, 2006; Noden and Trainor, 2005)

La myogenèse embryonnaire se fait à travers plusieurs étapes successives finement régulées par des acteurs à la fois extrinsèques, les morphogènes (Wnts, Shh, BMP, Notch), et intrinsèques via l'expression de différents facteurs myogéniques (Pax, MRFs, Six/Eya, Tbx, Pitx) qui opèrent de manière orchestrée et selon une cascade de régulation bien définie. Tous ces facteurs permettent d'obtenir une grande diversité de muscles. Tout d'abord, je



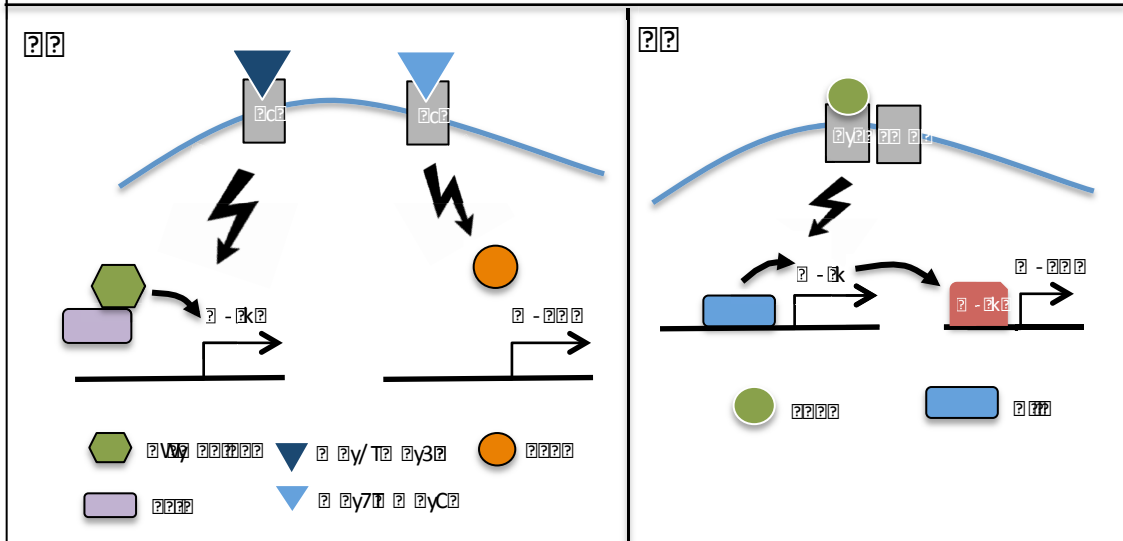
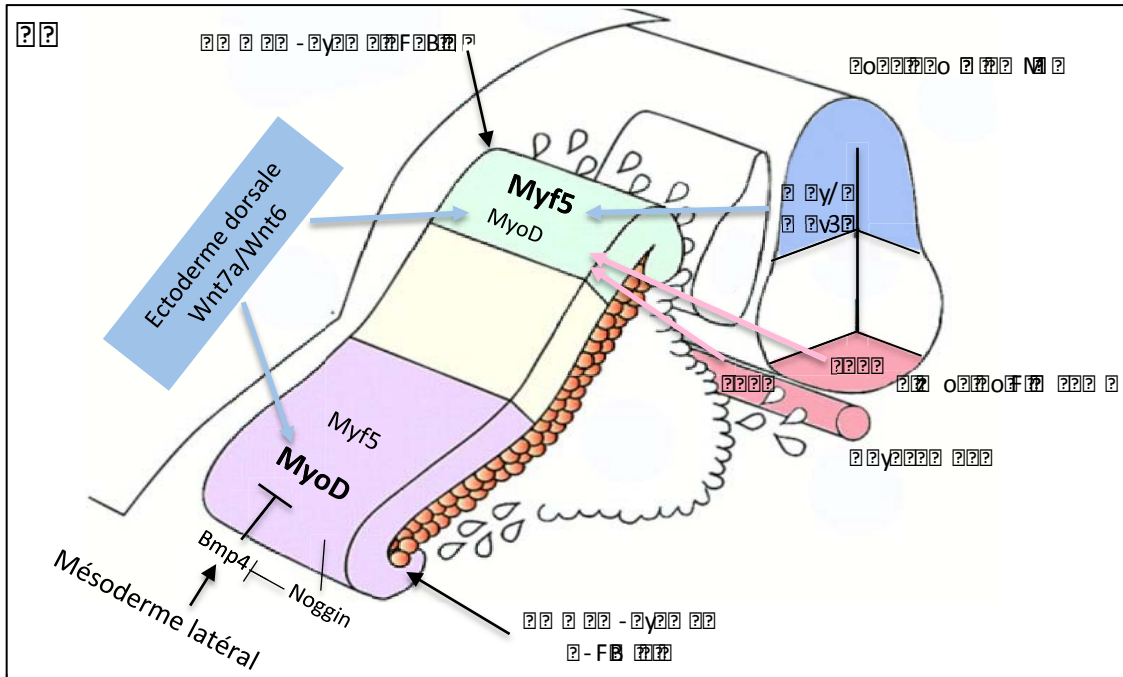
développerai les morphogènes pour me consacrer par la suite au contrôle de la diversité des muscles embryonnaires à travers la régulation transcriptionnelle contrôlée par les facteurs de transcription.

### III- Rôle des morphogènes au cours de la myogenèse embryonnaire

#### 1. La voie Wnt

La famille de protéines Wnt est requise au cours de la myogenèse pour la formation du dermomyotome et du myotome. Les protéines Wnt se lient à leurs récepteurs cellulaires Frizzled (Fzd) ce qui induit l'activation de la voie canonique via le complexe transcriptionnel  $\beta$ -caténine/TCF mais aussi l'activation de la voie non canonique (van Amerongen and Nusse, 2009). Les molécules Wnt impliquées dans la régionalisation des somites sont Wnt4, Wnt6 et Wnt7a secrétées par l'ectoderme dorsal et les protéines Wnt1 et Wnt3 qui proviennent de la partie dorsale du tube neural (Parr et al., 1993) (Figure 5A). Dans les souris mutantes pour *Wnt1* et *Wnt3a* une partie du dermomyotome (épaxial) ne se forme pas notamment à cause d'une réduction de l'expression de *Myf5* (Ikeya and Takada, 1998). En ce qui concerne les récepteurs de Wnt, Fzd1 et Fzd6 sont présents dans la partie épaxiale des somites alors que Fzd7 est exprimé dans la partie hypaxiale (Borello et al., 1999a). L'action de Wnt1 via les récepteurs Fzd1/Fzd6 dans le domaine épaxial des somites induirait le contrôle de l'expression de *Myf5* de manière canonique, alors que l'induction de *MyoD* à travers Wnt7a et Fzd7 est indépendante de la voie canonique  $\beta$ -caténine et requiert la Protéine Kinase C (PKC) (Borello, 2006; Brunelli et al., 2007) (Figure 5B). L'identification de séquences de fixation spécifiques des facteurs de transcription Tcf/Lef, effecteurs de la voie Wnt, dans la partie 5' de l'enhancer épaxial de *Myf5* a permis de déterminer l'expression précise spatiotemporelle de ce dernier dans le domaine épaxial du somite (Borello, 2006). La voie Wnt est donc indispensable à la formation des muscles et cela semble être de façon dose dépendante au cours du développement embryonnaire. En effet, lorsque un antagoniste de la voie Wnt, Frzb1 (récepteur soluble) est délivré de manière trans-placentaire chez la souris, la sévérité des défauts n'est pas la même. Le mésoderme présomitique semble fortement affecté avec une inhibition de la myogenèse (perte de *Myf5*) alors que les cellules provenant de somites plus matures sont moins touchées (Borello et al., 1999b).





The diagram illustrates the signaling pathways involved in the expression of Myf5 and MyoD. In the dorsal region, Wnt7a/Wnt6 signaling leads to the activation of Myf5 and MyoD. In the lateral region, Bmp4 and Noggin signaling leads to the activation of Myf5 and MyoD. The legend indicates that Myf5 is represented by a green hexagon and MyoD by a purple rectangle.

## 2. La voie Sonic Hedgehog (*shh*)

La voie Shh contribue également à la spécification des progéniteurs musculaires dans les somites. Shh est sécrétée par la notochorde ainsi que par la plaque du plancher du tube neural (Figure 5A). Ce facteur fait partie de la famille de protéines Hedgehog (Echelard et al., 1993). Shh interagit avec le récepteur Patched ce qui induit le relargage de Smoothed (Smo). Par la suite Smo régule l'expression de gènes à travers les facteurs de transcription GLI (Lum, 2004). Dans les souris mutantes pour *Shh* ou *Smo* on observe une formation perturbée du sclérotome mais aussi une diminution de l'expression de Myf5 et MyoD au niveau épaxial du somite entraînant une perte des muscles épaxiaux (Zhang et al., 2001). Dans des explants de MPS de souris sauvages cultivés avec du Shh soluble ou en présence de notochorde on observe une induction de l'expression de MyoD alors que dans des explants issus d'embryons de souris *Myf5* mutantes la régulation de MyoD est défectueuse, ceci suggérant que Shh régule MyoD *via* l'intermédiaire de Myf5 (Borycki et al., 1999). Shh ne semble pas être nécessaire pour l'activation de Myf5 et MyoD au niveau des cellules du dermomyotome hypaxial. La fonction activatrice de Shh sur l'expression de *Myf5* se fait *via* les facteurs GLI (Figure 5C). En effet un enhancer de 651 paires de bases localisé à 6.6 kb en amont du site d'initiation de la transcription (TSS) de *Myf5* dans la région intragénique *MRF4/Myf5* permet l'expression de *Myf5* au niveau épaxial dans les embryons de souris 9.5 (Tajbakhsh et al., 1996). Dans les mutants *Shh* l'activité de cet enhancer est perdue (Gustafsson et al., 2002). La voie Shh intervient donc directement sur la régulation de *Myf5* entraînant ainsi la spécification des muscles épaxiaux.

## 3. La voie FGF

La famille FGF (Fibroblast Growth Factor) comprend des protéines sécrétées qui vont se lier à des récepteurs FGF de type tyrosine kinase afin de phosphoryler certains facteurs pour les activer et induire l'expression de gènes cibles. La voie FGF a été reportée comme agissant sur le destin des cellules souches dans différents organes (Calmont et al, 2006) comme le foie, l'épithélium sensoriel de l'audition ou les dents (Kelin et al, 2006). La voie de signalisation FGF est impliquée dans la myogenèse *in vivo* en promouvant la prolifération et la différenciation des cellules. Dans l'embryon, *Fgfr4* est transcrit aux différents sites de myogenèse dans les myoblastes proliférant (Marcelle et al, 1995). Il a été montré que chez la souris, *Fgfr4* est directement régulé par Pax3 (Lagha et al., 2008). Dans les mutants *Fgfr4*,



une perte importante des muscles des membres ainsi qu'une réduction de l'expression de marqueurs musculaires (Myf5, MyoD, Myosin) sont observées. De manière intéressante l'expression forcée de *Fgf8* dans les somites augmente l'expression de *Fgfr4* et favorise la différenciation myogénique alors que l'électroporation d'un dominant négatif de *Fgfr4* inhibe cette différenciation (Marics et al, 2002). La voie FGF est donc nécessaire à la myogenèse embryonnaire.

Les voies Wnt, Shh et FGF sont donc indispensables à la bonne spécification et/ou différenciation des différents progéniteurs musculaires. Cependant des acteurs inhibiteurs sont aussi requis pour la régulation de la myogenèse.

#### 4. La voie BMP

Les BMPs (Bone Morphogenetic Protein) font partie de la superfamille des TGF- $\beta$  et sont connues pour réguler négativement la myogenèse (Tsumaki and Yoshikawa, 2005). Ces derniers exercent leur fonction à travers deux récepteurs serine/thréonine kinase induisant l'activation des protéines SMAD qui sont à l'origine de l'activation ou la répression des gènes cibles (Miyazono et al., 2005). La protéine Bmp4 qui est exprimée dans la lame latérale du mésoderme maintient certaines populations de progéniteurs musculaires dans un état indifférencié. Cela se fait notamment par le maintien de Pax3 et le retard d'induction de Myf5 et MyoD (Pourquié et al., 1995). L'activité de Bmp4 est elle même antagonisée dans la partie dorso-médial du dermomyotome par Noggin dont l'expression est induite par Wnt et Shh (Hirsinger et al., 1997; Marcelle et al., 1997) (Figure 5A). Les souris mutantes pour *Noggin* présentent un développement anormal du myotome avec des somites restant dans un état épithélial et exprimant de manière forte Pax3 (McMahon et al., 1998).

#### 5. La voie Notch

La voie Notch est très étudiée depuis des décennies et son rôle dans la régulation négative de la myogenèse est très important (Lindsell et al., 1995). Notch qui est une protéine transmembranaire induit la communication cellule-cellule en interagissant avec ces ligands Jagged/Serrate et Delta présents au niveau des cellules avoisinantes. Lorsque Notch est activé, la partie intracellulaire va se localiser au noyau et se lier à la protéine RBP-J (homologue de Suppressor of Hairless) permettant ainsi l'activation du répresseur



transcriptionnel Hes1. Ce dernier induit alors la répression de *MyoD* (Kuroda et al., 1999; Jarriault et al, 1995). La surexpression de Hes1 inhibe *MyoD* en faisant des hétérodimères inactifs (Sasai et al., 1992; Sun et al., 2001). Des mutations du ligand *Delta1* et *RBP-J* induisent une différenciation myogénique prématurée et une perte de précurseurs musculaires menant à une hypotrophie des muscles (Schuster-Gossler et al., 2007; Vasyutina et al., 2007).

### 6. Morphogènes et muscles craniofaciaux

Sans entrer dans la complexité on retrouve au niveau de la tête les mêmes morphogènes cependant leurs rôles diffèrent. Alors que les signaux BMP inhibent à la fois la myogénèse dans les somites et les arcs branchiaux, les signaux Wnt et Shh sont des inhibiteurs de la myogénèse crâniale (Tzahor, 2003).

La myogénèse dans la tête est induite par une combinaison d'inhibiteurs de BMPs, Noggin-Gremlin et de Wnts, Frzb. Ces molécules sont sécrétées par les cellules de la crête neural crâniale et par les tissus entourant les myoblastes crâniens. Elles permettent la différenciation myogénique en bloquant les signaux Wnts et BMPs provenant du tube neural (Tzahor, 2003).

Dans le mésoderme de la tête l'absence d'acide rétinoïque dans la partie antérieure et l'induction de FGF dans la partie postérieure permettent l'expression respective de *Pitx2* et *Tbx1*. La partie antérieure est restreinte par l'augmentation de BMP et FGF induisant ainsi l'expression de *Musculin* (MyoR). Dans la partie postérieure le niveau important de FGF renforce l'expression de *Tbx1*. Le FGF active la prolifération des cellules progénitrices et favorise l'identité des muscles branchiaux *versus* les extraoculaires (Bothe and Dietrich, 2006; Bothe et al., 2011).

## IV- Contrôle génétique de la diversité musculaire

### 1. Les facteurs clés de la spécification des muscles du tronc et des membres

#### 1.1. Les MRFs (Myogenic Regulatory Factors)

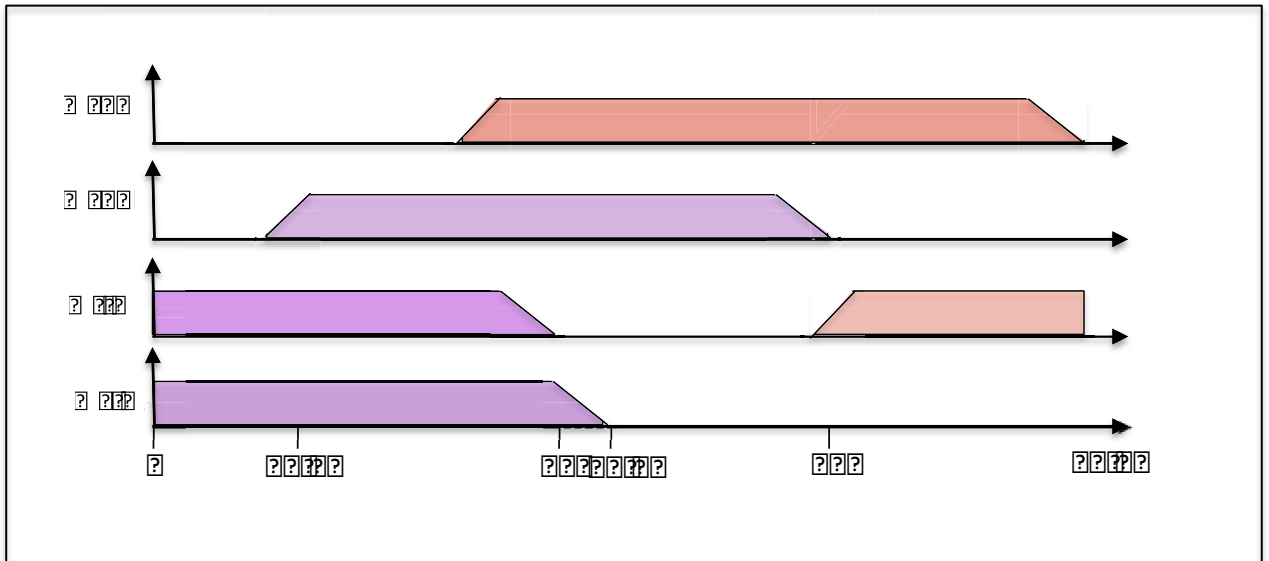
A la fin des années 80, des expériences d'hybridation « soustractive » ont permis d'identifier et d'isoler des transcrits spécifiques capables d'orchestrer la conversion myogénique de fibroblastes 10T1/2. Ce travail a permis l'identification de *MyoD* qui est capable de convertir une variété de types cellulaires (fibroblastes, chondrocytes, neurones,



amniocytes) en myoblastes (Choi et al., 1990; Davis et al., 1987; Weintraub et al., 1989). Peu de temps après la découverte de *MyoD*, trois autres gènes ont été identifiés : *Myf5*, myogénine et *MRF4* (aussi connu sous le nom *Myf6*) capables eux aussi d'induire des caractéristiques myogéniques dans des cellules non musculaires (Braun et al., 1990; Edmondson and Olson, 1989; Miner and Wold, 1990). Les MRFs contiennent trois domaines conservés : un domaine de transactivation dans la région N-terminale incluant un domaine riche en histidine/cystéine, un domaine bHLH (hélice-boucle-hélice) dans la région centrale et un domaine de transactivation C-terminal (Singh and Dilworth, 2013). *MyoD*, *Myf5*, myogénine et *Mrf4* sont nommés collectivement facteurs de régulation myogénique ou MRFs (Myogenic Regulatory Factors). *In vitro*, chaque MRF reconnaît et se fixe sur un site consensus CANNTG appelé E-Box qui est présent au niveau des promoteurs et enhancers de gènes musculaires (Massari and Murre, 2000). Les MRFs sont des facteurs de transcription de classe II et sont capables, soit de s'homodimériser entre eux ou bien de s'hétérodimériser avec des facteurs de classe I. Les facteurs de classe I incluent les E-protéines HEB/HTF4, E2-2/ITF-2 et E12/E47, exprimées de manière ubiquitaire dans de nombreux tissus à différents stades de développement (Conway et al., 2004; Parker et al., 2006). La région basique est requise pour la liaison à l'ADN alors que le domaine bHLH permet la dimérisation avec les autres facteurs (Blackwell and Weintraub, 1990; Perry and Rudnick, 2000; Puri and Sartorelli, 2000). Un « code myogénique » ou un « motif de reconnaissance musculaire » conservé et composé des acides aminés Alanine, Thréonine et Lysine permet de distinguer les MRFs par rapport à tous les autres facteurs de transcription bHLH. Ces trois AA confèrent aux MRFs la capacité d'activer des gènes spécifiques du muscle en interagissant avec des cofacteurs ou en induisant un changement de conformation ou bien les deux (Black and Olson, 1998; Molkentin et al., 1995).

Au cours de la myogenèse, les MRFs présentent une expression spatio-temporelle différente ainsi que des fonctions diverses. Par exemple, *Myf5* et *Mrf4* sont exprimés tout d'abord dans la partie centrale du dermomyotome et sont requis pour l'activation initiale de *MyoD* qui s'exprime un peu plus tard dans le domaine hypaxial puis épaxial (Buckingham, 2006). La mutation nulle de *Myf5* ou *MyoD* résulte en un développement presque normal du muscle squelettique, démontrant ainsi un degré de redondance fonctionnelle (Braun et al., 1992; Rudnicki et al., 1992). Cependant, différents mutants *Myf5* présentent des défauts moyens au niveau des muscles du tronc alors que l'on observe chez les mutants *MyoD*, un





QUESTION 1: A signal  $x(t)$  is shown in the figure. The signal is a trapezoidal pulse with a peak value of 1. The signal is zero for  $t < 0$  and  $t > 10$ . The signal is 1 for  $0 < t < 5$ . The signal is 0.5 for  $5 < t < 10$ . The signal is 0 for  $t > 10$ .

QUESTION 2: A signal  $y(t)$  is shown in the figure. The signal is a trapezoidal pulse with a peak value of 1. The signal is zero for  $t < 2$  and  $t > 8$ . The signal is 1 for  $2 < t < 6$ . The signal is 0.5 for  $6 < t < 8$ . The signal is 0 for  $t > 8$ .

QUESTION 3: A signal  $z(t)$  is shown in the figure. The signal is a trapezoidal pulse with a peak value of 1. The signal is zero for  $t < 0$  and  $t > 10$ . The signal is 1 for  $0 < t < 5$ . The signal is 0.5 for  $5 < t < 10$ . The signal is 0 for  $t > 10$ .

QUESTION 4: A signal  $w(t)$  is shown in the figure. The signal is a trapezoidal pulse with a peak value of 1. The signal is zero for  $t < 0$  and  $t > 10$ . The signal is 1 for  $0 < t < 5$ . The signal is 0.5 for  $5 < t < 10$ . The signal is 0 for  $t > 10$ .

retard de développement des muscles des membres. Ceci suggère que *Myf5* et *MyoD* auraient des rôles plus spécifiques, respectivement au cours du développement des muscles épaxiaux et hypaxiaux (Kablar et al., 1998). Dans le double mutant *Myf5/MyoD*, une perte complète des muscles est observée (Rudnicki et al., 1993). Ce double mutant s'est avéré être un triple mutant *Myf5/Mrf4/MyoD*. En effet, la localisation en tandem de *Mrf4* et *Myf5* dans le génome entraîne une dépendance d'expression de *Mrf4* et *Myf5* : les perturbations de l'un affectent l'expression de l'autre en cis. L'absence totale de muscle est donc seulement retrouvée en contexte triple mutant (Kassar-Duchossoy et al., 2004). Dans ce mutant une perte d'expression de *Myogénine (Myog)* est aussi observée. *Myog* est requise pour la différenciation terminale car les souris *Myog*<sup>-/-</sup> développent une faible portion de muscle squelettique. De plus, l'expression de marqueurs de différenciation comme Myosine ou *Mrf4* est réduite alors que l'expression de *MyoD* reste normale (hasty et al 1993 ; Venuti et al, 1993, (Nabeshima et al, 1993). La perte de *Myog* dans les mutants peut être partiellement restaurée par l'insertion d'un transgène qui contient *Mrf4* sous le contrôle du promoteur *Myog* (Zhu et al, 1997). Enfin, le rôle de *Mrf4* est plus complexe : il est exprimé de manière transitoire dans le myotome de l'embryon E9 de souris, juste après *Myf5*. Son expression se réduit à partir du stade E11.5 pour se renforcer au stade E16 dans les fibres différenciées (Kassar-Duchossoy et al., 2004). Cette expression temporelle complexe suggère que *Mrf4* possède des rôles différents au cours de la myogenèse embryonnaire (Zhu et al, 1997). En tenant compte de toutes ces études, on peut facilement suggérer que *MyoD* et *Myf5* sont des facteurs de spécification alors que *Myog* un facteur de différenciation. Enfin, *Mrf4* est impliqué dans la spécification des myoblastes et plus tardivement dans la différenciation terminale (Figure 6).

## 1.2. Les Pax

Les gènes *Pax* jouent un rôle clé dans la spécification des tissus et le développement des organes. Cette famille de gènes code pour des facteurs de transcription qui sont caractérisés par la présence d'un domaine « paired » (PD), séquence spécifique de liaison à l'ADN de 128 acides aminés. Plusieurs membres de la famille *Pax* (*Pax3*, *Pax4*, *Pax6* et *Pax7*) contiennent un domaine de liaison à l'ADN supplémentaire : un homéodomaine (HD). Un dernier domaine hautement conservé de 8 AA est retrouvé entre le PD et le HD (Chi and



Epstein, 2002; Lang et al., 2007). Ce dernier servirait de motif de liaison pour des cofacteurs afin d'inhiber la transcription de gènes cibles (Eberhard et al., 2000).

*Pax3* et *Pax7* sont les seuls *Pax* nécessaires à la myogenèse. Ils ne sont pas exprimés de manière tissu-spécifique mais sont retrouvés dans le neurectoderme, dans des sous domaines du cerveau, dans la crête neurale et dans le muscle squelettique.

De manière individuelle, *Pax3* est exprimé dans le mésoderme pré-somitique et dans les somites épithéliaux, avant d'être restreint dans la partie dorsale et ventrale du dermomyotome. Il est requis pour la formation des muscles du tronc et pour la délamination et la migration des cellules myogéniques progénitrices au niveau du territoire des membres (Kassar-Duchossoy, 2005). Les embryons de souris présentant une mutation spontanée homozygote de *Pax3* (mutation *splotch*) ne développent pas la partie hypaxiale du dermomyotome, ayant pour conséquence une absence de muscles des membres et du diaphragme. Concernant les muscles dérivés du domaine épaxial ces derniers sont affectés mais moins fortement (réduits, désorganisés) (Tremblay et al, 1998, Relaix et al., 2004).

En comparaison, *Pax7* s'exprime dans la partie centrale du dermomyotome et ne semble pas essentiel au développement musculaire embryonnaire. Cependant, la formation des muscles des doubles mutants *Pax3/Pax7* est plus affectée que dans le mutant *splotch* puisque seul le myotome précoce se développe. Ceci suggère que *Pax7* est capable de compenser l'absence de *Pax3*. En effet, l'introduction de la séquence codante de *Pax7* dans un allèle de *Pax3* a permis de valider que *Pax7* peut remplacer les fonctions de *Pax3* au cours de la myogenèse du tronc et des membres (Buckingham and Relaix, 2007). La génération de mutants conditionnels pour *Pax3* et *Pax7* a démontré que la perte des cellules exprimant *Pax3* (progéniteurs embryonnaires) engendre une létalité embryonnaire et empêche l'émergence de cellules *Pax7*-positives alors que la perte du lignage exprimant *Pax7* engendre la formation de muscles plus petits (Hutcheson et al., 2009). Les cellules exprimant *Pax3* sont donc nécessaires à la formation des fibres primaires des muscles des membres alors que les cellules *Pax7*-positives (progéniteurs foetaux) permettent le développement des fibres secondaires *via* la fusion et la croissance. Dans les doubles mutants *Pax3/Pax7*, aucune activation de *Myf5* ou *MyoD* n'est détectée, les cellules acquièrent donc d'autres destins cellulaires ou meurent. Dans les muscles épaxiaux, *Pax3* semble activer *Myf5* en contrôlant l'expression de *Dmrt2*, qui à son tour active l'enhancer épaxial de *Myf5* dans les somites (Sato et al., 2010). Dans le myotome hypaxial, deux



éléments enhancers de *Myf5*, localisés à -110 kb et -57.5 Kb du site de transcription, sont des cibles directes de Pax3 (Bajard, 2006). En ce qui concerne *MyoD*, l'activation directe par Pax3/7 n'a pas encore été démontrée en embryon. Cependant, une séquence régulatrice en amont de *MyoD* est une cible de Pax7 dans les cellules satellites (Buckingham and Vincent, 2009). De plus, aucune formation des muscles des membres et du tronc n'est observée dans les mutants *Myf5(MRF4)/Pax3*. L'absence de *Myf5* et *Mrf4* cause alors une perte d'expression de *MyoD*, requis pour la détermination musculaire, suggérant que l'activation initiale de *MyoD* dépend de Myf5 et de Pax3 (Tajbakhsh et al., 1997) (Figure 7).

Pax3 intervient dans la délamination et la migration des cellules progénitrices des membres. En effet, les souris mutantes pour *Pax3* présentent une absence de muscles des membres, due à la perte d'expression du récepteur tyrosine kinase c-Met. C-Met et son ligand HGF (Hepatocyte Growth Factor), sont essentiels à la délamination et à la migration des cellules vers le territoire des membres. c-Met est exprimé par les progéniteurs musculaires alors que HGF est sécrété par les cellules mésenchymateuses situées autour (Birchmeier and Brohmann, 2000; Dietrich et al., 1999). Un crible impliquant des cellules sur-exprimant *Pax3* a permis d'identifier les gènes régulés positivement ou négativement dans les somites et les cellules précurseurs des muscles des membres (Lagha et al., 2010). Parmi ces gènes, on retrouve *Lbx1* et *Pitx2* qui codent pour des facteurs de transcription nécessaires à la spécification des cellules des muscles des membres dont je détaillerai les rôles ultérieurement. Les cellules exprimant Pax3/Pax7 constituent aussi un « pool » de cellules satellites indispensables à toute la vie adulte du muscle (Kassar-Duchossoy, 2005; Schienda et al., 2006).

### 1.3. Les Six

Les membres de la famille des gènes *Six* ont été identifiés comme homologues du gène *sine oculis (so)* de drosophile, gène essentiel à la formation des yeux. Les protéines de la famille Six sont des facteurs de transcription caractérisés par deux domaines conservés : l'homéodomaine de type Six qui se lie à l'ADN et le domaine Six N-terminal qui interagit avec des co-activateurs ou des co-répresseurs de transcription (Tessmar et al., 2002; Zhu et al., 2002). Les protéines Six se lient aux protéines Eya1 et Eya2 (Eyes-absent homologs) permettant ainsi leur translocation dans le noyau où ils agissent comme cofacteurs pour activer les gènes cibles des Six que sont *Pax3*, *MyoD*, *Mrf4* et *Myog*. Chez la souris, *Six1*, *Six4*,

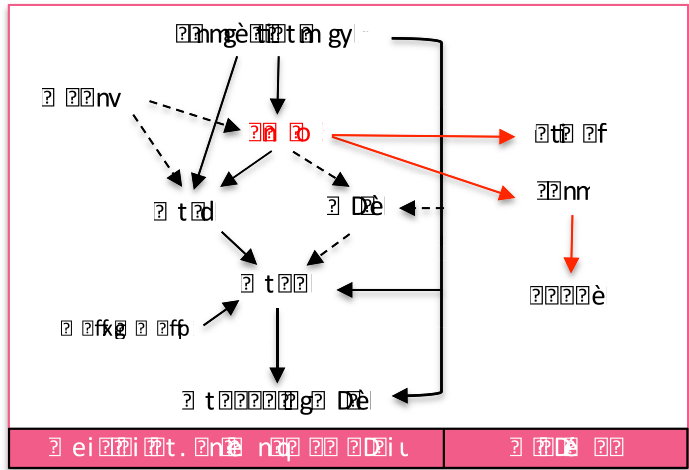
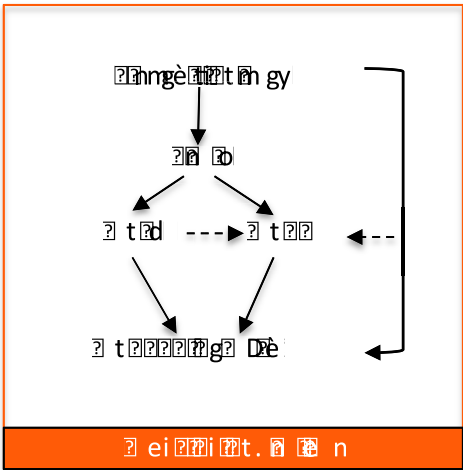
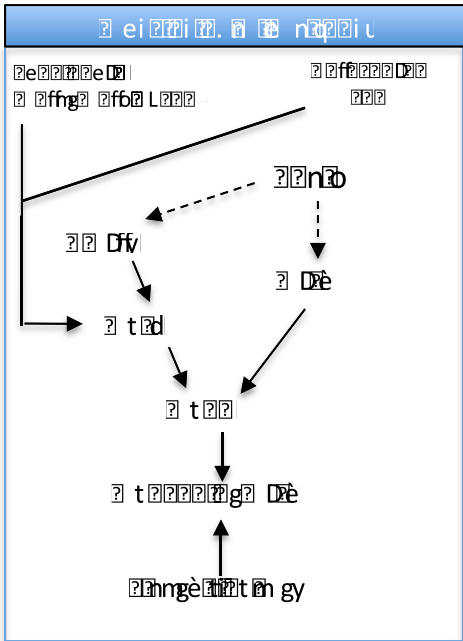


Diagram illustrating the embryonic development of the heart, showing the relationship between the heart tube and the septum transversum. The diagram shows the following components and their relationships:

- Heart Tube (Tubus cordis):** The central structure that forms the heart.
- Septum transversum (Septum transversum):** A horizontal partition that separates the foregut from the midgut.
- Septum secundum (Septum secundum):** A second horizontal partition that forms the right and left ventricles.
- Septum primum (Septum primum):** A vertical partition that forms the interatrial septum.
- Septum inferius (Septum inferius):** A vertical partition that forms the interventricular septum.
- Septum transversum (Septum transversum):** A horizontal partition that forms the diaphragm.

*Eya1* et *Eya2* sont exprimés dans le dermomyotome et ensuite dans les cellules Pax3 positives. De plus, on retrouve leur expression dans le muscle squelettique différencié (Grifone, 2005). Concernant *Eya1* et *Eya2*, ils sont principalement exprimés dans le domaine épaxial et hypaxial du dermomyotome. Le rôle de Six et Eya au cours de la myogenèse est critique. En effet, les souris doubles mutantes *Six1/Six4* et *Eya1/Eya2* présentent des phénotypes très sévères. Notamment, on retrouve une perte des muscles dérivant du dermomyotome hypaxial, les muscles des membres et une partie des muscles du tronc, ceci étant dû à l'absence de *Pax3* et de ces cibles directes *Lbx1* et *C-met* (Grifone et al., 2007). La myogenèse épaxiale ne semble pas affectée car celle-ci débute avant que la régulation fonctionnelle des *Six* se mette en place (Figure 7).

Six et Eya régulent aussi directement l'expression des gènes *MyoD* et *Myf5* via leurs enhanceurs respectifs. En effet, *Six1/4* se fixent sur une séquence d'un élément enhanceur de 145 bp, localisée à -57,5 kb et qui contrôle l'expression de *Myf5* dans les membres et dans le dermomyotome hypaxial mature (Giordani et al., 2007). L'expression de *MyoD* est contrôlée par un enhanceur situé à -20 kb et par un second distal à -6 kb du gène (Goldhamer et al., 1995; Asakura et al., 1995). Ces deux enhanceurs contiennent aussi des sites de liaison pour *Six1/4* et *Six2*. La mutation de ces sites *in vivo* abolit presque complètement l'expression du transgène contrôlé par ces deux enhanceurs et le promoteur proximal démontrant ainsi la régulation directe de *MyoD* par le complexe Six/Eya (Relaix et al., 2013). Les gènes *Six* ne sont pas uniquement exprimés dans les muscles du tronc et des membres mais aussi dans ceux de la tête. Cependant, les doubles mutants *Six1/Six4* ne présentent aucun phénotype musculaire (tête), probablement lié à une compensation par *Six2*. Pour finir, les *Six* sont exprimés dans les cellules satellites adultes et jouent un rôle important dans la différenciation des fibres musculaires (Le Grand et al., 2012; Niro et al., 2010).

#### 1.4. Les facteurs impliqués dans la formation des muscles des membres

D'autres facteurs, *Lbx1* et *Meox2* sont exprimés dans les myoblastes qui migrent dans les bourgeons des membres (Yokoyama and Asahara, 2011). *Lbx1* contrôle la migration des cellules progénitrices à travers l'activation de CXRC4, qui code pour le récepteur SDF1 (Vasyutina et al., 2005). Son absence induit une perte de la masse musculaire dorsale des membres supérieurs et la perte totale des muscles des membres inférieurs (Jagla et al., 1995; Brohmann et al., 2000). *Lbx1* est une cible directe de Pax3 (Lagha et al., 2010). *Meox2*





est exprimé dans les somites et les cellules en migration (Mankoo et al., 1999). Dans les mutants *Meox2*, la migration semble s'effectuer normalement mais une population restreinte de muscles des membres est absente. De plus, l'expression de *Myf5* et *Pax3* est réduite dans les cellules myogéniques des membres, ce qui suggère un rôle d'activateur transcriptionnel de *Meox2* dans ces cellules (Buckingham, 2001) (Figure 7).

### 1.5. Les facteurs répresseurs des MRFs

Afin de limiter l'activité prématurée ou excessive des MRFs, des facteurs de régulation négatifs existent pour contrôler ces derniers. La plupart de ces facteurs sont des protéines à domaine HLH comme les protéines Id (Inhibitor of differentiation, Id1-4), Twist et Musculin (MyoR) ou encore Mist-1 (Puri and Sartorelli, 2000).

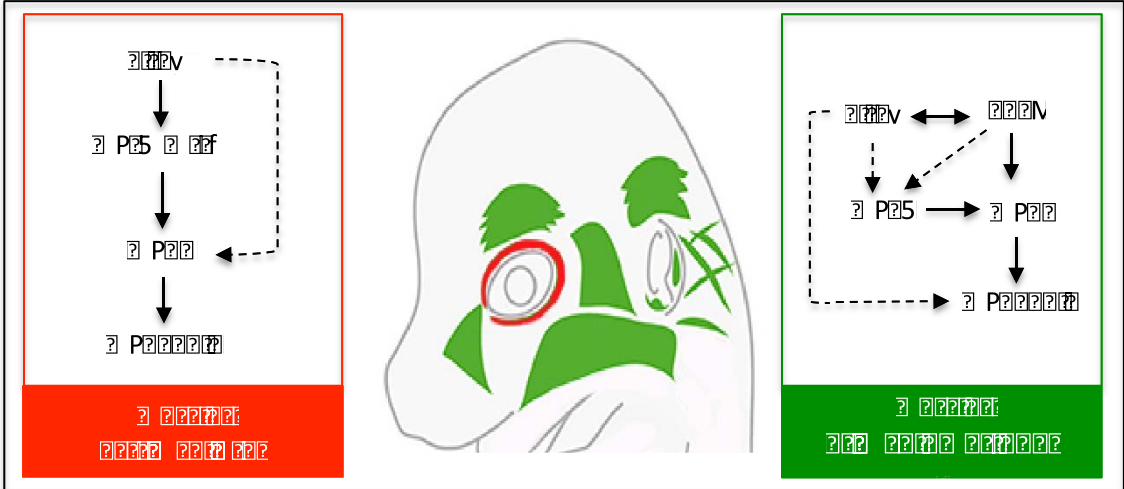
Twist forme des hétérodimères avec les E-protéines, entraînant leur séquestration et empêchant les interactions avec les MRFs. Twist peut aussi inhiber la myogenèse en se liant directement aux protéines Mef2 ou aux MRFs. MyoR et Mist-1 forment des hétérodimères avec les MRFs et se lient aux E-boxes, cependant ces dimères ne sont pas capables d'activer la transcription (Lermercier et al., 1998; Lu et al., 1999).

Similairement, les protéines Id s'hétérodimérisent avec les E-protéines ou avec les MRFs pour empêcher leur activité transcriptionnelle (Neuhold and Wold, 1993). Une étude a montré que les protéines Id seraient régulées positivement par les MRFs dans les myoblastes (Wyzykowski et al., 2002).

## 2. Les facteurs clés de la spécification des muscles craniofaciaux

### 2.1 Pitx2

Les trois gènes *Pitx* présents chez les vertébrés codent pour des facteurs de transcription à homéodomaine (classe paired). Ils jouent un rôle important dans la formation de plusieurs organes au cours du développement et notamment celle du muscle squelettique. Le facteur de transcription à homéodomaine *Pitx2* est présent précocement dans le mésoderme crânial et est un acteur clé de la spécification des muscles extra-oculaires (MEO), en régulant la survie du mésoderme pré-myogénique (Gage et al., 1999; Zacharias et al., 2011). Par la suite, il est nécessaire à l'activation de *Myf5*, *Mrf4* et *MyoD* (Zacharias et al., 2011) (Figure 8). En l'absence de *Mrf4* et *Myf5*, *Pitx2* seul ne peut assurer la survie des cellules et l'activation de l'expression *MyoD* (Sambasivan et al., 2009). Ce rôle



ସର୍ବମୁଖ୍ୟତଃ, ଏହି ପ୍ରକାରର ରୋଗ ହେଉଛି ଏକ ପ୍ରକାରର ପ୍ରତିରକ୍ଷା ପ୍ରଣାଳୀର ଅଭାବ ଯାହାକି ଏହି ପ୍ରକାରର ରୋଗକୁ ସମ୍ଭବ କରେ । ଏହି ପ୍ରକାରର ରୋଗ ହେଉଛି ଏକ ପ୍ରକାରର ପ୍ରତିରକ୍ଷା ପ୍ରଣାଳୀର ଅଭାବ ଯାହାକି ଏହି ପ୍ରକାରର ରୋଗକୁ ସମ୍ଭବ କରେ । ଏହି ପ୍ରକାରର ରୋଗ ହେଉଛି ଏକ ପ୍ରକାରର ପ୍ରତିରକ୍ଷା ପ୍ରଣାଳୀର ଅଭାବ ଯାହାକି ଏହି ପ୍ରକାରର ରୋଗକୁ ସମ୍ଭବ କରେ । ଏହି ପ୍ରକାରର ରୋଗ ହେଉଛି ଏକ ପ୍ରକାରର ପ୍ରତିରକ୍ଷା ପ୍ରଣାଳୀର ଅଭାବ ଯାହାକି ଏହି ପ୍ରକାରର ରୋଗକୁ ସମ୍ଭବ କରେ ।

dans la survie et la spécification myogénique suggère que ce dernier assume le même rôle que *Pax3* au niveau de la myogenèse des muscles du corps. De manière intéressante, la myogenèse des MEO semble sensible à la dose de *Pitx2*, comme le montre la présence de défauts de morphogenèse et d'expression des MRFs dans les souris *Pitx2* hétérozygotes (Diehl et al., 2006). Il a été démontré que l'expression de *Pitx2* est plus importante dans les MEO que dans les muscles des membres, forte expression qui est maintenue dans les MEOs dystrophiques et âgés (Hebert et al., 2013). Pour finir, dans le premier arc branchial l'inactivation de *Pitx2* induit la perte des marqueurs de spécification des myoblastes (*Tbx1*, *MyoR*, *Capsulin*), ce qui démontre que *Pitx2* est aussi requis pour la différenciation et la survie des cellules progénitrices dans cet arc branchial (Shih et al., 2007, 2008).

*Pitx2* est aussi exprimé dans les progéniteurs musculaires des membres et est contrôlé par *Pax3* (Lagha et al., 2010; L'Honore et al., 2010). Dans les mutants *Pitx2*, le début de la myogenèse des membres est affecté et cela est dû à la perte d'expression de *MyoD*, suggérant que *Pitx2* contrôle l'expression de *MyoD* dans les membres (L'Honore et al., 2010)

## 2.2 *Tbx1*

*Tbx1* est un facteur de transcription T-box, exprimé dans les arcs branchiaux avant l'activation des MRFs (Kelly, 2004), et requis pour la formation des muscles branchiomériques. Son inactivation induit l'atrophie des arcs branchiaux secondaires et le développement sporadique des muscles du premier arc (Kelly, 2004). Un mutant *Tbx1* spécifique du mésoderme montre que *Tbx1* est requis de manière cellule-autonome pour la myogenèse des arcs branchiaux (Kong et al., 2014). En l'absence de *Tbx1*, les muscles squelettiques sont hypoplasiques et souvent unilatéraux. De plus, ce phénotype est plus sévère lorsque *Myf5* est absent. Ceci indique que *Myf5* contrôle la myogenèse indépendamment de *Tbx1* et que *MyoD* qui normalement compense la perte de *Myf5* est dépendant de *Tbx1* (Grifone et al., 2008). *Tbx1* agit donc de manière synergique avec *Myf5* pour assurer l'expression de *MyoD* (Sambasivan et al., 2009) (Figure 8). *Tbx1* n'est donc pas absolument requis pour l'activation précoce de la myogenèse des arcs branchiaux mais plutôt pour que les étapes de spécification se passent de manière robuste (Grifone et al., 2008; Sambasivan et al., 2009).



### 2.3 Islet1

Le facteur à homéodomaine LIM, Islet1 (*Isl1*) est un marqueur du mésoderme splanchnique et du premier arc branchial chez la souris et le poulet. Des expériences de suivi des cellules dans l'embryon de poulet ont permis de montrer que les cellules *Isl1* positives sont présentes dans certains muscles branchiomériques et dans les muscles laryngés (Nathan et al., 2008). En effet, les cellules exprimant *Isl1* contribuent à la formation de sous-population de muscles branchiomériques dont certains de la mâchoire (Harel et al., 2009; Nathan et al., 2008). Le rôle de *Isl1* au niveau des muscles de la tête reste encore à élucider mais celui-ci pourrait être nécessaire au maintien de cellules progénitrices spécifiques de sous-populations de muscles. En effet, il a été montré que la surexpression de *Isl1* réprime l'expression de marqueurs de différenciation tels que *Myog* ou *Myosin* (Harel et al., 2009).

### 2.4 Les facteurs répresseurs et la myogenèse crâniofaciale.

Dans la myogenèse branchiomérique, deux facteurs répresseurs à domaine bHLH sont requis, *Musculine* (*MyoR*) et *Capsuline* (*Tcf21*) (Sambasivan et al., 2011). Ce sont des marqueurs des précurseurs musculaires indifférenciés du premier arc branchial (von Scheven, 2006). La mutation chez la souris de l'un ou l'autre n'affecte pas le développement des muscles branchiomériques alors qu'on observe chez les doubles mutants *MyoR/Capsulin*, une absence d'activation de l'expression de MRFs dans le premier arc. Cela induit une mort cellulaire importante et une perte des muscles principaux de la mastication. Ils sont donc tous les deux requis pour l'activation de *Myf5* dans l'arc mandibulaire, à l'origine des muscles de fermeture de la mâchoire (Lu et al., 2002). Pour finir *MyoR* et *Capsuline* sont connus pour contrôler le niveau transcriptionnel de *Myf5* et *MyoD* dans les arcs branchiaux à travers des enhanceurs spécifiques (Moncaut et al., 2012).

## V- La différenciation terminale des muscles squelettiques

### 1. Les facteurs : *Mef2*, *Myog* et *Nfix*

Les MRFs sont assistés par la famille de facteurs de transcription *Mef2* dans le but de favoriser l'expression de gènes spécifiques du muscle (Black and Olson, 1998). La protéine *Mef2* appartient à la famille des facteurs de transcription contenant un domaine MADS (*MCM1*, *Agamous*, *Deficiens*, *Serum response factor*), elle est composée de 4 membres : *Mef2A* à *Mef2D* qui codent chacun pour un gène différent. Par comparaison avec les MRFs,



les gènes *Mef2* sont plus largement exprimés au cours du développement et ne sont donc pas restreints aux cellules musculaires (Potthoff and Olson, 2007). Les protéines Mef2 se lient préférentiellement à des éléments de séquences d'ADN riches en nucléotides A/T (motif C/TTA(A/T)<sub>4</sub>TAG/A) qui sont retrouvés au niveau des promoteurs de gènes spécifiques du muscle (Black and Olson, 1998; McKinsey et al., 2002). A l'inverse des MRFs, Mef2 n'a pas la capacité d'induire *in vitro* le programme de différenciation myogénique, cependant il aurait un rôle critique dans l'assistance des MRFs au cours de la myogenèse. En effet, la formation d'hétérodimères Mef2/MRFs via leurs domaines respectifs MADS et bHLH potentialise le pouvoir myogénique des MRFs (Black and Olson, 1998; Penn et al., 2004). Les hétérodimères formés peuvent se fixer soit aux E-box spécifiques des MRFs soit aux sites de liaison à l'ADN spécifiques de Mef2 (Molkentin et al., 1995). Les gènes *Mef2A*, *2C* et *2D* exprimés dans le muscle squelettique, présentent une expression et un épissage dépendant de MyoD (Penn et al., 2004). La délétion conditionnelle précoce de *Mef2c* dans le muscle squelettique de souris révèle son importance dans la régulation des gènes sarcomériques. En effet, dans ce contexte mutant, les muscles se différencient mais forment des myofibres contenant des sarcomères anormalement assemblés. De plus, on observe une diminution de l'expression de gènes spécifiques de la bande M, Myomesin et Myomesin 2. Cela montre que *Mef2c* est essentiel pour la maturation des myofibres et donc pour la maintenance de l'intégrité musculaire (Potthoff et al., 2007). Mef2 régule aussi l'expression de *Myog* fournissant ainsi une boucle de régulation positive de la différenciation myogénique.

Un des acteurs clés de la différenciation terminale des muscles est *Myog*, un membre de la famille des MRFs. En effet, dans les mutants *Myog* la myogenèse secondaire est totalement affectée (Venuti et al., 1995). De plus, dans les cellules C2C12 mutées pour ce gène il y a une réduction drastique d'expression des gènes musculaires tardifs tels que la *MyHC* (*Myosin Heavy Chain*) et *MCK* (*Muscle Creatine Kinase*) alors que les gènes musculaires plus précoces ne sont pas affectés, montrant ainsi le contrôle direct de l'expression de ces gènes par *Myog* (Du et al., 2012).

A ce jour, les différentes analyses transcriptomiques réalisées sur des myoblastes embryonnaires et fœtaux (isolés par FACS) ont permis d'identifier des gènes codant pour des facteurs de transcription différentiellement exprimés dans ces deux populations de myoblastes (Biressi et al., 2007). Par exemple, le facteur de transcription Nfix (Nuclear factor one X), activé par Pax7, est fortement présent dans les myoblastes fœtaux où il forme un





complexe avec PKC theta, autre marqueur de la myogenèse fœtale (Messina et al., 2010). Ce complexe se lie à Mef2 et entraîne sa phosphorylation menant ainsi à l'activation transcriptionnelle des gènes fœtaux comme *MCK* et *Enolase-β* et en parallèle à la répression de gènes embryonnaires. En effet, la mutation de *Nfix* induit la maintenance de gènes embryonnaires et sa surexpression mène à la répression de l'isoforme lente de la myosine embryonnaire (*MyHC-I/β*). Les facteurs *Nfix* peuvent aussi se fixer sur le promoteur de *Myog* et par effet direct augmenter la transcription de gènes dépendants de *Myog*.

## 2. Fusion des myoblastes : Fibres primaires et secondaires

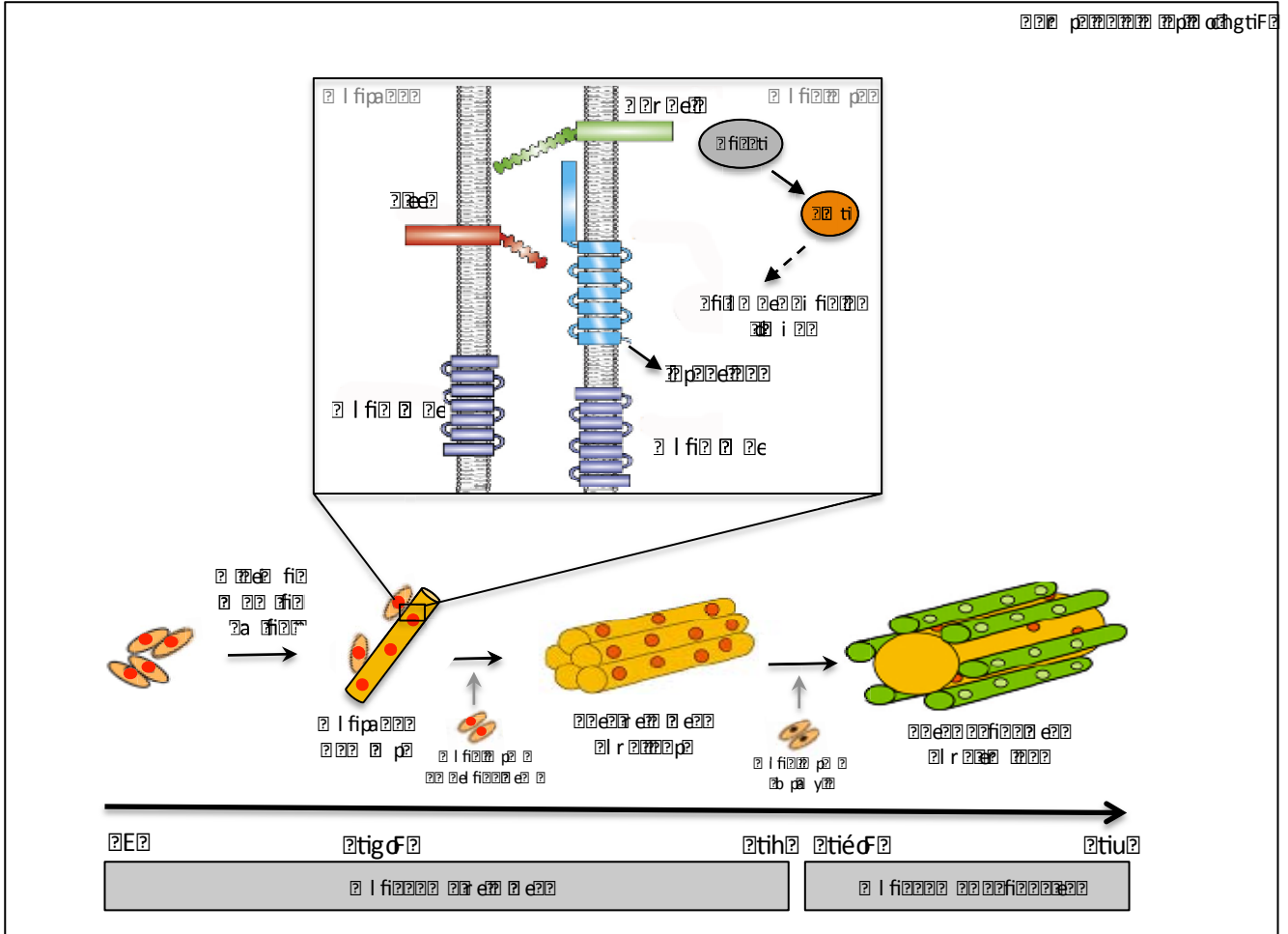
Toutes les fibres musculaires sont formées par la fusion de cellules myogéniques. La formation de ces fibres se produit au cours de deux phases du développement : une précoce et une plus tardive (Kelly and Zacks, 1969).

Durant la phase précoce (myogenèse primaire), les fibres primaires sont générées par la fusion de myoblastes embryonnaires entre eux. Les muscles consistent alors en un petit nombre de myotubes, contenant seulement quelques noyaux. Les fibres primaires servent d'échafaudage aux fibres secondaires.

Par la suite, les myoblastes fœtaux vont fusionner avec les myotubes primaires mais aussi entre eux pour donner les fibres secondaires (Evans et al., 1994 ; Duxson et al., 1999) (Figure 9). Les fibres secondaires sont initialement formées au site d'innervation des fibres primaires et sont entourées de lame basale (Duxson et al., 1989). La perte d'innervation ne semble pas affecter la myogenèse primaire mais impacte la survie des myotubes primaires (Ashby et al., 1993; Ross et al., 1987).

Les fibres primaires et secondaires diffèrent de part leur expression d'isoformes de *MyHC*. Chez les mammifères, les fibres primaires expriment la forme embryonnaire rapide et la forme lente (*MHC-I/β*), puis peu avant la fin de la formation des fibres primaires, l'isoforme périnatale/néonatale rapide. Par comparaison, les fibres secondaires expriment dès le début les isoformes rapides embryonnaires et périnatales et non pas l'isoforme lente *MHC-I/β*. En général, les fibres primaires sont programmées pour avoir un phénotype lent alors que les secondaires adoptent principalement un aspect rapide (Wigmore and Evans, 2002; Zhang and McLennan, 1998).

Au niveau moléculaire, différentes protéines permettent de réguler ces différentes phases. Les protéines membranaires Intégrine β1, caveolin-3 et VCAM (Vascular Cell



Microtubule plus end rescue and catastrophe cycles are regulated by microtubule rescue proteins (YBP and MTBP). YBP binds to the plus end to prevent rescue, while MTBP binds to the plus end to prevent catastrophe. The diagram shows a sequence of microtubule growth, rescue, catastrophe, rescue, and catastrophe, with corresponding protein binding events.

Adhesion Molecule) sont nécessaires à la fusion primaire des myoblastes entre eux alors que les protéines NFATc2, activée par le calcium, Calmoduline et la cytokine IL-4 sont critiques pour la seconde phase de fusion entre les myotubes et les myoblastes fœtaux (Horsley and Pavlath, 2004). Des études *in vitro* ont montré que les cadhérines permettent la reconnaissance et l'adhésion entre les myoblastes chez la souris (George-Weinstein et al, 1997). Cependant, les cadhérines ne se sont pas révélées indispensables *in vivo*, suggérant la contribution d'autres protéines telles que les intégrines et les protéines transmembranaires (ex :Immunoglobulines) (Hollnagel et al., 2002; Horsley and Pavlath, 2004; Schwander et al, 2003). En effet, le génome de souris code pour les protéines orthologues de Duf/Kirre/Rst et Sns/Hbs (Sellin, 2002; Ueno et al., 2003) qui sont nécessaires à la fusion chez la drosophile. Kirre1 et les protéines Nephrin sont connues pour former des complexes d'adhésion cellule-cellule. De plus, une étude a montré que Nephrin, un orthologue de Sns, avait un rôle conservé chez les mammifères. En effet, les myoblastes embryonnaires isolés de mutants nuls *Nephrin*, présentent des défauts de fusion *in vitro*. Une expérience de mélange de cellules a démontré la réquisition asymétrique de Nephrin durant la fusion, ce qui suggère que cette protéine est nécessaire dans les myoblastes et pas dans les myotubes (Sohn et al., 2009). Les deux régulateurs du cytosquelette d'actine, *Cdc42* et *Rac1*, sont aussi requis pour la fusion puisque chez la souris, la fusion est affectée chez les mutants conditionnels de chacun d'entre eux (Vasyutina et al., 2009). L'homologue murin de *Mbc*, *Dock1*, semble aussi impliqué dans la fusion. En effet, le mutant nul *Dock1* présente une réduction drastique du nombre de muscles causée par une myogenèse primaire défectueuse (Laurin et al., 2008). Une protéine transmembranaire appelée « Myomaker » retrouvée à la surface des myoblastes est également indispensable à la fusion (Figure 9). La surexpression de cette protéine favorise la fusion alors que chez le mutant murin, l'absence de fibres multinuclées cause une mort périnatale (Millay et al., 2013). De plus, l'expression de Myomaker dans des fibroblastes suffit à ces derniers pour fusionner avec des myoblastes. Pour finir, un récent criblage de perte de fonction *via* la technique CRISPR-Cas9 a permis d'identifier une nouvelle protéine de fusion appelée Myomixer. Les souris nulles pour *Myomixer* présentent des défauts ressemblant à ceux de *Myomaker*. Cependant, son expression dans des fibroblastes ne permet pas la fusion avec des myoblastes. Ceci suggère donc que *Myomixer* améliore l'activité fusogénique de *Myomaker* pour induire le mélange membranaire, potentiellement par l'établissement de pores de fusion (Bi et al., 2017). Ces différentes étapes de fusion sont



contrôlées en amont en grande partie par les MRFs mais aussi *via* d'autres facteurs tels que Mef2 ou Nfix précédemment décrits.

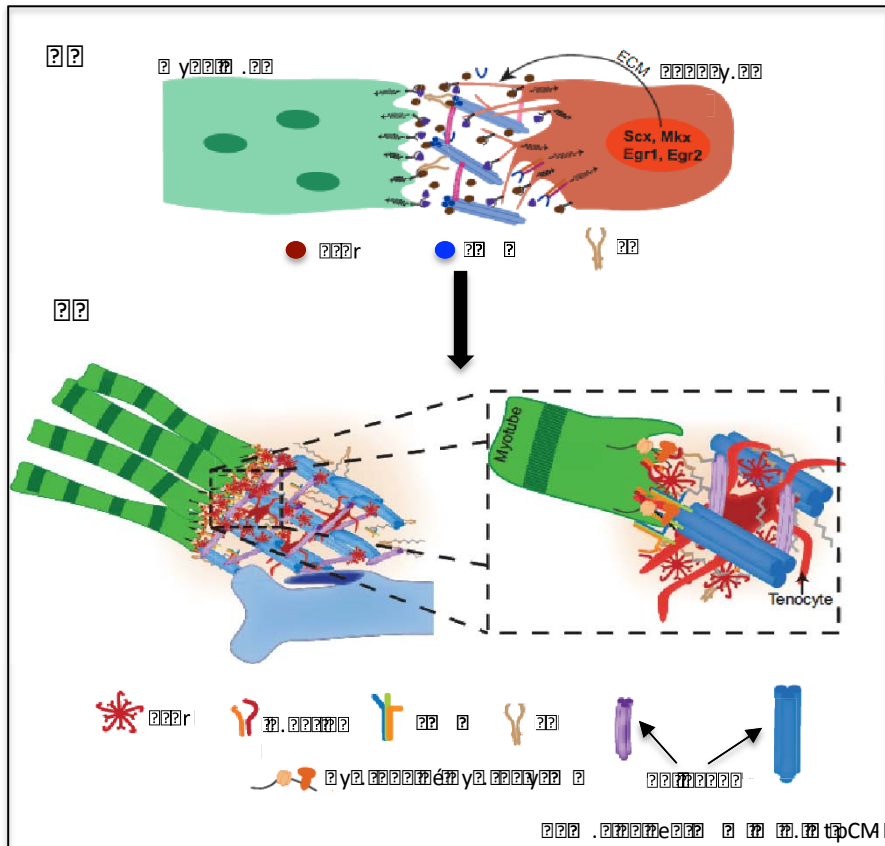
### 3. Attachement et jonction myotendineuse

Chez les vertébrés, les muscles s'attachent à l'endosquelette par l'intermédiaire des tendons. La formation de ce système musculo-squelettique, qui comprend les muscles, les tendons et les os se fait différemment selon les types de muscles (du tronc, des membres et de la face) et l'origine des tendons auxquels ils s'attachent n'est pas la même (Kardon, 1998).

Tout d'abord, il y a spécification, migration puis différenciation des tendons, et enfin les jonctions myotendineuses (MTJ) se forment. Les progéniteurs de tendon ou ténocytes expriment le gène *Scx* (*scleraxis*) qui code pour un facteur de transcription à domaine bHLH (Cserjesi et al., 1995). Sa découverte a permis de mettre en évidence que la spécification des progéniteurs de tendons dépend de leur interaction avec d'autres tissus. En effet, dans les futurs tendons des muscles du tronc, l'expression de *Scx* est liée à la sécrétion de facteurs de croissance aux fibroblastes (FGF) par le myotome, tandis que dans les bourgeons des membres, cette expression serait dépendante de l'ectoderme (Schweitzer et al., 2001).

L'étape suivante de différenciation des cellules de tendon est majoritairement dépendante des muscles chez les vertébrés. Il est connu que les facteurs de transcription *Egr1* et *Egr2* sont impliqués dans la différenciation des cellules tendons (Lejard et al., 2011). L'expression de *Egr1* et *Egr2* dans ces cellules est dépendante des signaux *Fgf4* provenant des extrémités des muscles et est suffisante pour induire l'expression des gènes de différenciation des tendons. Le facteur de transcription de la famille iroquois Mohawk (*Mkx*) est aussi nécessaire à la différenciation des ténocytes. Son expression tardive chez la souris permet de compléter l'effet de *Scx* en induisant l'expression de protéines de la matrice extracellulaire (ECM) et en réprimant le destin myogénique et squelettogénique des tendons (Liu et al., 2015; Otabe et al., 2015).

Les études réalisées sur différents modèles animaux (drosophile, poulet, souris et poisson zèbre) ont révélé le lien crucial entre les ténocytes se développant et la dynamique de la matrice extracellulaire (ECM, collagène, laminine et thrombospondine). La formation de la jonction myotendineuse peut être considérée comme un processus en deux étapes.



The diagram illustrates the differentiation of a myotube from a myoblast. The top part shows a myoblast (green) with nuclei and cytoplasmic organelles, and a myotube (red) with nuclei and cytoplasmic organelles. The myotube is labeled 'ECM' and contains transcription factors 'Scx, Mlx, Egr1, Egr2'. A legend below shows a red circle for 'Myoblast', a blue circle for 'Myotube', and a yellow Y-shape for 'Myotube'. A large black arrow points from the myoblast to the myotube. The bottom part shows a myotube with a green and white striped pattern, and a tenocyte (red) with a blue nucleus. A legend below shows a red star for 'Tenocyte', a yellow Y-shape for 'Myotube', a blue Y-shape for 'Myotube', and a yellow Y-shape for 'Myotube'. A legend below shows a red star for 'Tenocyte', a yellow Y-shape for 'Myotube', a blue Y-shape for 'Myotube', and a yellow Y-shape for 'Myotube'. A legend below shows a red star for 'Tenocyte', a yellow Y-shape for 'Myotube', a blue Y-shape for 'Myotube', and a yellow Y-shape for 'Myotube'.

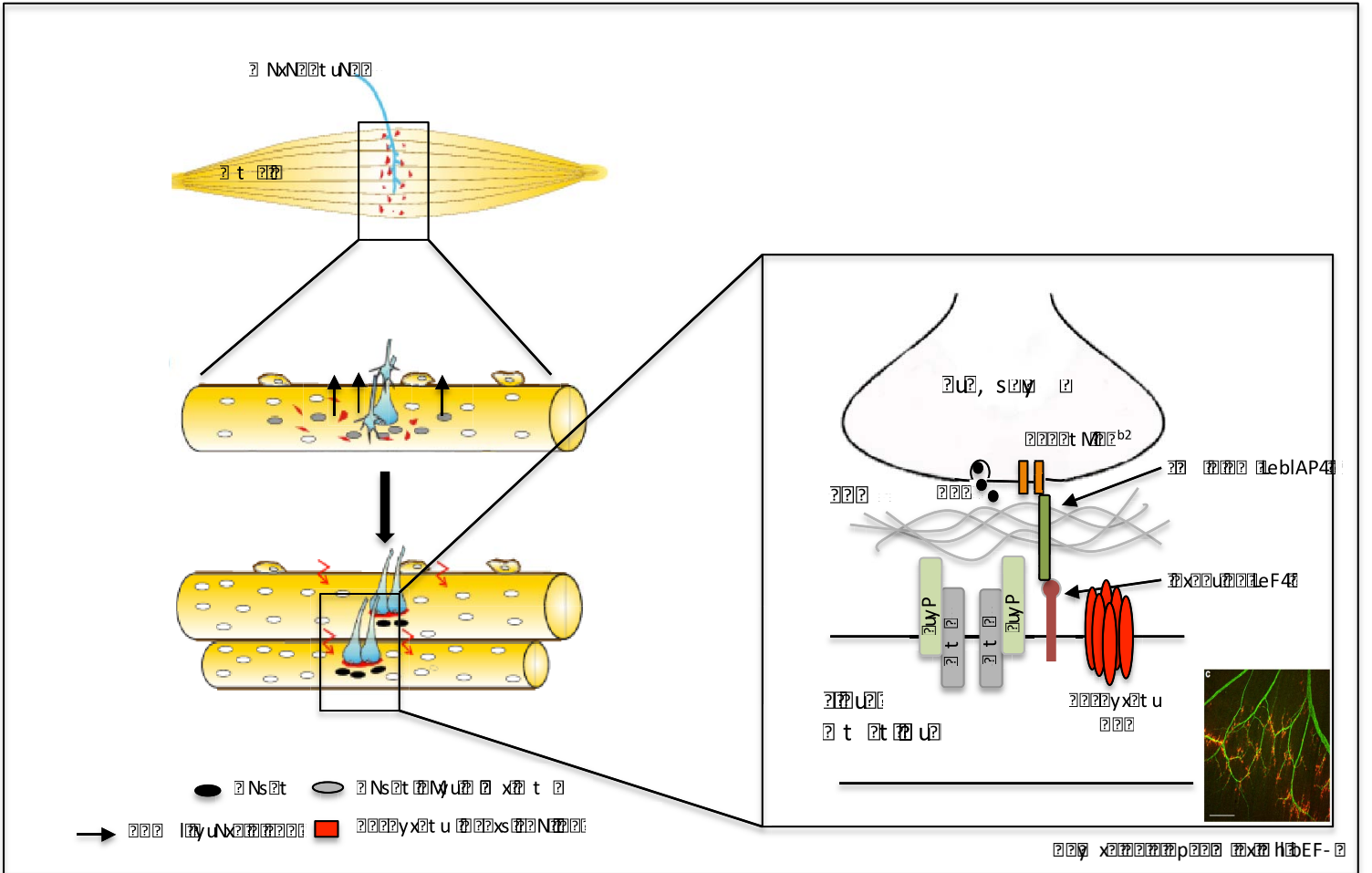
Dans la phase initiale, indépendante des tendons chez les vertébrés (chez le poisson zèbre par exemple), les myoblastes synthétisent une ECM « pré-tendon » qui inclue les ligands d'intégrines Tsp4, Lama2 et la Fibronectine (Fn) (Carmignac and Durbeej, 2012; Snow et al., 2008; Subramanian and Schilling, 2014). Dans les muscles de mammifères, Fn et Lama2 sont aussi nécessaires à la migration et l'attachement (Vaz et al., 2012). Fn et Lama2 se lient aux intégrines et au complexe Dystrophine/Dystroglycan à la surface des cellules musculaires. Ensuite, la mécanotransduction couplée au TGF- $\beta$  mène à l'expression de *Scx* et *Mkx* dans les ténocytes qui à leur tour induisent l'expression sélective de gènes de l'ECM des tendons (Figure 10A). De plus, *Mkx* inhibe l'activité de facteurs myogéniques pendant la différenciation des ténocytes.

La phase tardive, dépendante des tendons, permet la production de protéines de l'ECM et plus particulièrement des différents types de collagène (Eliasson et al., 2008; Maeda et al., 2011) qui permettent le maintien et la robustesse des jonctions myo-tendineuses. Pour finir, les protéines Talin 1, Talin 2 et les récepteurs aux intégrines  $\alpha7\beta1D$  et  $\alpha7B\beta1D$  sont nécessaires à la bonne formation des jonctions myo-tendineuses chez les vertébrés (Conti et al., 2008, 2009) (Figure 10B).

#### 4. Innervation et jonction neuromusculaire

L'innervation des muscles commence en même temps que la formation des fibres. La formation de la jonction neuromusculaire (JNM) est initiée au cours du développement embryonnaire (chez les rongeurs entre les stades E12 et E13.5), précédée par la guidance des axones vers les fibres musculaires qui doivent être innervées, impliquant des signaux moléculaires précis pour assurer le parfait contact muscle-nerf (Sanes and Lichtman, 1999; Wu et al., 2010). Par la suite, une période de maturation permet de conduire à des changements fonctionnels et morphologiques d'éléments pré et post-synaptiques. La mise en place primaire ainsi que la guidance des neurones se font très précocement *via* un pool de neurones dans la corde spinale, qui forment les colonnes motrices (au stade E6-7 chez l'embryon de poulet)(Landmesser, 1978). Ces colonnes sont des groupes de motoneurones qui innervent de manière sélective un ensemble de muscles et peuvent être classifiées selon leur localisation dans la corde spinale (Landmesser, 2001). Cette organisation en colonne semble être cruciale pour l'établissement du bon ciblage de chaque classe de motoneurones





The diagram illustrates the molecular mechanism of a plant defense response. It shows a cross-section of a plant stem with a central vascular cylinder. A blue arrow labeled "Nxn2t uN2t" points to the center. The middle panel shows a longitudinal section of the stem with a blue structure and red arrows indicating movement. The bottom panel shows a longitudinal section with a blue structure and red arrows. A legend below the diagram identifies symbols: a black circle for "Ns2t", a grey oval for "Ns2t M2u2t x2t", and a red square for "y2t u 2xs2tN2t". A right-hand inset shows a detailed view of a cell wall with a blue structure labeled "N2t M2t b2", a green structure labeled "LebAP4", and a red structure labeled "E4". Below this inset are labels "yP", "t", "yP", and "y2t u". A small fluorescence microscopy image is in the bottom right of the inset. The text "x2t p2t x2t H2EF-" is at the bottom right of the inset.

à leurs muscles spécifiques afin de contrôler le comportement locomoteur très complexe (Tripodi and Arber, 2012). L'organisation très structurée suggère que des mécanismes très élaborés sont requis pour guider le bon groupe de motoneurones vers leurs cibles finales. Il est maintenant reconnu que cette mise en place est contrôlée par des molécules de guidance (Phenail and Holliday, 1990 ; Lavigne and Goodman, 1996). L'identité des molécules de reconnaissance muscle-nerf reste encore largement méconnue. Cependant, la topographie neuromusculaire perturbée de souris mutées pour les *Ephrins A2/A5* ou sur-exprimant *Ephrin A5* dans le muscle squelettique montre l'implication du système *EphA* dans le processus de contact muscle-nerf (Feng et al., 2000; Lampa et al., 2004). Grâce à différents modèles d'étude comme le poisson-zèbre et la drosophile, un nombre important d'acteurs impliqués dans la reconnaissance muscle-nerf a été découvert comme les protéines LRR (Leucine Reach Repeat), les CAMs (Cell Adhesion Molecules), des membres de la famille des immunoglobulines et des cadhérines, mais aussi des molécules d'attraction ou de répulsion telles que les nétrines et les sémaphorines (Sanes and Yamagata, 2009). La fonction de ces différentes protéines dans la mise en place du contact muscle-nerf reste à valider chez les mammifères (Bonanomi and Pfaff, 2010). Les laminines synaptiques, telles que les Laminins  $\beta 2$ ,  $\alpha 4$  et  $\alpha 5$  sont relarguées par les fibres musculaires. Elles forment alors un complexe glycoprotéique hétérotrimérique, inclus dans la lame basale et sont nécessaires avec le complexe musculaire membranaire Musk (récepteur tyrosine kinase musculaire)/Lrp4 (récepteur lipoprotéique 4) pour la maturation pré et post-synaptique. (Darabid et al., 2014) (Figure 11).

Bien qu'une partie des mécanismes moléculaires et génétiques de spécification et de différenciation des myoblastes au cours de l'embryogenèse chez les vertébrés a été découverte, le contrôle de l'acquisition des propriétés individuelles (taille, forme, attachement et innervation) des différents muscles du tronc des membres ou de la face reste encore mal détaillé.

Afin d'avoir une meilleure compréhension des mécanismes de diversification musculaire, il est nécessaire d'avoir une vision cellule-spécifique, dans le but d'identifier de nouveaux gènes qui participent à l'acquisition de l'identité musculaire individuelle au cours de l'embryogenèse. Pour cela, nous avons utilisé comme modèle d'étude la drosophile, dont les mécanismes et les facteurs-clés myogéniques sont conservés.



## C – La drosophile comme modèle d'étude de la diversification musculaire

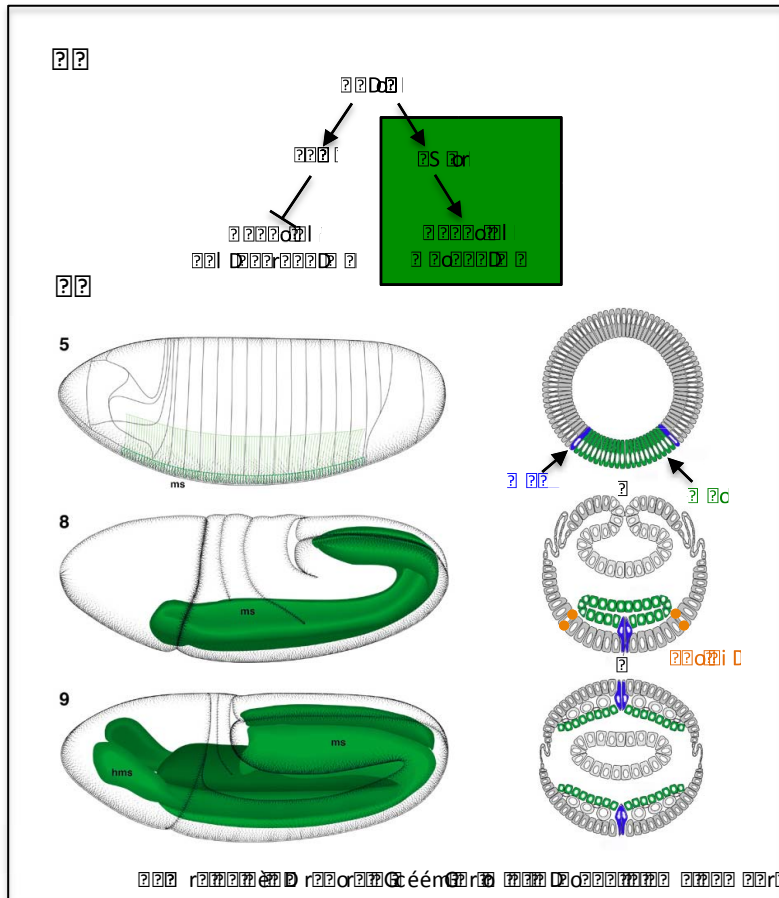
### I. La drosophile

La drosophile et en particulier *Drosophila melanogaster*, aussi appelée mouche de vinaigre, est un insecte très étudié depuis plus d'un siècle. En 1908 Thomas Hunt Morgan étudiait déjà cette espèce afin de mieux comprendre la génétique et l'hérédité. Il obtiendra le prix Nobel pour ses recherches en 1933. La drosophile est un diptère holométabole dont le cycle de développement est de 10 jours environ à 25°C. Les différentes étapes de vie sont l'embryon, la larve, la pupa puis enfin la mouche adulte. La rapidité du développement, l'entretien peu coûteux, mais surtout la puissance des outils génétiques et une conservation extensive de son génome avec celui de l'homme fait de la drosophile un modèle d'étude très efficace et pertinent.

### II. Généralités sur le muscle somatique

La myogenèse est un processus développemental clé qui permet de mettre en place l'ensemble de la musculature d'un organisme. Chez la drosophile il y a deux vagues de myogenèse. La première, myogenèse embryonnaire, permet la mise en place de la musculature larvaire et la deuxième au cours de la métamorphose aboutit à la formation des muscles adultes (Figure 12). En accord avec mon projet je présenterai seulement les différentes étapes de la myogenèse embryonnaire des muscles somatiques.

Les muscles somatiques qui constituent la paroi musculaire larvaire assurent la locomotion de la larve. Chacun de ces muscles est composé d'une seule fibre multinucléée ce qui est la différence principale avec les muscles vertébrés qui sont multifibrillaires. L'embryon présente une structure segmentée et dans chaque hémisegment on retrouve un patron stéréotypé d'environ 30 muscles. Les segments abdominaux A1-A7 présentent un patron musculaire similaire alors que le A8 et les segments thoraciques T1-T3 ont un patron spécifique (Bate, 1990). Le développement de chaque muscle est sous contrôle d'un programme d'identité qui lui confère des caractéristiques individuelles de taille, forme, orientation, attachement et innervation facilement identifiables (Baylies et al., 1998) (Figure 13).



The diagram illustrates the developmental stages of the mesoderm in a fly embryo. At the top, a gene expression tree shows the relationship between *Dfd*, *E*, *S*, *P*, and *C*. Below this, three stages of an embryo are shown: stage 5 with a green mesoderm (ms) at the posterior; stage 8 with the mesoderm (ms) beginning to fold; and stage 9 with the formation of the heart mesoderm (hms) and the gut tube. To the right, a circular diagram illustrates the folding of the mesoderm, with labels 'P' and 'C' indicating different regions.

### III. Le mésoderme

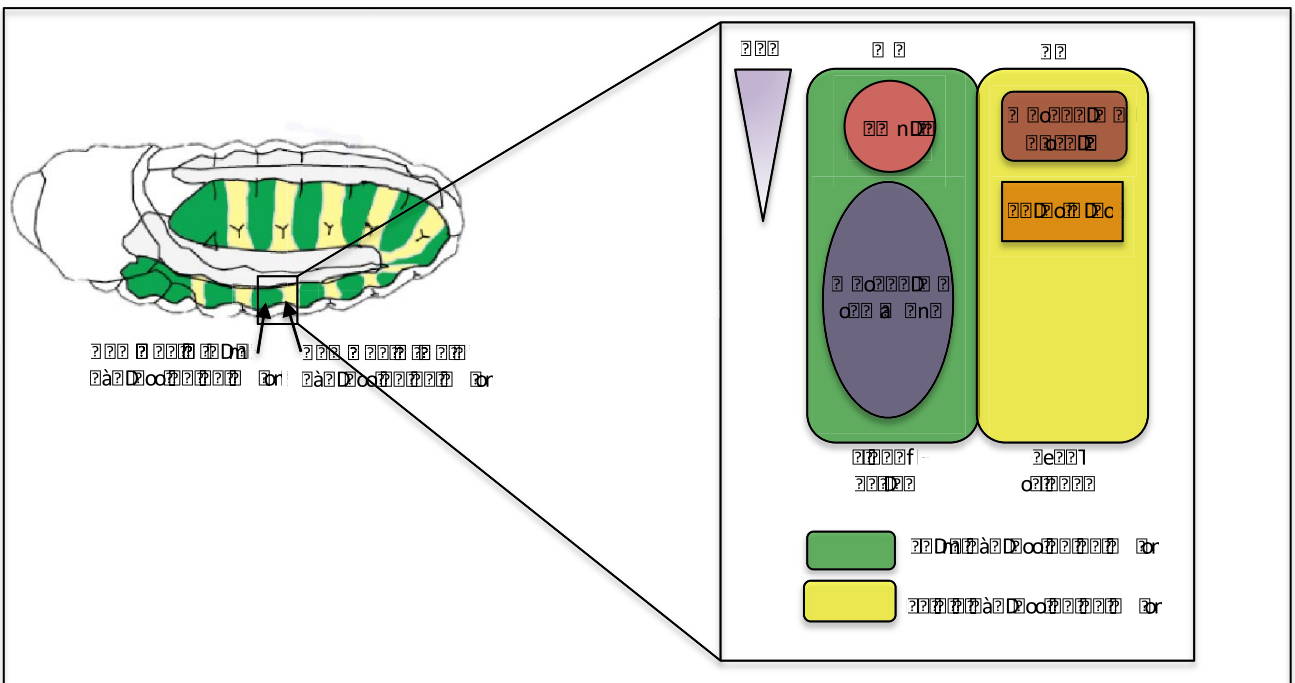
#### 1. Origine et détermination du mésoderme

Les muscles somatiques trouvent leur origine au sein du mésoderme. Le mésoderme est un feuillet embryonnaire formé au cours de la gastrulation. Celui-ci est spécifié en position ventrale de l'embryon au stade blastoderme. Cette étape se fait sous le contrôle du morphogène maternel Dorsal (NF- $\kappa$ B chez les vertébrés) au stade 5. Dorsal induit l'expression de deux facteurs de transcription requis pour la formation du mésoderme : Twist (Twi) et Snail (Sna). Twi est nécessaire pour l'activation d'un grand nombre de gènes impliqués dans la mise en place du programme myogénique tels que *Dmef2*, *Tin* et *Poxm* alors que Sna lui réprime des gènes neurogéniques dans le mésoderme (Cripps et al., 1998; Ip et al., 1992) (Figure 14A).

Ces deux facteurs sont nécessaires pour la migration des cellules mésodermiques au cours de la gastrulation. En effet, les cellules exprimant Twi et Sna changent de forme et s'invaginent dans le sillon ventral, se divisent et subissent une transition épithélio-mésenchymateuse (Groshans and Wieschaus, 2000; Leptin and Grunewald, 1990). Les cellules mésodermiques se répandent ensuite dorsalement le long de la partie intérieure de l'embryon. A la fin elles forment une monocouche dans la partie sous-jacente de l'ectoderme (Leptin and Grunewald, 1990). Ce processus est contrôlé par l'expression du récepteur FGF Heartless (Htl) au niveau du mésoderme et de ses ligands Thisbee et Pyramus au niveau du neurectoderme (Shishido et al., 1997; Stathopoulos et al., 2004) (Figure 14B). Par la suite le mésoderme va être régionalisé pour donner naissance aux différents dérivés mésodermiques.

#### 2. Subdivision du mésoderme

Une fois le mésoderme formé celui-ci est subdivisé dans chaque segment en quatre domaines distincts. Cette étape correspond à la régionalisation des somites chez les vertébrés. Le mésoderme est défini selon l'axe antéro-postérieur (A/P) et dorso-ventral (D/V). Cette régionalisation se fait sous l'influence de différents signaux provenant de l'ectoderme : Decapentaplegic (Dpp) (homologue de BMP), Wingless (Wg) (homologue de Wnt) et Hedgehog (Hh) (homologue de Shh). Cette étape permet de définir les différents dérivés mésodermiques à l'origine des muscles somatiques, viscéraux, du cœur et du corps gras (Reuter and Leptin, 1994; Bate, 1993). La limite antéro-postérieure de chaque segment



The diagram shows a cross-section of a stem with vascular bundles. The bundles are arranged in a ring. Each bundle consists of xylem (inner part, green) and phloem (outer part, yellow). The xylem is responsible for water transport, and the phloem is responsible for nutrient transport. The diagram also shows the cambium, which is the tissue between the xylem and phloem that produces secondary xylem and secondary phloem.

est déterminée par les signaux Wg en antérieur et Hh en postérieur. Wg et Hh contrôlent respectivement l'expression de Sloppy paired (Slp) et Even-skipped (Eve) qui par la suite permettent l'expression de gènes mésodermiques (Azpiazu et al., 1996; Riechmann et al., 1997). Wg permet l'activation de Slp dans la partie antérieure ce qui maintient l'expression de Twi. En comparaison Eve agit à travers la signalisation Notch pour réprimer Twi dans la partie postérieure de l'hémisegment. Cette régulation se fait *via* Suppressor of Hairless (Su(H)) qui est un effecteur transcriptionnel de la voie Notch. Ces deux mécanismes permettent de partitionner l'expression de Twi en deux domaines, un à forte expression et un à faible expression de Twi (Baylies et al., 1998; Frasch, 1999).

La combinaison de Wg, Slp et Twi permet la formation des muscles somatiques et du cœur en induisant l'expression de gènes du mésoderme somatique tels que *muscle segment homeobox (msh)*, *slouch (slou/S59)* ou encore *ladybird (lb)* et en réprimant les gènes du mésoderme viscéral (D'Alessio and Frasch, 1996; Jagla et al., 1998; Knirr et al., 1999). Dans la partie postérieure la combinaison de Hh et de la faible expression de *twi* induit l'activation de *bagpipe (bap)* et *serpent (srp)* permettant ainsi la mise en place de la musculature viscérale et du corps gras (Azpiazu et al., 1996) (Figure 15).

La division dorso-ventrale du mésoderme est contrôlée par l'expression du morphogène Dpp qui provient de l'ectoderme dorsal (Frasch, 1995 ; Staehling et al., 1994). Dpp a comme cible le facteur de transcription Tinman (Tin). Celui-ci est tout d'abord exprimé dans tout le mésoderme sous le contrôle de Twi avant d'être maintenu par Dpp dans la partie dorsale uniquement. (Frasch, 1995). Dpp restreint en même temps l'expression des facteurs de transcription dSix4 (*Drosophila Six4*) et Pox-meso à la partie latérale et ventrale du mésoderme (Clark et al., 2006; Duan et al., 2007). L'activation de ces derniers permet de réguler l'expression de gènes spécifiques tels que *slou* ou *lb* à l'origine de certains muscles latéraux et ventraux (Clark et al., 2006; Duan et al., 2007; Sandmann et al., 2007).

La spécification du mésoderme somatique est donc sous le contrôle de facteurs myogéniques clés tels que Twi et Slp auxquelles s'ajoutent Tin, dSix4 et Pox-meso en dorsal, latérale et ventrale respectivement. La perte de fonction d'un de ces facteurs entraîne soit l'absence totale de muscles comme dans les mutants *Slp* soit n'affecte qu'une partie des muscles (mutants *pox-meso*, *dSix4* et *tin*) en fonction du domaine d'expression de ces facteurs au sein de chaque hémisegment (Bodmer, 1993; Duan et al., 2007; Riechmann et al., 1997).





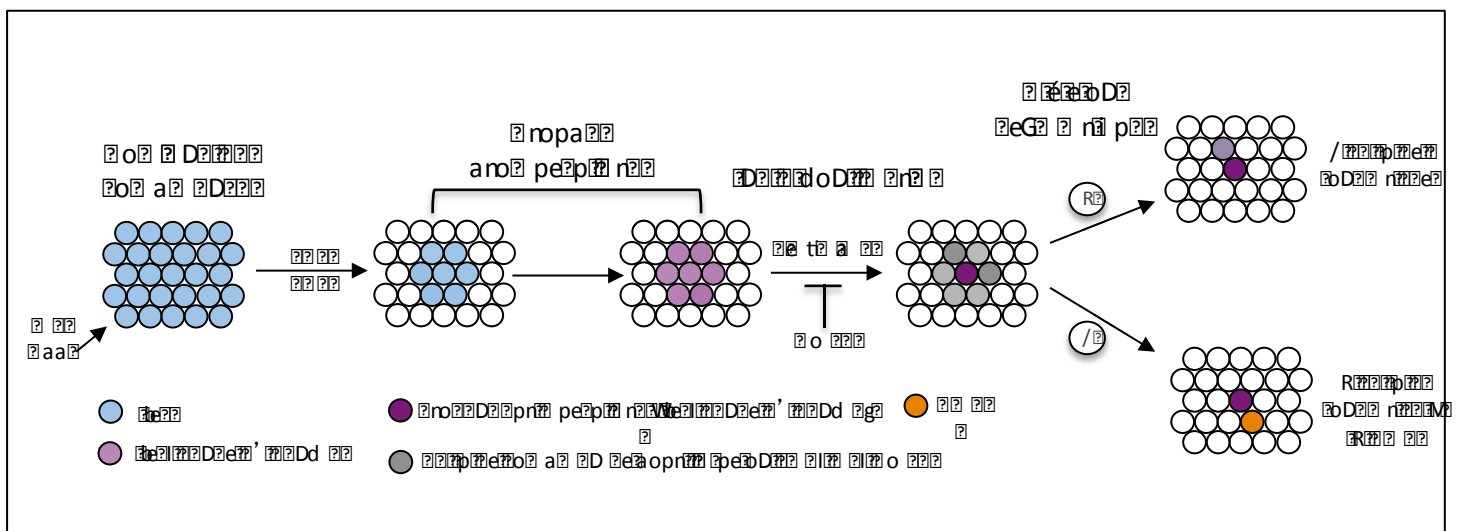
## IV. Etapes de formation du muscle

Après la subdivision et la spécification du mésoderme somatique plusieurs étapes clés sont nécessaires à la formation des muscles. Il y a tout d'abord formation de groupes (clusters) promusculaires qui sont des groupes de cellules équivalentes puis sélection d'un progéniteur (cellule progénitrice, CP) au sein de chaque groupe. Chaque cellule progénitrice va ensuite subir une division asymétrique menant à la formation soit de deux cellules appelées « fondatrices » soit d'une cellule fondatrice et un précurseur de muscle adulte (AMP)(Bate, 1990). Les cellules non sélectionnées dans les différents clusters deviennent des myoblastes dit « naïfs » appelés aussi myoblastes compétents pour la fusion (FCM). Par la suite il y aura fusion des myoblastes pour générer les myofibres multinuclées, puis attachement aux cellules tendons et enfin innervation par les motoneurones pour donner la musculature fonctionnelle larvaire. Toutes ces étapes vont être décrites par la suite en suivant l'ordre chronologique.

### 1. Origine des cellules fondatrices et des FCMs

Les cellules progénitrices à l'origine des cellules fondatrices et des FCM proviennent d'un groupe de cellules équivalentes retrouvées dans le mésoderme somatique.

Tout d'abord au sein du mésoderme somatique la signalisation Wg ou l'intersection Wg/Dpp dans la partie dorsale permet de définir plusieurs régions qui sont compétentes pour répondre à la signalisation Ras/MAPK (Mitogen Activated Potein Kinase). Les cellules composant ces régions ont été identifiées grâce à l'expression du gène proneural *lethal of scute (l'sc)* et sont appelées « pre-cluster » (Carmena et al., 1995). En effet la signalisation Wg (Wg/Dpp dans le mésoderme dorsal) permet tout d'abord l'expression de *l'sc* dans les cellules présentes au sein des domaines de compétence. Ensuite l'activation de la voie Ras/MAPK dans ces cellules par deux récepteurs tyrosine kinase DER (EGF) et Heartless (Htl, FGF) restreint l'expression de *l'sc* à des sous groupes de cellules équivalentes appelés groupes promusculaires. (Buff et al., 1998; Michelson et al., 1998). Il existe 19 groupes promusculaires qui ont été cartographiés et numérotés en fonction de leur position dorso-ventrale. A ce stade certaines cellules commencent à exprimer des facteurs de transcription appelés aussi facteurs d'identité (iTFs) nécessaires au contrôle de la diversité des muscles (Carmena et al., 1998a). Les cellules de chaque groupe d'équivalence ont toutes le potentiel pour devenir une cellule progénitrice. La sélection du progéniteur se fait grâce au



The diagram illustrates the evolution of a lattice structure. It starts with a 4x4 lattice of blue particles. An arrow labeled 'aa' points to a 4x4 lattice with one blue and three white particles. A second arrow points to a 4x4 lattice with two blue and two white particles. A third arrow points to a 4x4 lattice with three purple and one white particle. A fourth arrow points to a 4x4 lattice with one purple, one grey, and two white particles. From this lattice, two arrows branch out: one labeled 'R' points to a 4x4 lattice with one purple and three white particles, and another labeled 'I' points to a 4x4 lattice with one purple, one orange, and two white particles. A legend below identifies the colors: blue for 'a', purple for 'b', grey for 'c', and orange for 'd'. Labels above the lattice stages include 'ano', 'pep', 'n', 'D', 'G', 'r', 'p'.

mécanisme d'inhibition latérale médié par la voie Notch. En effet le maintien de la voie Ras/MAPK dans la cellule progénitrice induit la surexpression du ligand Delta ce qui provoque une activation de la voie Notch et une répression de la voie Ras dans les cellules avoisinantes (Halfon et al., 2000). Dans un mutant *delta*, toutes les cellules d'un groupe promusculaire maintiennent une forte expression de *I'sc*, indiquant que toutes les cellules sont compétentes pour devenir des progéniteurs empêchant ainsi la sélection d'un seul d'entre eux (Carmena et al., 1995). Par la suite seul le progéniteur sélectionné continue d'exprimer *I'sc*. Les cellules non sélectionnées peuvent par la suite soit rejoindre un autre groupe promusculaire soit se différencier en FCM et exprimer des facteurs spécifiques tels que *Lame duck* (*Lmd*) et *Tramtrack69* (*Ttk*) (Ciglar et al., 2014; Duan et al., 2001) (Figure 16).

Les embryons mutants pour *I'sc* présentent seulement une perte de quelques muscles suggérant que d'autres gènes peuvent compenser sa perte. En effet la perte de fonction de *I'sc* peut être sauvée dans les domaines de compétences ventraux par l'expression de *pox-meso* (Duan et al., 2007).

Je reviendrai plus en détail sur les facteurs de transcription exprimés dans les cellules fondatrices afin de mieux appréhender leurs fonctions durant l'acquisition de l'identité musculaire.

## 2. Spécification des cellules fondatrices : division asymétrique

Une fois qu'un progéniteur a été sélectionné au sein des différents groupes promusculaires celui ci va subir une division asymétrique pour donner naissance à deux cellules fondatrices (FCs) ou une FC et une AMP nécessaire à la myogenèse adulte (Carmena et al., 1998a; Ruiz and Bate, 1997). Le progéniteur musculaire peut également se diviser dans la partie dorsale en un précurseur cardiaque ou péri-cardiaque et une cellule fondatrice. Différentes protéines connues pour avoir une fonction cruciale dans la division asymétrique des cellules neurales sont aussi impliquées dans la division asymétrique des progéniteurs musculaires. En effet *Inscutable* (*Insc*), une protéine associée à la membrane localisée au cortex apical et *Numb* (*Nb*), une protéine adaptatrice du cytosquelette située au cortex basal sont nécessaires (Carmena et al., 1998b). *Insc* est impliqué dans l'orientation du fuseau mitotique et dans la localisation de *Numb* au cortex apical. A la suite de la division, *Numb* se retrouve donc dans une des deux cellules seulement, présentant par la suite des signalisations opposées. En effet, les cellules *Nb* positives répriment la signalisation Notch

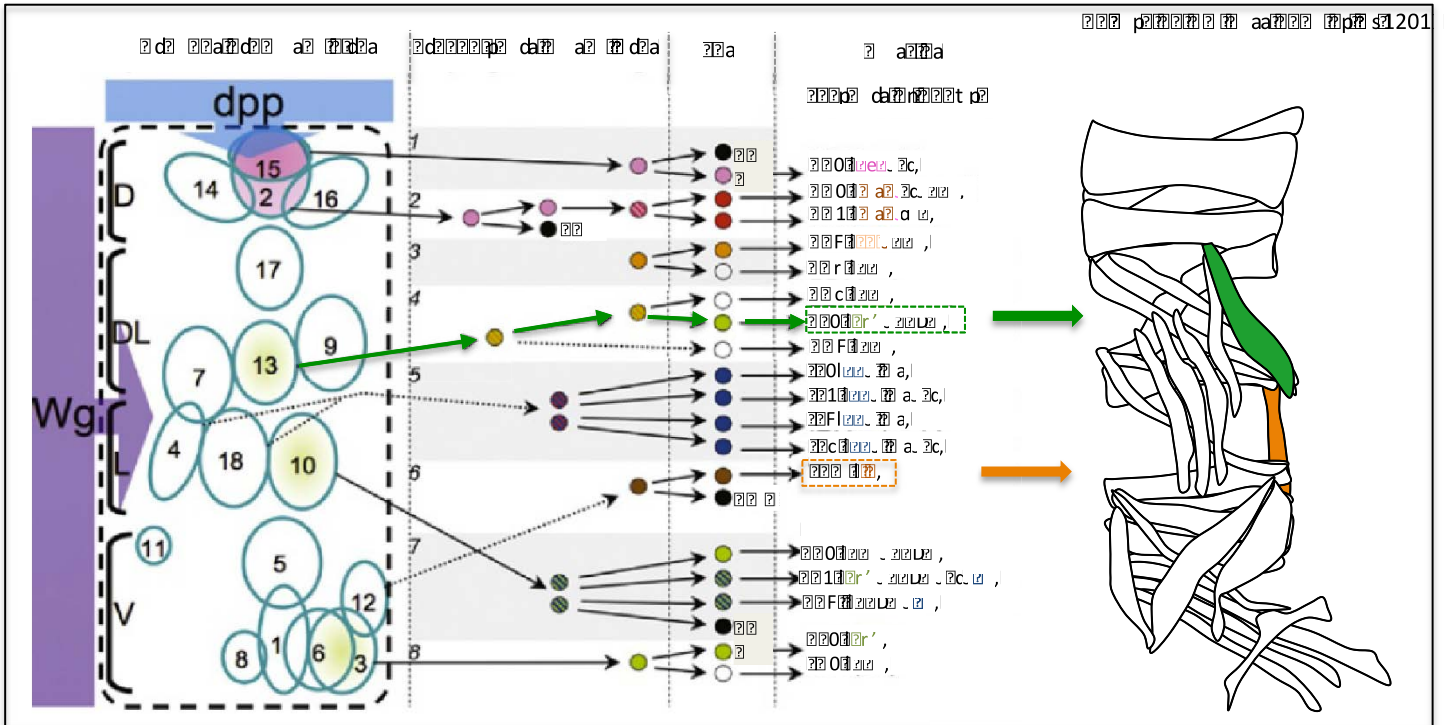


Diagram illustrating gene expression patterns in a fly embryo. The embryo is divided into compartments D, DL, L, and V. The **dpp** gene is expressed in the dorsal compartment (D), and the **Wg** gene is expressed in the ventral compartment (V). The diagram shows the expression of various genes in the 18 numbered cells, with arrows indicating the direction of gene expression. The legend at the bottom right shows the color-coded key for gene expression.

alors que celle-ci reste active dans les cellules Nb négatives. Un modèle propose que Sanpodo (protéine transmembranaire), un interacteur physique de Notch soit régulé par Nb. Nb empêche la localisation de Sanpodo (Spdo) à la membrane inhibant ainsi sa liaison à Notch et donc l'activation de ce dernier (O'Connor-Giles and Skeath, 2003; Park et al., 1998). Cette asymétrie procure donc des cellules fondatrices qui expriment différents iTFs à l'origine de la diversité musculaire. En effet, certains gènes d'identité présents au préalable dans le progéniteur se retrouvent maintenus seulement dans une des deux FCs. Par exemple, le facteur Slou est exprimé dans le progéniteur de muscles VA1 et VA2 puis se retrouve après la division seulement maintenu dans le VA2 (Dohrmann et al., 1990). Au contraire pour Ap qui reste présent de façon continue dans les deux FCs à l'origine des muscles latéraux (Bourgoin et al., 1992).

En fonction de la position des FCs selon l'axe D/V on distingue des FCs dorsales, dorso-latérales, latérales et ventrales qui expriment une combinaison spécifique de iTFs (Figure 17). Concernant les cellules AMPs, celles-ci maintiennent la voie Notch activée et restent quiescentes et indifférenciées jusqu'à leur réactivation au stade larvaire. Cet état est régulé *via* la persistance d'expression du facteur Twi (Anant et al., 1998; Figeac et al., 2007). Au final 30 FCs permettent le développement de 30 muscles individuels dans chaque hémisegment abdominal.

Concernant les FCMs, ces cellules vont exprimer leurs propres facteurs développementaux. En effet, les FCMs semblent présenter aussi des différences d'identité suggérant que ces derniers ne forment pas une population uniforme de cellules comme précédemment décrit (Artero, 2003; Duan et al., 2001; Estrada et al., 2006). Cette différence d'identité est principalement basée sur l'identification de gènes exprimés dans des sous-populations de FCMs (Artero et al., 2001; Estrada et al., 2006; Ruiz-Gómez et al., 2002). De plus, il a été montré qu'un autre groupe de FCMs localisé de manière répétée dans chaque hémisegment subissait une division cellulaire supplémentaire au stade 12-13 (Beckett and Baylies, 2007). Cela suggère que comme pour les FCs, la position des FCMs joue un rôle important dans l'acquisition de leur identité. Au niveau transcriptionnel deux facteurs clés de la diversification des FCMs ont été identifiés, Lmd et Ttk (Ciglar et al., 2014; Duan et al., 2001). Lmd fait partie de la superfamille de facteurs de transcription à doigt de zinc GLI et est exprimé dans les FCMs de la fin du stade 11 jusqu'au stade 14. Ttk qui possède également un domaine de liaison à l'ADN à doigt de zinc s'exprime dans le mésoderme



viscéral (stade 11-16) et transitoirement dans les FCMs au stade 11 (Ciglar et al., 2014; Ruiz-Gómez et al., 2002). La perte de *lmd* ou *Ttk* entraîne des défauts de fusion sévères. En effet, la majorité des myoblastes restent monoclués. Des expériences de CHIP ont été réalisées sur ces deux facteurs afin d'identifier leurs cibles. Il s'avère que *Lmd* favorisent l'expression de gènes de fusion et de différenciation terminale tels que *Mef2*, *stick and stones (sns)* et *hibris (hb)* mais aussi celle de *ttk* (Cunha et al., 2010). *Ttk* lui est connu pour être un répresseur transcriptionnel et apparemment il exerce ce même rôle au sein des FCMs où il est nécessaire pour réprimer l'expression de gènes spécifiques des FCs tels que *dumbfounded (duf)*, *jumeau* et *phyllopod*. Pour conclure *Lmd* active à la fois les gènes de fusion et *ttk*, qui lui réprime le programme transcriptionnel de FCs permettant au programme des FCMs de s'engager (Busser et al., 2012; Ciglar et al., 2014). Une fois toutes ces cellules spécifiées le processus de fusion s'engage.

### 3. La fusion des myoblastes et le positionnement des noyaux

Les 30 fibres musculaires sont produites à partir de la fusion d'une FC avec un nombre précis de FCMs et cela est contrôlé par les facteurs d'identité. (Bataillé et al., 2010). La fusion des myoblastes se déroule sur une période de temps de 5,5h c'est à dire entre le stade 12 et 15. La taille du muscle est déterminé par le nombre d'évènements de fusion qui vont de 2 à 25 en fonction du muscle ciblé (Bate, 1990). Bien qu'il y ait quelques variations de la taille finale du muscle, le nombre moyen de noyaux par muscle a été déterminé (Beckett and Baylies, 2007). L'analyse du profil de fusion de muscle individuel a montré que celui ci est divisé en deux phases. Au cours des 3 premières heures (Stade 12-13), entre 10 et 30% des évènements de fusion ont lieu alors que ceux restant (entre 73 et 91%) se déroulent dans les 2.5h suivantes (Beckett and Baylies, 2007; Richardson et al., 2008). La fusion primaire engage que deux ou trois FCMs maximum alors que la seconde tout les FCMs restant. Le processus de fusion se déroule en plusieurs étapes. Il y a tout d'abord migration des cellules FCMs vers les FCs puis reconnaissance et adhésion des cellules entre elles via des molécules d'adhésion. Des vésicules puis des plaques de préfusion visualisables en microscopie électronique sont formées à l'interface des deux cellules. Par la suite, les FCMs envahissent les FCs grâce aux protrusions membranaires produites à partir du réarrangement de leur cytosquelette d'actine. Lorsque les deux cellules sont très proches les membranes lipidiques sont déstabilisées permettant la création de pore de fusion. Par la





suite le FCM est endocyté pour former une seule cellule multinucléée (Kim et al., 2015a) (Figure 18).

### 3.1 Reconnaissance et adhésion FC/FCM

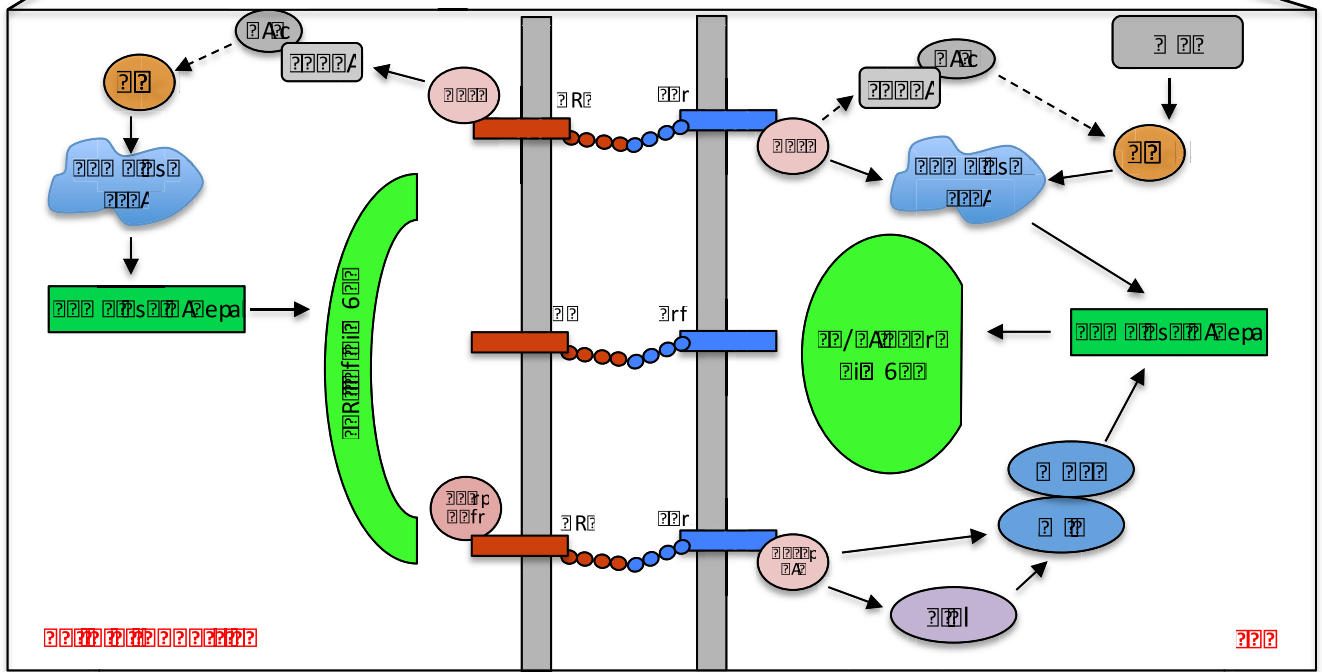
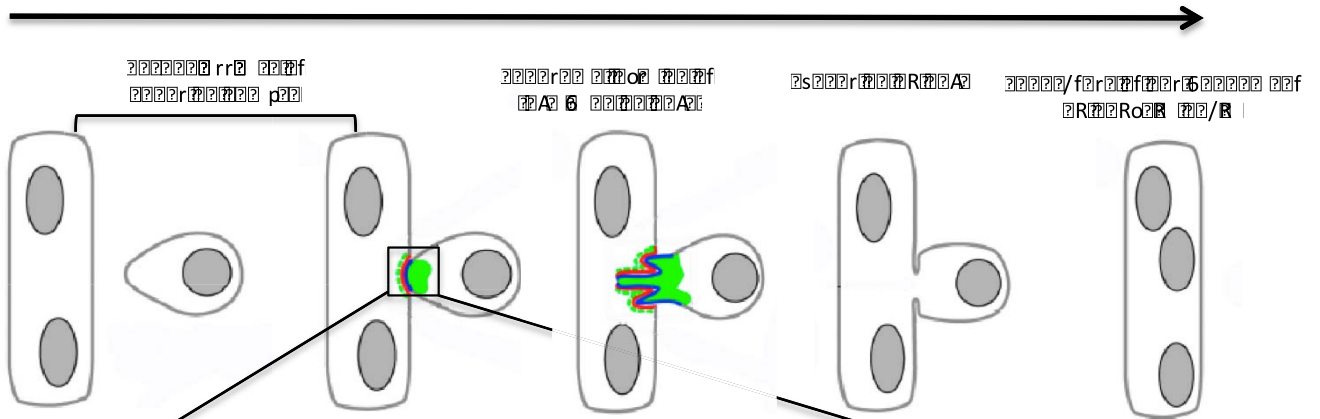
La migration des FCMs vers les FCs se fait via la formation de structures appelées « filopodes » qui permettent aux FCMs d'explorer leur environnement et d'identifier leur FC. Cette migration est dirigée aussi par l'expression de molécules d'adhésion par le FC. Ces molécules appartiennent à la famille des immunoglobulines transmembranaires (IgSF) et sont exprimées de manière dissymétrique entre le FC et le FCM. En effet, les FCs expriment Dumbfounded/Kin of IrreC (Duf/Kirre) et Roughest (Rst) alors que les FCMs, Sns et Hb (Artero et al., 2001; Bour et al., 2000; Ruiz-Gómez et al., 2000; Strünkelberg et al., 2001). Toutes ces protéines ont de larges domaines extracellulaires facilitant leur adhésion.

L'adhésion de ces deux types de cellules se fait via des interactions hétérophiliques entre Duf ou Rst et Sns. La perte de Sns entraîne des défauts majeurs de fusion qui ne peuvent être que partiellement sauvés par son paralogue Hb (Bour et al., 2000). L'expression ectopique de Rst ou Duf dans des cellules épithéliales permet l'attraction des FCMs vers ces dernières montrant ainsi que Duf ou Rst est suffisant pour attirer les cellules exprimant Sns. La perte simultanée de Duf et Rst abolit complètement la fusion. Cela est partiellement restauré lorsque l'un des deux gènes est ré-exprimé (Strünkelberg et al., 2001) Une fois que l'interaction entre ces protéines est établie il y a enrichissement de la protéine PI(4,5)P2 (phosphoinositide) au site de contact FC/FCM. Cette signalisation entraîne un remodelage du cytosquelette d'actine au niveau du site de fusion (Bothe et al., 2014).

### 3.2 Transduction du signal de fusion

Le lien entre les protéines transmembranaires et le cytosquelette est primordial pour transduire le signal de fusion. Pour cela plusieurs types de protéines sont nécessaires. En effet dans les FCs le domaine cytoplasmique de Duf recrute différentes protéines à la membrane telles que Rolling pebbles/Antisocial (Rols/Ants), Dreadlock (Dock) ainsi qu'une protéine à activité GEF, Schizo/Loner (Guanine-nucleotide Exchange Factor)(Chen and Olson, 2001; Chen et al., 2003; Kaipa et al., 2013; Menon and Chia, 2001).

Rols fonctionne à travers une boucle de rétroaction positive qui permet de stabiliser Duf à la membrane des FCs/myotubes. Rols interagit avec Myoblast city (Mbc) dont l'activité



PLP  
 RrRr Aepa rrrrrr Aepa frrrrr Aepa RrRr RrRr/600  
 RrRr Aepa rrrrrr Aepa frrrrr Aepa RrRr RrRr/600  
 RrRr Aepa rrrrrr Aepa frrrrr Aepa RrRr RrRr/600

GEF dépendante de PI(4,5)P2 permet d'activer des GTPases de la famille Rho, Rac1 et Rac2 au site de fusion. Cependant la pertinence de cette interaction n'est pas claire car l'expression de *Mbc* est seulement requise dans les FCMs (Haralalka et al., 2013). Schizo/Loner facilite le recrutement à la membrane de Rac1 par l'intermédiaire de la GTPase Arf6 activant ainsi la voie Scar (Suppressor of cAMP Receptor)(Chen et al., 2003; Ibarra et al., 2005). *Loner* est aussi exprimé dans les FCMs où il exerce probablement la même fonction mais indépendamment de Duf (Richardson et al., 2007). Dans les FCMs le cytodomaine de Sns se lie à deux protéines adaptatrices contenant des domaines SH2-SH3, Dock et Dcrk (Kaipa et al., 2013; Kim et al., 2007). Ces deux adaptateurs recrutent les protéines spécifiques des FCMs, Verprolin1 (Wasp interacting protein) et Blown fuse (Blow) nécessaires à l'activation du complexe Wasp (Wiskott-Aldrich Syndrome protein)(Schäfer et al., 2007) (Figure 18).

Après l'adhésion, les membranes plasmiques s'alignent et des groupes de vésicules (dense en électron) se rangent par paire de part et d'autres des cellules qui s'apposent pour former des complexes de préfusion (ou plaque de préfusion)(Doberstein et al., 1997). La fusion de ces vésicules aux membranes permet de délivrer par exocytose des molécules nécessaires à la fusion comme des lipides, des fusogènes ou des protéases. Cette régulation est contrôlée par Singles-Bar (*Sing*), protéine transmembranaire dans les FCMs et les FC. Dans les embryons mutants pour *Sing* il y a une accumulation de vésicules entraînant des défauts de fusion (Estrada et al., 2007a).

### 3.3 Remodelage du cytosquelette d'actine et formation de pores de fusion

Par la suite, le recrutement de ces différentes protéines induit la formation de foyer (foci) dynamiques d'actine filamenteuse (actine-F) nécessaires à la fusion. Cela se fait notamment à travers le recrutement et l'action des facteurs NPFs (Nucleation Promoting Factors) Scar et WASp sur le complexe Arp2/3 (actin related protein 2/3)(Berger et al., 2008). Ce dernier permet la construction d'un réseau d'actine ramifié en liant les filaments d'actine préexistants et en promouvant la création de nouveaux (Padrick et al., 2011). Dans les FCs l'activation de Scar par Rac1 permet sa fixation au complexe Arp2/3 favorisant la formation d'un feuillet d'actine au site de fusion. Dans les FCMs l'activation du complexe WASp/Verprolin1 et Scar permet de réguler Arp2/3 induisant ainsi la polymérisation de



l'actine (foyer dense) au site de contact cellule-cellule (Kim et al., 2007; Massarwa et al., 2007; Schroter, 2004) (Figure 18).

Il est maintenant clair que les foci ne sont pas symétriquement distribués entre les deux types cellulaires. En effet au niveau des FCs on retrouve une fine couche d'actine alors que dans les FCMs l'accumulation dense d'actine forme des structures ressemblant à des podosomes ayant des capacités invasives (PLS : podosome like structure) (Haralalka et al., 2013; Sens et al., 2010; Shilagardi et al., 2013). Ces structures protrusives poussent les FCMs à l'intérieur des FCs. Jusqu'à récemment, on ne savait pas comment les FCs répondaient à l'invasion par ces podosomes. Il s'avère que le stress mécanique imposé par les PLS induit le recrutement et l'activation au niveau de la région cortical des FCs de la myosine II (MyoII)(Kim et al., 2015b). MyoII agit comme un senseur direct du stress mécanique induit par les PLS. Elle fournit alors une résistance (tension corticale) au niveau de la membrane plasmique des FCs pendant l'invasion, promouvant ainsi le rapprochement des membranes FC/FCM et la formation de pore. En l'absence d'accumulation de MyoII aucun pore de fusion n'est présent (Kim et al., 2015b). La formation de pores de fusion se fait donc à travers des structures invasives (PLS) générées par l'action de Arp2/3 sur l'actine via les facteurs Wasp/Vrp1/Scar dans les FCMs et par des forces de résistance induites par la combinaison de l'actine ramifié et de la myosine II dans les FCs. L'équipe d'Elisabeth Chen pose l'hypothèse que la formation d'un seul pore à l'extrémité d'un projection invasive permet l'agrandissement du pore, dépolymérisation de l'actine et l'endocytose du FCM à l'intérieur de la cellule fondatrice/myotube, contribuant ainsi à un noyau supplémentaire (Sens et al., 2010). Le noyau du FCM nouvellement incorporé adopte par la suite le destin transcriptionnel du myotube/FC. En effet, le suivie de noyaux de FCMs dans les myotubes a montré que ces derniers ont la capacité d'être reprogrammés selon une identité musculaire précise post-fusion. Des hybridations *in situ* d'ARN naissants de FCM et de gènes d'identité ont montré que le reprogrammage est progressif et commence par la répression de gènes spécifiques des FCMs (Bataillé et al., 2017).

### 3.4 Positionnement des noyaux au sein du muscle

Au terme de la fusion chaque myotube contient un nombre défini de noyaux qui adoptent une distribution spécifique afin de permettre le bon fonctionnement du muscle (Folker et al., 2012; Schulman et al., 2014). Les noyaux musculaires subissent plusieurs



mouvements gouvernés par des mécanismes distincts. Ces mécanismes ont été étudiés dans les muscles latéraux transverses (LTs). Après les étapes de fusion les noyaux dans le muscle se retrouvent au sein d'un seul groupe situé proche du pôle ventral de chaque LT. Peu de temps après (stade 14, 10h20-11h20 AEL (After Egg Laying)) le groupe se sépare en deux, un ventral et un dorsal. Au stade 15 (11h20-13h AEL) et 16 (13h-16h AEL) ces groupes migrent vers leurs pôles respectifs. Au stade final 17 les noyaux deviennent uniformément répartis au sein du muscle (Metzger et al., 2012). Ces différents mouvements nucléaires semblent mécaniquement distincts. En effet ces derniers sont contrôlés par plusieurs protéines associées aux microtubules du cytosquelette. Par exemple la protéine Enscin semble nécessaire à la division du premier groupe en un ventral et un dorsal. En effet dans les mutants Enscin ce premier groupe persiste tout au long du développement embryonnaire. Deux autres protéines motrices sont particulièrement importantes, Kinesin et Dynein (Folker et al., 2012; Metzger et al., 2012). Ces dernières ne jouent pas de rôles clés dans la première phase mais sont nécessaires pour que les groupes ventraux et dorsaux migrent correctement à leur pôle respectif. La dynéine et la kinésine agissent ensemble à travers deux mécanismes spatialement séparés pour faire migrer les noyaux dans une direction unique. Lorsque la dynéine est localisée au cortex cellulaire proche du pôle par la kinésine, celle-ci tire les microtubules et le noyau attaché vers le pôle musculaire. Dans un second cas la kinésine exerce une force à l'avant du noyau pendant que la dynéine pousse à l'arrière afin de transférer le noyau vers le pôle (Folker et al., 2014). D'autres acteurs comme les protéines transmembranaires nucléaires KASH et SUN sont aussi impliqués dans le positionnement des noyaux au sein des muscles et cela se fait notamment grâce à leur interaction avec le cytosquelette d'actine (Malone et al., 1999; Starr and Becker, 2002). Par exemple la protéine Kash, Msp300, permet une fois que le noyau est arrivé à sa position finale de sécuriser son emplacement en l'attachant à des composants du cytosquelette d'actine environnant (Elhanany-Tamir et al., 2012)

Le positionnement des noyaux au sein du muscle est très important. Une localisation anormale des noyaux est souvent corrélée avec des pathologies musculaires ou des faiblesses musculaires (Romero, 2010).





### 3.5 Régulation muscle spécifique de la fusion

Bien que l'ensemble des gènes présentés ci-dessus soit nécessaire au bon déroulement des différentes étapes de fusion, une régulation muscle spécifique de ce processus est mise en place pour former des muscles avec un nombre précis de noyaux. Cela se fait à travers l'expression de différents acteurs du cytosquelette d'actine et de l'adhésion cellulaire comme Muscle protein 20 (Mp20), Paxilin (Pax) et M-spondin (Msp0) qui sont régulés entre autres par les facteurs d'identités Eve, Slou et Lb (Bataillé et al., 2010; Junion et al., 2007). Mp20, Msp0 et Pax s'expriment à des niveaux différents dans plusieurs muscles et c'est leur combinaison d'expression qui détermine le nombre d'évènements de fusion qui aura lieu dans chaque muscle. Au cours de la fusion l'intensité d'expression de Pax est augmentée dans seulement quelques muscles ventraux. Au contraire de Pax, Msp0 est un régulateur négatif qui lorsqu'il est absent induit une augmentation du nombre de noyaux dans un muscle dorsal. Pour Mp20 sa surexpression induit le recrutement de deux FCMs au maximum et son absence la perte de deux dans une sous population de muscles. Cette étude a permis de mettre en valeur l'impact des facteurs d'identité sur la différenciation terminale des muscles individuels et notamment *via* le processus de fusion (Bataillé et al., 2010). Cependant le contrôle du processus de fusion par ces trois facteurs n'est vrai que pour un sous ensemble de muscle. En effet Eve, Lb et Slou sont suffisants pour contrôler le nombre de fusions dans le muscle DA1, SBM et VT1, cependant dans le VA2 et le DT1 Slou fonctionne différemment et ne semble pas avoir de rôle décisif dans ce processus. Cela soulève une question importante sur la hiérarchie des gènes d'identité au cours du programme myogénique ainsi que sur leur rôle dans l'acquisition des propriétés individuelles.

### 4. Elongation et attachement aux cellules tendineuses

Par comparaison avec les vertébrés la drosophile présente un exosquelette à la place d'un squelette interne (cartilage et os). Les connections avec le muscle se font à travers des cellules épithéliales appelées cellules tendineuses qui proviennent de l'ectoderme et qui sont les analogues des tendons des vertébrés (Fernandes et al., 1991).

Au cours du processus de fusion, la fibre en incorporant des noyaux supplémentaires s'allonge pour atteindre les cellules tendineuses qui sont des sites d'attachement au sein de l'épiderme. L'allongement des myotubes se fait de manière bidirectionnelle c'est à dire en



direction de sites d'attachement opposés et cela implique la mise en place de filopodes à chaque extrémité des muscles (Schnorrer and Dickson, 2004).

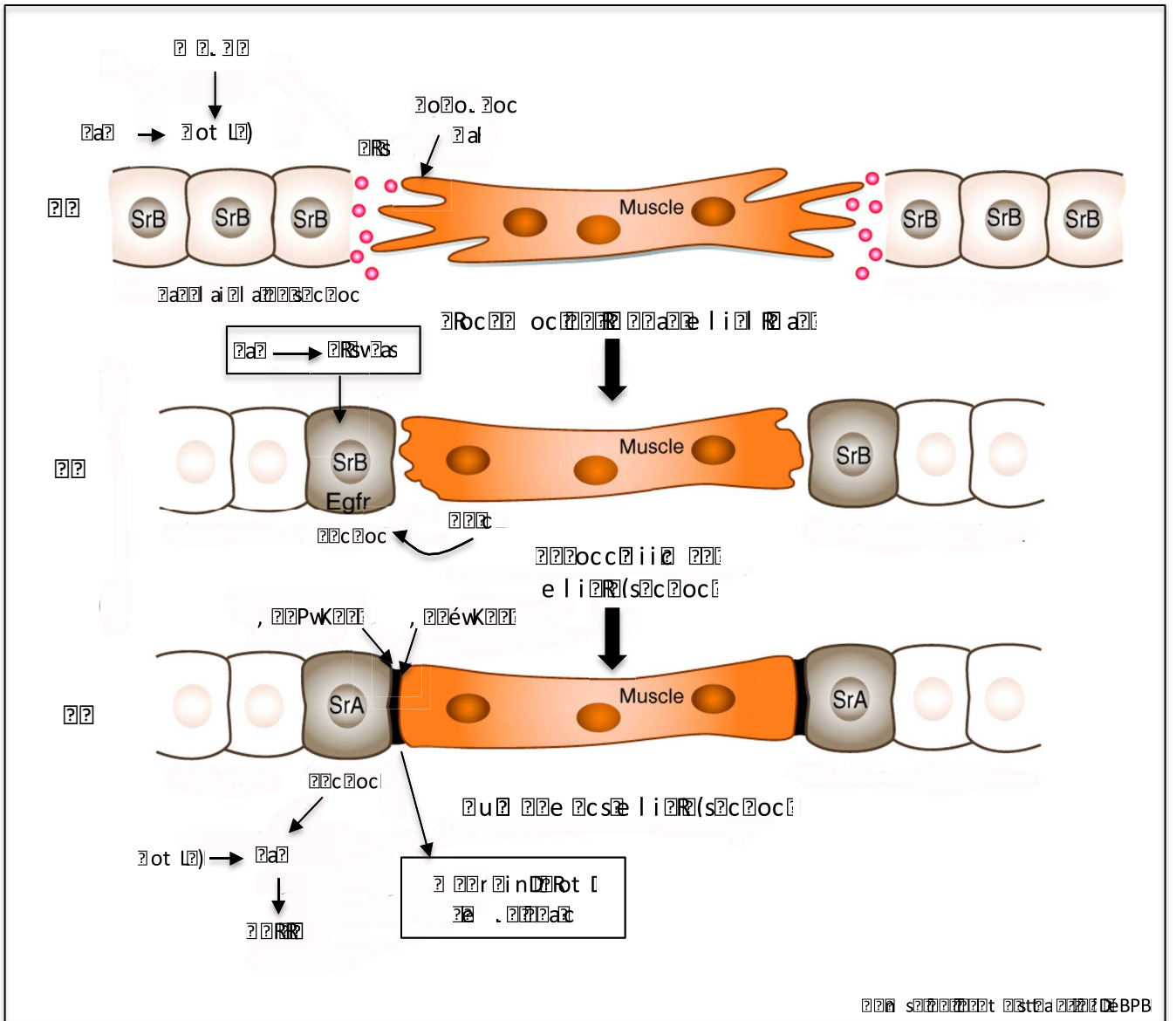
#### 4.1 Détermination des cellules de tendon

La détermination précoce des cellules tendineuses se fait indépendamment du muscle. En effet, les précurseurs de tendon qui proviennent de l'ectoderme arrivent en parallèle des cellules fondatrices musculaires au cours de l'embryogenèse. Un des premiers gènes qui induit la spécification de ces cellules est le facteur de transcription de la famille EGR (Early Growth Response), Stripe B (SrB), un des deux isoformes produits par le gène *stripe* (*Sr*). La perte de *Sr* induit l'absence de formation de cellules tendineuses menant à la désorganisation complète de la musculature somatique (Vorbrüggen and Jäckle, 1997). Le patron d'expression de SrB dans l'ectoderme est assez complexe et se fait selon les signalisations Hh, Wg, Egfr et Notch (Hatini and DiNardo, 2001). Cette régulation induit la formation au sein de l'ectoderme de plusieurs rangées de cellules tendineuses. Une fois les cellules ectodermes transformées en cellules précurseurs de tendon, celles ci dirigent les cellules musculaires en migration vers leurs sites spécifiques d'attachement et cela se fait notamment par l'expression de molécules d'attraction.

#### 4.2 Guidance et attachement

Dans les cellules précurseurs de tendon, SrB régule l'expression de Slit et Lrt (leucine rich tendon specific protein)(Wayburn and Volk, 2009; Kramer et al, 2001). Ces protéines permettent de guider les myotubes vers leurs sites d'attachement respectifs situés en bordure de segment. Slit est une molécule d'attraction sécrétée par les précurseurs et Lrt est une protéine transmembranaire qui permet l'arrêt de la migration au niveau des cellules tendineuses. Slit et Lrt interagissent directement avec le récepteur Roundabout (Robo) situé à l'extrémité des myotubes (Kramer et al, 2001) (Figure 19A). A ce jour l'action de Slit comme chemoattractant reste à élucider.

Dans les muscles latéraux transverses (LTs) la signalisation de guidance vers les sites d'attachement internes est différente. En effet ces muscles n'expriment pas de manière importante Robo et sa surexpression entraine un attachement anormale des LTs en bordure de segment. A la place de Robo, les LTs expriment le récepteur tyrosine kinase Derailed (Drl) (Callahan et al., 1996). Dans les mutants Drl les muscles latéraux ignorent les cellules



MyoD is a transcription factor that binds to DNA and initiates muscle-specific gene expression. It is a key regulator of muscle development and is expressed in muscle precursor cells. MyoD binds to specific DNA sequences, known as MyoD binding sites, and recruits other transcription factors to initiate the expression of muscle-specific genes. This process is essential for the differentiation of muscle cells from stem cells.

tendineuses exprimant *Sr* et continuent de s'allonger empêchant la formation de la jonction myotendineuse. Le ligand de ce récepteur pourrait être la molécule sécrétée *Wnt5*. En effet dans les mutants *Wnt5* les phénotypes observés ressemblent à ceux de *Drl*, les LTs dépassent leurs sites d'attachement et se fixent ailleurs au niveau de l'épiderme. La restauration de l'expression de *Wnt5* dans les tendons ou dans les muscles restaure le phénotype mutant (Lahaye et al., 2012).

Une autre protéine, *Kon Tiki (Kon)*, est nécessaire à la migration des muscles ventraux latéraux (VLs) vers les tendons, (Estrada et al., 2007b; Schnorrer et al., 2007). *Kon* est une protéine transmembranaire accumulée au niveau des filopodes. La perte de *Kon* dans les muscles ventraux engendre des projections de filopodes anormales altérant la migration des myotubes (Schnorrer et al., 2007). *Kon* se lie à la protéine cytoplasmique *Grip* (Glutamate receptor interacting protein) via son domaine PDZ (Swan, 2004). La perte de *Grip* induit le même type de phénotype que *Kon* lorsqu'il est muté. Les défauts spécifiques des muscles ventraux suggèrent que l'activité de *Kon* dépend donc de sa liaison avec *Grip*. Une dernière protéine, *Echinoid (Ed)* de la famille des Igs transmembranaires semble aussi nécessaire à l'attachement des VLs (Swan et al., 2006). *Ed* se lie à *Grip* de la même façon que *Kon*, avec son domaine PDZ. De plus *Ed* colocalise avec *Grip* à l'extrémité des muscles VLs et ces deux éléments présentent des interactions génétiques fortes du fait des défauts d'extension des myotube VLs qu'ils engendrent respectivement. Comment ce complexe permet la migration des myotubes est encore inconnu.

Au cours de l'élongation la signalisation *Slit/Robo* stimule la sécrétion du ligand EGF *Vein* au niveau de la jonction myotendineuse par les muscles (Yarnitzky et al., 1997) (Figure 19B). *Vein* se lie aux récepteurs EGF au niveau d'un précurseur de tendon sélectionné pour l'attachement. Cette liaison active la voie EGFR que dans cette cellule précurseur et cela entraîne la perte des marqueurs de tendon dans les cellules non sélectionnées (Becker et al., 1997; Yarnitzky et al., 1997). Dans le précurseur, la signalisation *Ras* médié par EGFR régule l'expression des différentes formes de *Held out wing (How)* (Nabel-Rosen et al., 2002). *How* est une protéine de liaison à l'ARN nécessaire pour le contrôle post-transcriptionnel de *SrA* ce stade la surexpression de la forme courte *How(S)* et la diminution de la forme longue *How(L)* permet la stabilisation des transcrits *Stripe A (SrA)* (Vолоhonsky et al., 2007). *SrA* active des gènes de différenciation des tendons tels que *Delilah (Dei)*, qui code pour un facteur de transcription qui régule l'expression d'une intégrine, et *Shortstop (Shot)*, une



plakine qui connecte le cytosquelette d'actine au microtubules régulant les propriétés élastiques des cellules tendons (Schweitzer et al., 2010; Subramanian et al., 2003).

#### 4.3 Jonction myotendineuse

Une fois que les myotubes ont atteint les cellules de tendon, les deux types cellulaires sécrètent des protéines de la matrice extracellulaire (ECM) pour former la jonction myotendineuse (MTJ). Les cellules tendineuses déposent de la Laminine (Lam) de la Trombospondine (Tsp) et sécrètent aussi Slowdown (slown) qui séquestre temporairement Tsp jusqu'à ce que ces partenaires, les intégrines musculaires ( $\alpha$ PS2/ $\beta$ PS) soient présents. Cela permet d'éviter la formation immature de MTJ (Gilsohn and Volk, 2010a, 2010b; Subramanian et al., 2007). Les myotubes sécrètent la Tigrin au niveau de l'ECM (Fogerty et al., 1994). Par la suite chaque type cellulaire exprime une combinaison de différentes protéines transmembranaires appelées intégrines qui permettent l'adhésion à l'ECM. Les myotubes se lient à l'ECM grâce aux intégrines  $\alpha$ PS2 et  $\beta$ PS produites respectivement par les gènes *inflated (inf)* et *myospheroid (mys)*. Ces protéines forment des hétérodimères  $\alpha$ PS2/ $\beta$ PS qui interagissent avec Tsp et Tigrin. Dans les cellules tendineuses les gènes *mys* et *multiple edematous wings (mew)* codent respectivement pour  $\beta$ PS et  $\alpha$ PS1. Les hétérodimères  $\alpha$ PS1/ $\beta$ PS se fixent à Lam (Figure 19C). Les jonctions myotendineuses non fonctionnelles observées dans les différents mutants des intégrines, de la laminine ou Tsp entraînent une dissociation complète des muscles et des tendons menant à une létalité embryonnaire. L'interaction entre les intégrines et leurs différents partenaires se fait à l'aide de la protéine Talin. La Talin induit un changement de conformation des intégrines permettant ainsi leur activation (Kim et al., 2012). La formation de dimères d'intégrines hétérophiliques ainsi que leur liaison aux composants de l'ECM permet l'ancrage des muscles et des tendons (Horton et al., 2016). Dans les muscles, la perte de Talin engendre un détachement et un arrondissement des muscles. Talin est donc nécessaire au maintien des connections entre les intégrines et le cytosquelette.

La MTJ est une structure indispensable pour la génération de force et le maintien des attachements muscles tendons au cours de la contraction musculaire (Charvet et al., 2012). Au final l'attachement des muscles aux cellules tendineuses suit un processus général cependant la diversité des sites d'attachement ainsi que des molécules impliquées laisse





suggérer que la mise en place de la jonction myotendineuse soit contrôlée de manière muscle spécifique par les iTFs.

## 5. Innervation

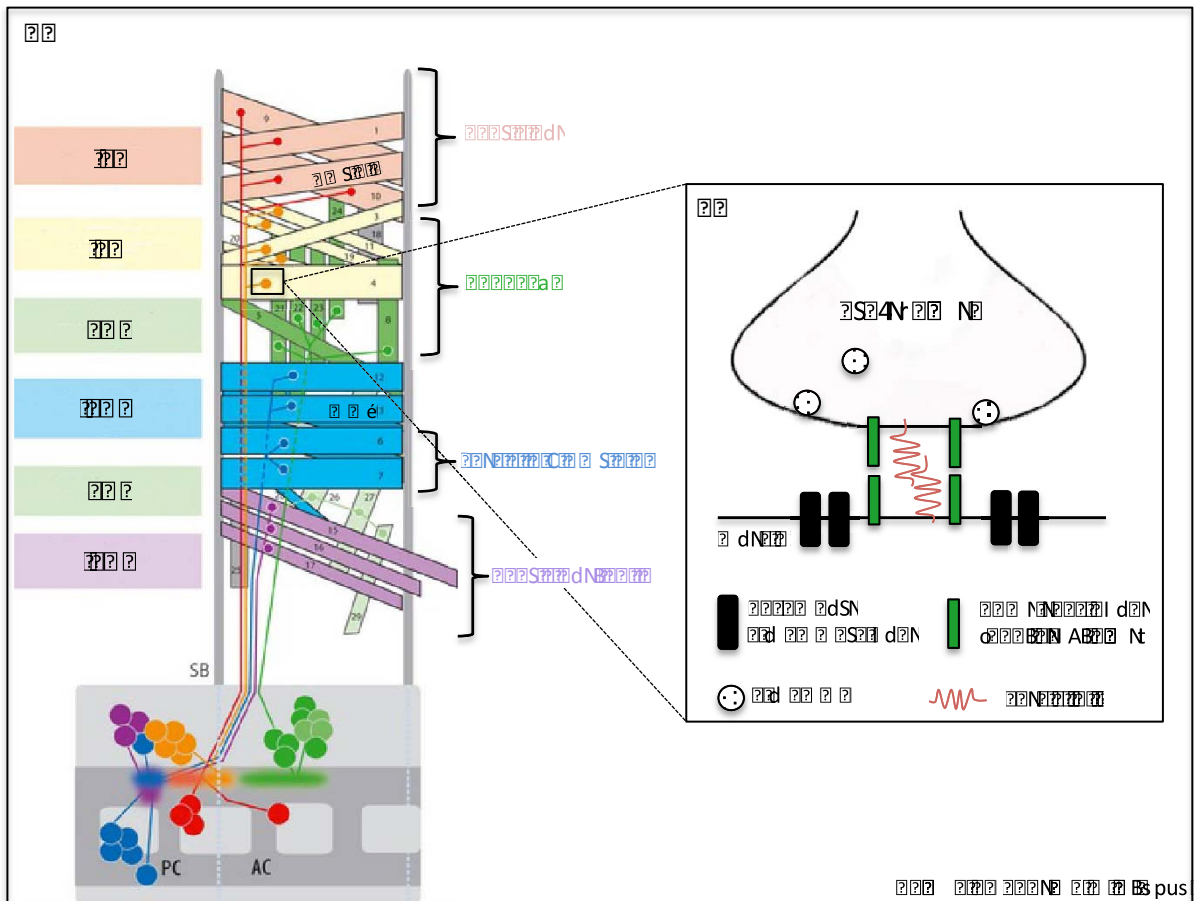
### 5.1 Formation des motoneurones

Un fois les muscles attachés aux cellules tendineuses ces derniers vont être innervés pour répondre aux signaux du système nerveux central (CNS). En effet, la capacité de se mouvoir de la larve provient de la contraction des muscles qui se fait grâce à un ensemble de motoneurones qui dérivent de 15 neuroblastes. Dans chaque hémisegment de l'embryon, les 30 muscles sont innervés par 36 motoneurones via six branches du système nerveux périphérique : les nerfs intersegmentaux (ISN, ISNb et ISNd), les nerfs segmentaux (SNa et SNc) et le nerf transverse (TN) (Halpern et al., 1991; Landgraf et al., 1997; Ruiz-Cañada and Budnik, 2006)(Figure 20A). Les motoneurones qui proviennent de la corde nerveuse centrale du CNS sont caractérisés par leur position, leur projection axonale ainsi que par le type de muscles qu'ils innervent (Landgraf et al., 1997). L'innervation des muscles se fait donc tout d'abord par la formation des motoneurones, puis migration de ces derniers vers leurs muscles respectifs et enfin établissement de la jonction neuromusculaire.

Au cours du développement les motoneurones produisent des filopodes au niveau des régions d'innervation puis les restreignent plus tardivement à leur cible musculaire spécifique (Sink and Whittington, 1991). Les muscles jouent aussi un rôle clé dans la mise en place de l'innervation. En effet les informations contenues dans les FC sont nécessaires pour diriger la défasciculation des axones individuels aux sites d'innervation spécifiques de chaque muscles (Landgraf et al., 1999). Cela se fait notamment grâce à la formation par les muscles de filopodes post-synaptiques appelés « myopodes ». Ces myopodes à la surface des muscles s'étendent initialement de manière aléatoire puis deviennent progressivement restreints au site d'innervation où ils interagissent avec les filopodes pré-synaptiques (Ritzenthaler and Chiba, 2003; Ritzenthaler et al., 2000).

### 5.2 Migration et Reconnaissance muscle-motoneurone

Le degré important de précision et l'habileté des motoneurones pour sélectionner leur muscle d'intérêt suggère la présence de molécules actives côté musculaire et nerveux (Van Vactor et al., 1993). Différentes études ont permis d'identifier plusieurs molécules de



The diagram shows a vertical column of colored boxes on the left, representing different layers or cell types. On the right, a vertical stack of numbered layers (1-10) shows the anatomical structure. At the bottom, a synaptic terminal (SB) contains presynaptic (PC) and postsynaptic (AC) components. An inset diagram shows a cross-section of a synapse with a presynaptic terminal and postsynaptic membrane, illustrating the transmission of a signal (red wavy line) across the synaptic cleft. A legend at the bottom right defines the symbols used in the diagram.

reconnaissance exprimées par des sous populations de muscles (Rose et al., 1997; Shishido et al., 1998; Winberg et al., 1998). Parmi ces molécules on retrouve les protéines Connectin (Con), Capricious (Caps) de la famille des LRR (leucine reach repeat protéine) et Fasciclin 3 (Fas3) de la famille des immunoglobulines Ig(SF) (Chiba et al., 1995; Nose et al., 1997; Shishido et al., 1998). Ces protéines sont des molécules d'adhésion cellulaire homophiliques qui sont exprimées dans des sous populations musculaires et dans les motoneurones qui innervent ces mêmes muscles. *Fas3* est exprimé dans le motoneurone RP3 et dans les muscles VL3-4, *Con* est présente dans le DT1, LT1-4, LO1, VA2, VA3 et dans les motoneurones 5, 18, 21-24, 27, 29 et pour finir *Caps* est retrouvée dans plusieurs motoneurones associés à différents muscles (DO1, DO2, DA1, DA2, VL1, VO1, VO4-6, VA3). Dans les mutants de ces différentes molécules les phénotypes observés ne sont pas très forts suggérant ainsi que ces molécules aient des fonctions redondantes. C'est le cas pour *Caps* qui peut être compensée par la protéine Tartan (LRR)(Kurusu et al., 2008).

D'autres protéines sécrétées par les muscles *NetrinA* (*NetA*) et *NetrinB* (*NetB*) ont une action soit répulsive soit attractive sur la guidance. *NetA* est exprimé dans 2 muscles dorsaux et *NetB* dans un muscle dorsal et 2 ventraux (Mitchell et al., 1996). La signalisation attractive ou répulsive de *NetB* se fait respectivement à travers les récepteurs *Frazzled* et *UNC-5* qui sont différentiellement exprimés au niveau des motoneurones (Keleman and Dickson, 2001; Kolodziej et al., 1996; Labrador et al., 2005). La protéine *Toll* qui est présente différentiellement dans les muscles ventraux a aussi un effet répulsif sur certains motoneurones (ex : RP3) (Halfon et al., 1995; Inaki et al., 2010). Lorsque *Toll* est surexprimé dans un muscle innervé normalement par le motoneurone 3 ce dernier échoue à trouver ses sites d'innervation (Rose et al., 1997). *Wnt4*, une glycoprotéine de la famille *Wnt* est sécrétée par le muscle VL2 et prévient la formation de synapses par les motoneurones qui ciblent les muscles voisins (Inaki et al., 2007) (Figure 20A).

La protéine homophile *Fasciclin2* (*Fas2*) et la protéine sécrétée *Semaphorin 2a* (*Sema2a*) sont aussi présentes au sein de tous les muscles mais ne sont pas nécessairement requises pour la spécificité neuromusculaire. Cependant leurs surexpressions entraînent des changements de sites d'innervation ou une inhibition de la formation des synapses par les motoneurones (Davis et al., 1997; Matthes et al., 1995). D'autres acteurs identifiés à travers un crible gain de fonction semble affecter un nerf dans son ensemble (Kurusu et al., 2008). C'est le cas de *Hattifater* et *Convoluted* (protéines LRR) retrouvées de manière ubiquitaire



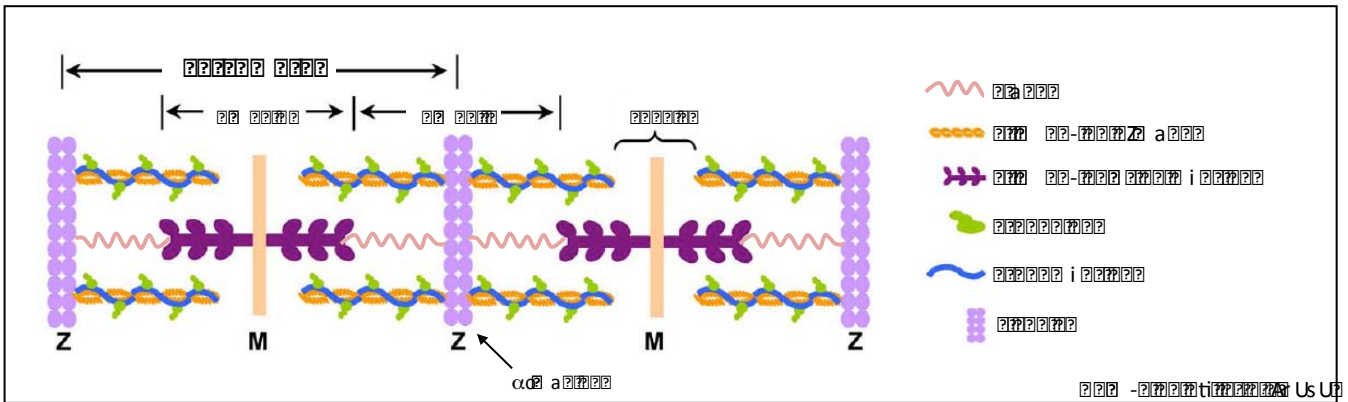
dans le muscle et qui sont nécessaires à la régulation de la guidance de l'ISNb. En leurs absences les projections axonales produites par l'ISNb sont anormales et ont une mauvaise localisation.

Bien qu'un certain nombre d'acteurs de la mise en place de la jonction neuromusculaire aient été découverts la régulation ou les facteurs qui contrôlent l'expression des molécules de reconnaissance de manière muscle spécifique reste encore à définir. Un facteur de transcription a été identifié comme un régulateur négatif de Toll, Teyrha-Meyhra (Tey). *Tey* est exprimé seulement dans le VL1 ou il réprime Toll. En effet lorsque Tey est absent le VL1 est moins innervé du fait de la persistance de Toll (Inaki et al., 2010).

### 5.3 Jonction neuromusculaire

Une fois que les motoneurones sont arrivés au niveau du muscle d'intérêt, l'axone développe une synapse fonctionnelle entre le muscle et la branche terminale du neurone (Prokop, 2006). Parmi les trois types de motoneurones, les neurones glutamatergiques de type I influencent directement leurs partenaires post-synaptiques et forment des jonctions neuromusculaires classiques pour réguler la contraction musculaire. A chaque stade de développement de chaque muscle individuel, la synapse de type I adopte une morphologie stéréotypée : un point d'entrée du nerf bien défini, un patron de ramification reproductible (nombre, localisation et taille) et un nombre spécifique de boutons synaptiques qui arborescent la synapse (Broadie and Bate, 1993; Johansen et al., 1989; Yoshihara et al., 1997). Les boutons sont des structures pré-synaptiques qui permettent la transmission du potentiel exciteur aux muscles. Pour répondre aux signaux des synapses de type I, les muscles expriment localement les récepteurs au glutamate (GluRs) à la membrane post-synaptique. Il a été montré que ces récepteurs sont sous le contrôle transcriptionnel du facteur de transcription à doigt de zinc Longitudinals lacking (*Lola*). En effet, la perte de fonction engendrée par des ARN interférents dirigés contre *Lola* montre une diminution des GluRs à la NMJ ainsi qu'une diminution de leurs transcrits (Fukui et al., 2012).

Les récepteurs deviennent actifs après le relargage par la membrane pré-synaptique du neurotransmetteur glutamate (Figure 20B). Par la suite les canaux ioniques des GluRs facilite l'influx de  $Na^{2+}$  dans le muscle ce qui induit un relargage du  $Ca^{2+}$  et la contraction musculaire. Les GluRs présentent une localisation et des activités différentes au niveau des



The diagram illustrates the structure of a sarcomere, the basic contractile unit of muscle. It shows two vertical purple bars representing Z-discs (Z) and two vertical orange bars representing myofibrils (M). The space between the Z-discs is divided into several regions: the Z-disc, the I-band, the A-band, and the H-band. The A-band is the region containing the thick filaments (myosin), while the I-band contains the thin filaments (actin). The H-band is the region containing only thick filaments. The diagram also shows the arrangement of various proteins: actin (red wavy lines), myosin (orange beaded lines), titin (purple branched structures), and other proteins like nebulin (green) and alpha-actinin (blue). The legend on the right identifies these components: red wavy line for actin, orange beaded line for myosin, purple branched structure for titin, green shape for nebulin, blue shape for alpha-actinin, and purple bar for Z-disc.

synapses. La distribution à la membrane post-synaptique de ces différents GluRs est contrôlée par la protéine Discs large (Dlg) et Calmoduline dependent protein kinase II (CaMKII)(Chen and Featherstone, 2005; Morimoto et al., 2010). La première contraction du muscle dépendante du système nerveux est observée au stade 17 (Crisp et al., 2008).

## 6. Sarcomères

En parallèle des étapes d'attachement et d'innervation le muscle a besoin de former un appareil contractile qui lui permette de supporter les contractions musculaires. Cela se fait au cours de la myofibrillogenèse. Cette étape définit la formation et l'assemblage des myofibrilles qui sont des structures cylindriques du muscle squelettique composées d'éléments répétés appelés sarcomères, unité contractile du muscle.

Un sarcomère, défini comme une région entre deux lignes Z, contient deux types de filaments, des filaments épais de myosine et des filaments fins d'actine associés à leurs protéines respectives. Les filaments fins d'actine sont associés au complexe tropomyosine-troponine et sont ancrés à la ligne Z par l' $\alpha$ -Actinin (molécule de liaison à l'actine) (Clark et al., 2002). Au sein du sarcomère les filaments épais occupent la bande sarcomérique A alors les filaments fins se retrouvent dans la bande I (Sparrow and Schöck, 2009). Au milieu du sarcomère les protéines de la ligne M telles que Obscurin et Titin permettent d'ancrer les filaments épais de myosine entre eux (Katzemich et al., 2012; Trinick, 1996). La protéine Titin s'étend de la ligne M à Z et fournit durant la contraction une élasticité nécessaire au maintien de l'organisation des sarcomères (Gregorio et al., 1998) (Figure 21). Obscurin permet l'organisation de la ligne M et sa fixation à la membrane plasmique musculaire, le sarcolemme.

Les myofibrilles s'attachent à l'ECM au niveau de leurs lignes Z par l'intermédiaire de structures appelées costamères (Ervasti, 2003). Le complexe ligne Z/costamère est morphologiquement similaire à la jonction myotendineuse et contient le même type de protéines telles que les intégrines. La protéine Zasp (Z-band alternatively spliced PDZ motif protein) est nécessaire à l'assemblage des costamères (Jani and Schöck, 2007). En effet en son absence les sites d'adhésion des intégrines sont perturbés. La connexion des lignes Z aux tissus environnants permet d'assurer une contraction musculaire synchrone et uniforme.

Une fois les sarcomères formés, le chevauchement des filaments fins et des filaments épais induit un glissement de l'un sur l'autre créant ainsi une compression du muscle dans





une direction longitudinale, générant la force et le mouvement. Ce mécanisme est initié par le réticulum sarcoplasmique qui sous l'effet du système nerveux va relarguer du calcium faisant augmenter sa concentration cytoplasmique. Il y a alors levée de l'inhibition de la troponine sur l'interaction actine/myosine et donc contraction. Ensuite lors de la relaxation musculaire, les pompes calcium ATP-dépendantes permettent le retour du calcium cytoplasmique dans le réticulum afin de diminuer sa concentration. Cette boucle de régulation se répète autant de fois que le mouvement est nécessaire.

## **V. identité et diversification musculaire**

La spécification du mésoderme somatique puis la mise en place d'un programme d'identité au sein de chaque FCs est nécessaire à la diversification musculaire. Cela se fait notamment à travers l'expression de facteurs généraux tels que Twist (Twi) ou Myocyte enhancer factor (Mef2) qui sont requis pour les étapes précoces puis par la combinaison de facteurs dits d'identité (iTFs) qui fournissent toutes les informations de taille, de fusion, d'attachement et d'innervation aux muscles dans lesquels ils sont exprimés. Je détaillerai tout d'abord les facteurs Twi et Mef2 pour me consacrer ensuite à quelques iTFs qui régissent la formation de certaines sous populations musculaires dorsales, latérales et ventrales.

### *1. Les facteurs de régulation transcriptionnelle généraux*

#### **1.1. Twist**

Le premier acteur nécessaire à la myogenèse embryonnaire est Twist. Ce dernier fait partie de la famille des facteurs de transcription à domaine bHLH. Twi est un facteur myogénique dit « master » car il est capable d'induire la formation de fibres musculaires ectopiques lorsqu'il est surexprimé dans l'épiderme. De plus, en son absence les muscles somatiques ne sont pas formés (Baylies and Bate, 1996).

Comme démontré à travers différentes analyses génétiques, le profil d'activité de Twi ainsi que les ensembles de gènes qu'il régule varient au cours du développement (Cripps et al., 1998; Sandmann et al., 2007; Zeitlinger et al., 2007). En effet afin de comprendre comment Twi orchestre le programme transcriptionnel des cellules du mésoderme des expériences d'immunoprécipitation de la chromatine suivies d'analyses par puces à ADNc (Chip on Chip) ont été réalisées à différents stades de développement allant de la



gastrulation à la subdivision précoce du mésoderme (2-6h AEL, stade 5-9)(Sandmann et al., 2007). Cette étude a permis d'identifier plus de 2000 CRMs (cis regulatory modules) et 500 gènes cibles directs qui varient en fonction des étapes de développement. Les 500 nouvelles cibles identifiées font partie de groupes de gènes impliqués dans la morphogenèse, la migration et la prolifération cellulaire. Les facteurs de transcription constituent aussi un large groupe de cibles directes de Twi. Parmi eux on retrouve des gènes « gap », des gènes de segmentation, des gènes homéotiques ainsi que des gènes d'identité musculaires qui sont régulés de manière hiérarchique.

Au niveau moléculaire Twist se fixe sur des séquences E-box (CANNTG) et son domaine HLH permet de s'homodimériser ou de s'hétérodimériser avec la protéine HLH Daughterless (Da) qui est ubiquitairement exprimée dans l'embryon (Cabrera and Alonso, 1991; Cronmiller and Cummings, 1993). Lorsque Twist est présent sous forme homodimère il se lie préférentiellement aux E-box ayant pour séquence CATATG alors que sous la forme hétérodimère avec Da, il se fixe aussi bien sur des E-box CACCTG que CATATG (Castanon et al., 2001). En fonction de sa forme son activité n'est pas la même. En effet les homodimères Twi/Twi ont une fonction activatrice des gènes nécessaires à la myogenèse somatique alors que les hétérodimères Twi/Da ont une fonction répressive. L'activité répressive des hétérodimères est assurée par le domaine REP (Répression) de la protéine Da et cette dernière se fait de manière tissu et stade dépendant (Wong et al., 2008).

Au final, dans la partie antérieure de chaque hémisegment on retrouve majoritairement des homodimères de Twist permettant l'expression de *Sna*, *Htl*, *Mef2* et *Tin* nécessaire à la spécification du mésoderme somatique alors que dans la partie postérieure les hétérodimères Twi/Da répriment les gènes myogéniques *Mef2* et *Isc* favorisant ainsi la spécification du mésoderme viscéral et du corps gras (Castanon et al., 2001; Cripps et al., 1998; Zeitlinger et al., 2007). La surexpression de Da dans le mésoderme inhibe les étapes précoces de la myogenèse somatique induisant une perte importante de muscles.

Twist interagit aussi génétiquement et physiquement avec Akirin, un cofacteur qui permet l'interaction de Twi avec le complexe de remodelage de la chromatine Brahma. Cette interaction faciliterait les changements de conformation de la chromatine et donc favoriserait l'expression de manière sélective de quelques gènes cibles de Twi au cours de la myogenèse (Nowak et al., 2012). Une fois les myoblastes différenciés, ces derniers perdent l'expression de Twi et seules les AMPs maintiennent son expression (Anant et al., 1998).

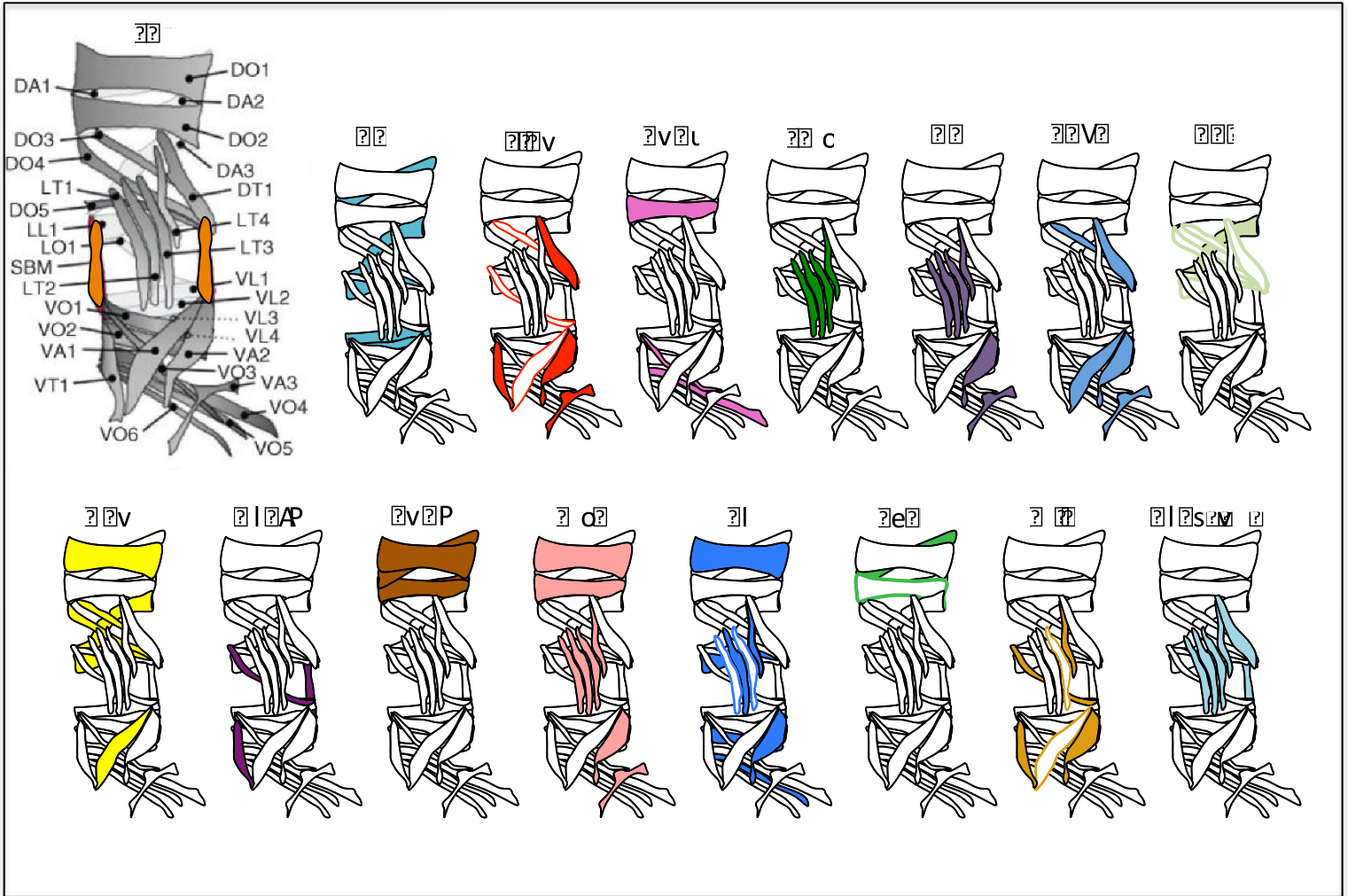


## 1.2 Mef2

La protéine Mef2 fait partie de la famille des facteurs de transcription MADS (Shore and Sharrocks, 1995). Chez la drosophile un seul gène *Mef2* est retrouvé. Il code pour une protéine qui contient dans sa partie N-terminale un domaine MADS hautement conservé et un domaine adjacent appelé « domaine Mef2 » qui ensemble permettent la dimérisation, la liaison à l'ADN et l'interaction avec des cofacteurs tel que Twi (Black and Olson, 1998; McKinsey et al., 2002). La partie C-terminale fonctionne comme un domaine d'activation transcriptionnelle. Mef2 se lie à la séquence d'ADN consensus YTA(A/T)4TAR sous la forme homo ou hétérodimère (Molkentin et al., 1995). Des mécanismes transcriptionnels et post-traductionnels gouvernent les fonctions de Mef2.

Le rôle central de Mef2 est d'orchestrer en aval de Twi le développement du muscle. *Mef2* est exprimé dans le mésoderme précoce puis dans différents lignages de cellules musculaires (somatique, viscérale et cardiaque) où il est requis pour la différenciation (Bour et al., 1995; Lilly and al., 1995). L'expression de *Mef2* dans le mésoderme précoce se fait via la liaison de Twi sur son enhancer mésodermale (Cripps et al., 1998). Plus tardivement Twi et Lmd agissent sur deux enhancers différents de Mef2 pour contrôler son expression dans des cellules musculaires spécifiques (Duan et al., 2001). Mef2 maintient aussi sa propre expression au cours de la différenciation en activant son propre enhancer distal spécifique du muscle somatique (Cripps et al., 2004). Dans les mutants *Mef2* les FCs sont spécifiés de manière appropriée cependant il y a un blocage complet de la fusion des myoblastes et une absence d'expression des marqueurs de différenciation (Bour et al., 1995; Lilly and al., 1995; Prokop et al., 1996). Similairement les cellules cardiaques sont formées correctement cependant aucune expression de protéines contractiles cardiaques n'est observée.

Dans le but d'identifier les rôles de Mef2 au cours de la myogenèse des expériences de Chip on Chip via des puces couvrant plus de 50% du génome ont été réalisées à différents stades de développement (2-12h AEL). Cela a mené à l'identification de 670 régions fixées par Mef2 et de 200 gènes cibles directs montrant ainsi le rôle central de Mef2 dans la hiérarchie transcriptionnelle au cours de la myogenèse (Junion et al., 2005; Sandmann et al., 2006). Mef2 contrôle l'expression de gènes nécessaires à la division asymétrique (Ex : *Delta*, *mam*, *bib*, *E(spl)*), l'expression de gènes d'identité dans les FCs (Ex : *Slou*, *eve*, *ap*, *lbe*, *Kr*) mais aussi l'expression de gènes de différenciation impliqués dans la fusion, l'attachement,



The diagrams illustrate the complex arrangement of muscles in the lumbar spine. The legend identifies the following muscles: DA1, DA2, DA3, DO1, DO2, DO3, DO4, DO5, LL1, SBM, LT1, LT2, LT3, LT4, VL1, VL2, VL3, VL4, VO1, VO2, VO3, VO4, VO5, VO6, VA1, VA2, VA3, and VT1. The main diagrams show color-coded muscle groups: blue (DO1-5), red (VL1-4), purple (VT1), green (VA1-3), yellow (LL1, SBM), brown (DA1-3), pink (DO1-5), blue (VL1-4), green (VA1-3), yellow (VT1), purple (DA1-3), brown (DO1-5), pink (VL1-4), blue (VA1-3), and yellow (VT1).

l'innervation ou la contraction (Ex : *Mhc*, *Tm1*, *Pax*, *B3-tubulin*, *if*, *blow*, *sns*, *Tig*, *Mbl*, *NetrinA/B*, *Lola*, *Titin*, *sls*) (Junion et al., 2005; Sandmann et al., 2006, 2007). L'expression ectopique de Mef2 dans l'ectoderme permet l'activation de ces gènes cibles suggérant que Mef2 est essentiel et suffisant pour l'expression d'un large pourcentage de ses gènes cibles.

Twi et Mef2 régulent coopérativement l'expression de gènes musculaires de manière similaire à MyoD et Mef2 chez les vertébrés.

## 2. Les facteurs d'identité

Chez la drosophile une vingtaine de gènes ont été identifiés comme étant nécessaires à la formation de la musculature embryonnaire (Figure 22). Ces gènes d'identité partagent des caractéristiques communes. En effet, ce sont tous des gènes qui codent pour des facteurs de transcription exprimés seulement dans des sous populations de FCs et précurseurs musculaires. Alors que la plupart des gènes sont exprimés dans plusieurs FCs selon une combinatoire particulière certains comme *ladybird (lb)* ou *eve* ont une expression spécifique à un seul FC/muscle, respectivement le DA1 et le SBM (segment border muscle) (Jagla et al., 1998; Knirr and Frasch, 2001). Cette combinaison de facteurs au sein de différentes FCs est à l'origine de toute la diversification musculaire.

### 2.1 Le facteur dorsal Tailup/islet1

Un des facteurs clés de la diversification de la musculature dorsale est Tailup/Islet1 (Tup). Tup fait partie de la famille des facteurs de transcription à LIM domaine (Lin11, Islet1 et Mec3). Il est composé de deux domaines à doigt de zinc et d'un domaine LIM nécessaire à l'interaction protéine-protéine. Il est connu pour avoir un rôle dans la spécification et le guidage des motoneurones mais aussi dans la formation du cœur et de l'organe hématopoïétique (glande lymphatique) (Mann et al., 2009; Santiago and Bashaw, 2017; Thor and Thomas, 1997).

Au niveau musculaire il est exprimé dans 4 cellules précurseurs à l'origine des muscles DO1, DO2, DA1, DA2 (dorsal acute et dorsal oblique). Dans les mutants *Tup* la musculature dorsale est sévèrement désorganisée et le muscle DA2 se transforme en DA3 (Boukhatmi et al., 2012). Cette transformation est aussi observée dans les embryons sur-exprimant le facteur Collier (Col) dans tout le mésoderme *via* le pilote Twi-Gal4 (Enriquez et al., 2012). Ce phénotype suggère que Tup réprime l'activité de Col dans le DA2. Cette



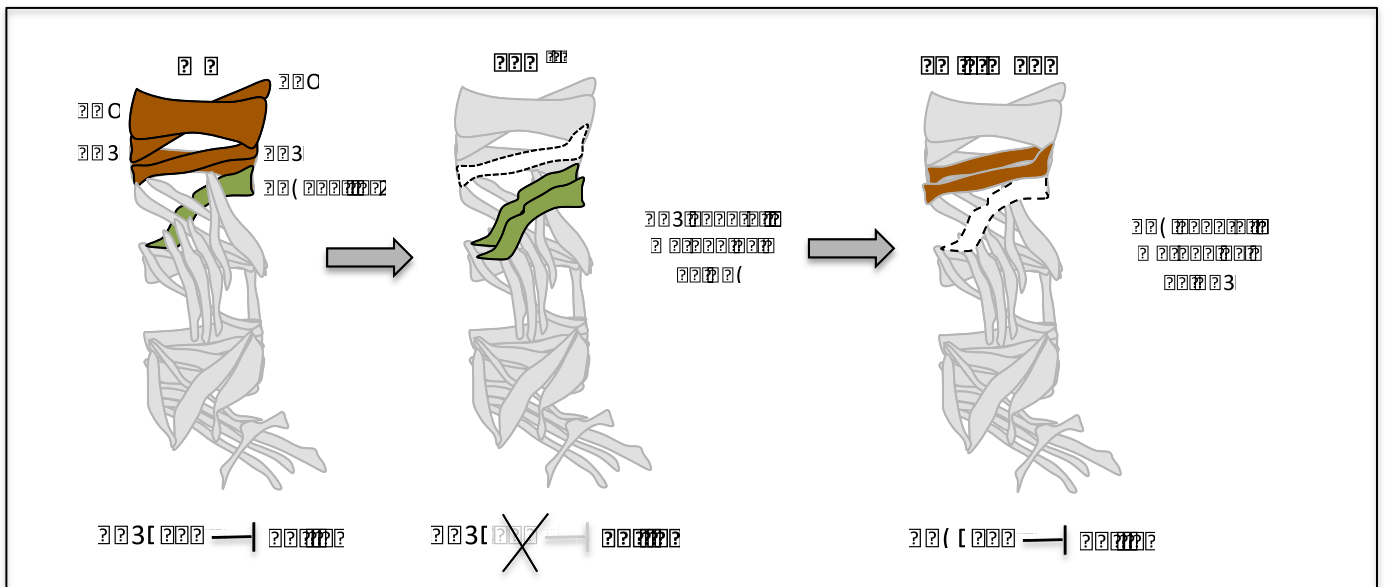


Diagram illustrating the relationship between the diaphragm and the lungs, showing three stages of diaphragm position relative to the lungs.

Stage 1: Normal anatomy. The diaphragm is positioned at the level of the T10/T11 vertebrae (C3). The lungs are positioned at the level of the T12/L1 vertebrae (C4).

Stage 2: Abnormal anatomy. The diaphragm is positioned lower than normal (dashed line), labeled as C3 (normal). The lungs are positioned lower than normal (dashed line).

Stage 3: Abnormal anatomy. The diaphragm is positioned higher than normal (dashed line), labeled as C4 (normal). The lungs are positioned higher than normal (dashed line).

hypothèse a été validée en sur-exprimant *Tup* dans tout le mésoderme. Dans cette condition, la perte d'expression de *Col* dans le DA3 est observée prouvant que *Tup* est capable de réprimer cet iTF dans le muscle DA2 (Figure 23). Au niveau transcriptionnel, l'activation précoce de *Tup* passe par la fixation du facteur *Tin* sur le CRM appelé *IsletH*. Son expression est ensuite maintenue au cours de la différenciation musculaire *via* un CRM tardif qui correspond à un enhancer de muscle dorsal (*DME*) positivement régulé par *Tup* lui-même. En l'absence de *Tup*, *IsletH* est toujours activé alors que l'expression contrôlée de *LacZ* par le *DME* est complètement perdue.

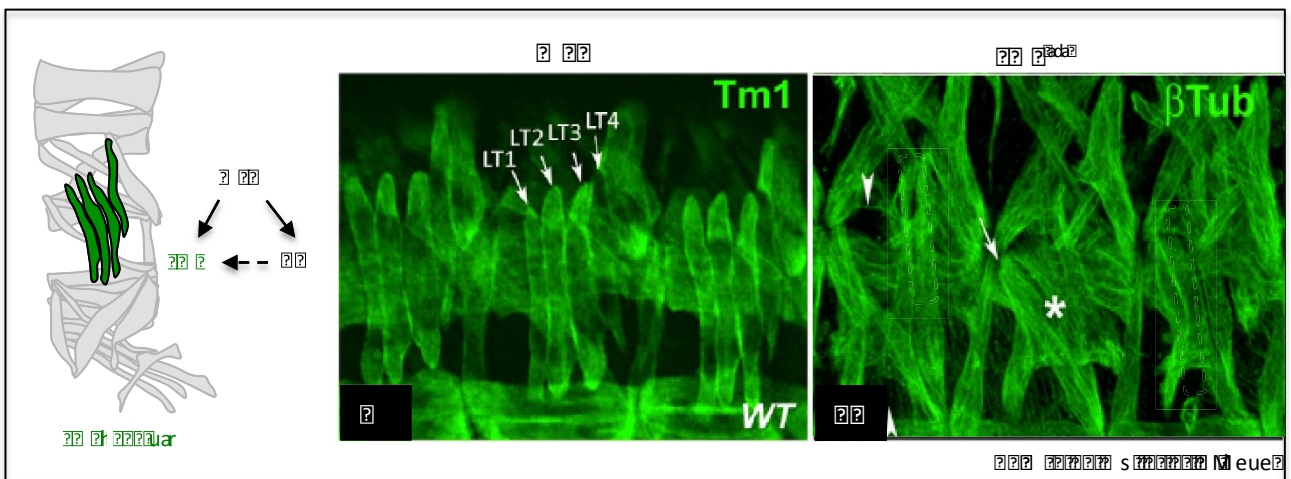
Malgré que *Tup* soit nécessaire à l'expression de *Kr* et *Col* respectivement dans le DO1 et le DA2 ce dernier ne semble pas à l'origine de la cascade de iTFs dans les muscles dorsaux. En effet la régulation d'autres iTFs dorsaux tels que *Eve*, *Runt* et *Vestigial* dans les muscles DA1, DO2 et DA2 et DA1 n'est pas dépendante de *Tup* (Boukhatmi et al., 2012; Mann et al., 2009).

La nécessité de *Tup/islet1* dans la spécification et la différenciation de sous populations musculaires rappelle la fonction de son orthologue vertébré *Islet1* dans la formation d'une sous population de muscle de la mâchoire (Harel et al., 2009; Nathan et al., 2008). Cela révèle que la diversification des muscles abdominaux chez la drosophile et des muscles craniofaciaux chez les vertébrés implique des facteurs conservés.

## 2.2 Les facteurs latéraux

### 2.2.1 *Lms*

Le gène *Lateral muscle scarcer* (*Lms*) fait partie de la famille de gènes d'identité qui codent pour des facteurs de transcription à homéodomaine (HD). Son expression est tout d'abord détectée au stade 11 dans un petit groupe de cellules du mésoderme situé dans les trois segments thoraciques (T1-T3). Par la suite il s'active dans une cellule de chaque segment abdominal puis dans 4 FCs latérales au début du stade 13. Au stade 15 il n'est retrouvé qu'au niveau des 4 muscles latéraux transverses (LT1-LT4)(Müller et al., 2010) (Figure 24). Il s'avère qu'un autre iTF à LIM domaine, *Apterous* (*Ap*), identifié 20 ans plutôt, présente un profil d'expression spatial et temporel similaire (Bourgouin et al., 1992). En effet l'utilisation du reporter *ap-GFP* a permis de suivre l'expression de *ap* en même temps que *Lms* (hybridation in situ). Bien que *Ap* soit exprimé tout d'abord seul dans un groupe de FCs latérales de chaque segment abdominal au stade 11, la première cellule exprimant *Lms* est



The figure illustrates the structure and localization of microtubule-associated proteins. The schematic on the left shows a microtubule lattice with green protofilaments. The right panels show fluorescence microscopy images of microtubule-associated proteins (Tm1 and βTub) in a WT cell. The Tm1 panel shows protofilaments labeled LT1, LT2, LT3, and LT4. The βTub panel shows a microtubule lattice with a white asterisk marking a specific region. The text 'War' is at the bottom left, 'Back' is at the top right, and 's' and 'eue' are at the bottom right.

détectée que peu de temps après (stade 12) et est Ap positive. Par la suite l'ensemble des cellules dans la région latérale co-expriment *ap* et *lms* ce qui suggère que *ap* et *lms* participent ensemble au code combinatoire nécessaire à la spécification et la différenciation des 4 muscles LT.

Bien que *Lms* semble spécifique des muscles LT il est connu que d'autres iTFs participent à la mise en place du programme d'identité de ces muscles. En effet, le gène *muscle segment homeobox (msh)* est présent dans les FCs des muscles LT1-4 et *Krüppel (Kr)* est exprimé entre autre dans les muscles LT2, LT4 et les FCs d'origine (Nose et al., 1998; Ruiz-Gómez et al., 1997). Afin de connaître la position de *lms* dans la cascade transcriptionnelle nécessaire aux muscles LTs, des analyses de perte et gain de fonction pour *ap* et *msh* ont été réalisées. Les défauts les plus drastiques sont observés dans les embryons mutés pour *msh*. En effet, une réduction importante de *Lms* et une perte d'expression de *Kr* dans les FCs des muscles LTs sont observées. De plus l'expression de *ap* dans les précurseurs et les FCs est fortement réduite. *Msh* est donc situé en amont de la cascade de régulation des muscles LTs et semble nécessaire à l'activation des gènes en aval *ap*, *kr* et *lms*. Afin de vérifier si *ap* peut aussi avoir un effet direct sur l'activation de *lms* celui-ci a été surexprimé dans tout le mésoderme via le pilote d'expression 24B-Gal4. Dans cette condition l'expression de *lms* est retrouvée dans les muscles ventraux (VA1-VA3) alors qu'elle y est normalement absente dans les embryons sauvages. Le même type de résultat est observé avec la surexpression ectopique de *msh* suggérant que l'activation et la maintenance de *lms* dans les muscles LTs est sous le contrôle transcriptionnel de *Msh* et de *Ap* (Müller et al., 2010). Cependant dans les mutants *ap*, *lms* est toujours exprimé, ce qui suggère que ce dernier n'est pas suffisant pour réguler *lms*.

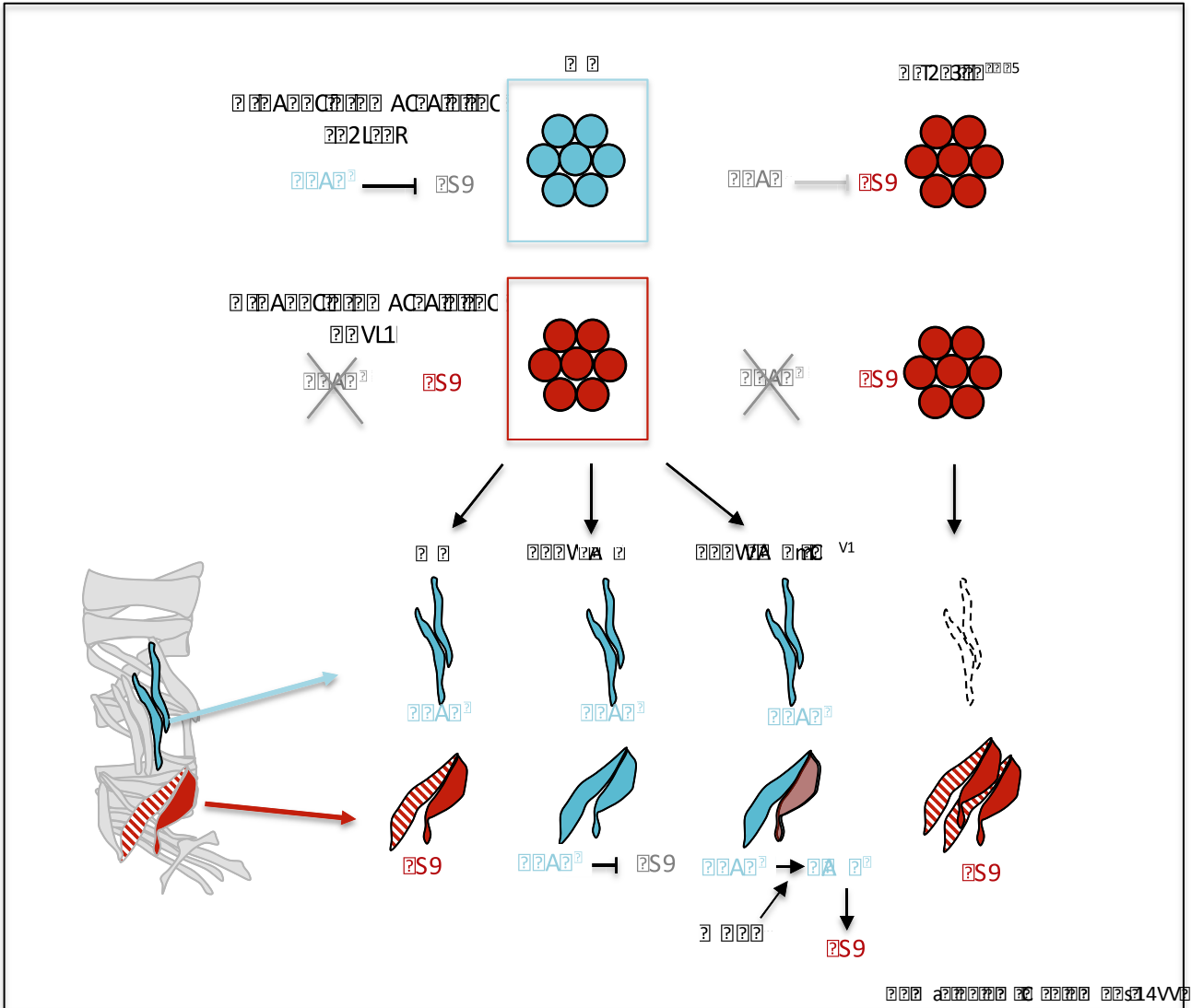
Dans les mutants nul pour *lms*, les phénotypes observés ne sont pas très forts, seulement un LT est généralement affecté au niveau de son attachement ou de sa structure alors que les autres sont toujours présents avec une morphologie normale (Figure 24). Seul les doubles mutants *lms/ap* présentent des défauts plus marqués avec l'absence régulière de 1 voir 2 LTs au niveau des hémisegments abdominaux (Müller et al., 2010).

Au final une faible pénétrance des phénotypes d'attachement ou de structure anormale des LTs chez les mutants *lms* suggèrent que *lms* joue un rôle en partie redondant dans l'acquisition du programme d'identité des muscles latéraux et l'acquisition de leurs caractéristiques finales.



### 2.2.2 Araucan/Caupolican

*Araucan* (*ara*) et *caupolican* (*caup*) sont deux gènes qui font partie du complexe *Iroquois* (*Iro-C*). Les gènes *Iro-C* codent pour des homéoprotéines conservées au cours de l'évolution (Gómez-Skarmeta et al., 1996). Elles participent à une grande variété de processus développementaux principalement liés à la spécification de plusieurs territoires au sein du corps (Cavodeassi et al., 2000, 2001). L'expression de *ara* et *caup* est détectée tout d'abord dans le mésoderme viscéral et dans l'ectoderme dorso-latéral puis à la fin du stade 11 dans plusieurs groupes promusculaires au sein du mésoderme. Leurs expressions sont ensuite maintenues dans les FCs à l'origine des muscles LT1-4, DT1 et SBM. Au niveau des muscles LTs, *ara* et *caup* sont présents tout au long de la spécification et de la différenciation de ces muscles (stade 11-15)(Carrasco-Rando et al., 2011). Afin de déterminer leurs rôles au sein des muscles, la déficience *Df(3L)iro<sup>EGP6</sup>* a été utilisée. Celle-ci supprime l'expression de *ara* et *caup*. Dans 95% des embryons portant la déficience les muscles LTs sont morphologiquement absents ou très affectés. Cette perte de muscles a été validée en regardant l'expression du gène spécifique des LTs, *lms*. L'expression de ce facteur est totalement perdue. L'absence de *ara* ou *caup* de manière individuelle n'a aucun impact sur le développement musculaire validant ainsi leurs fonctions redondantes au cours de l'embryogenèse. Dans les embryons déficients les FCs associées aux muscles LT3-4 expriment *Slou* qui est normalement exclusif des FCs du VA1-2. Dans ce contexte il y a transformation des LT3-4 en VA1-2. Cette analyse suggère que Ara/Caup agissent comme répresseurs de *Slou* dans les muscles LTs. Pour valider cela les auteurs ont recherché des sites de fixation de facteurs Iroquois sur les CRMs de *Slou* et en ont identifié deux à +129 et -1642 du TSS de *Slou*. Des expériences de dosage de luciférase en cellules ont montré qu'une augmentation de Caup favorise l'expression de la luciférase (sous le contrôle de la région *Slou* identifiée) prouvant que Caup est un activateur transcriptionnel de *Slou* (Carrasco-Rando et al., 2011). Cependant il est connu chez les vertébrés que les facteurs Iroquois passent d'un état répresseur à activateur en fonction de leur phosphorylation par la voie Ras/MAPK (Matsumoto et al., 2004). Pour vérifier cette hypothèse un inhibiteur de la voie Ras/MAPK, MEK1, a été ajouté et cela entraîne une diminution de l'activité de la luciférase prouvant que la régulation activatrice ou répressive de Caup sur *Slou* passe par sa phosphorylation. Afin de s'assurer de l'effet répresseur de Caup sur *Slou*, *ara* et *caup* ont été exprimé de manière ectopique dans les muscles *Slou* positifs VA1-2 (Con>Caup). Une perte



The diagram illustrates the embryonic development of the vertebral column and intervertebral discs. It shows the differentiation of somites from neural crest cells (blue) and somites (red), and their subsequent migration and differentiation into the axial skeleton (vertebrae and intervertebral discs).

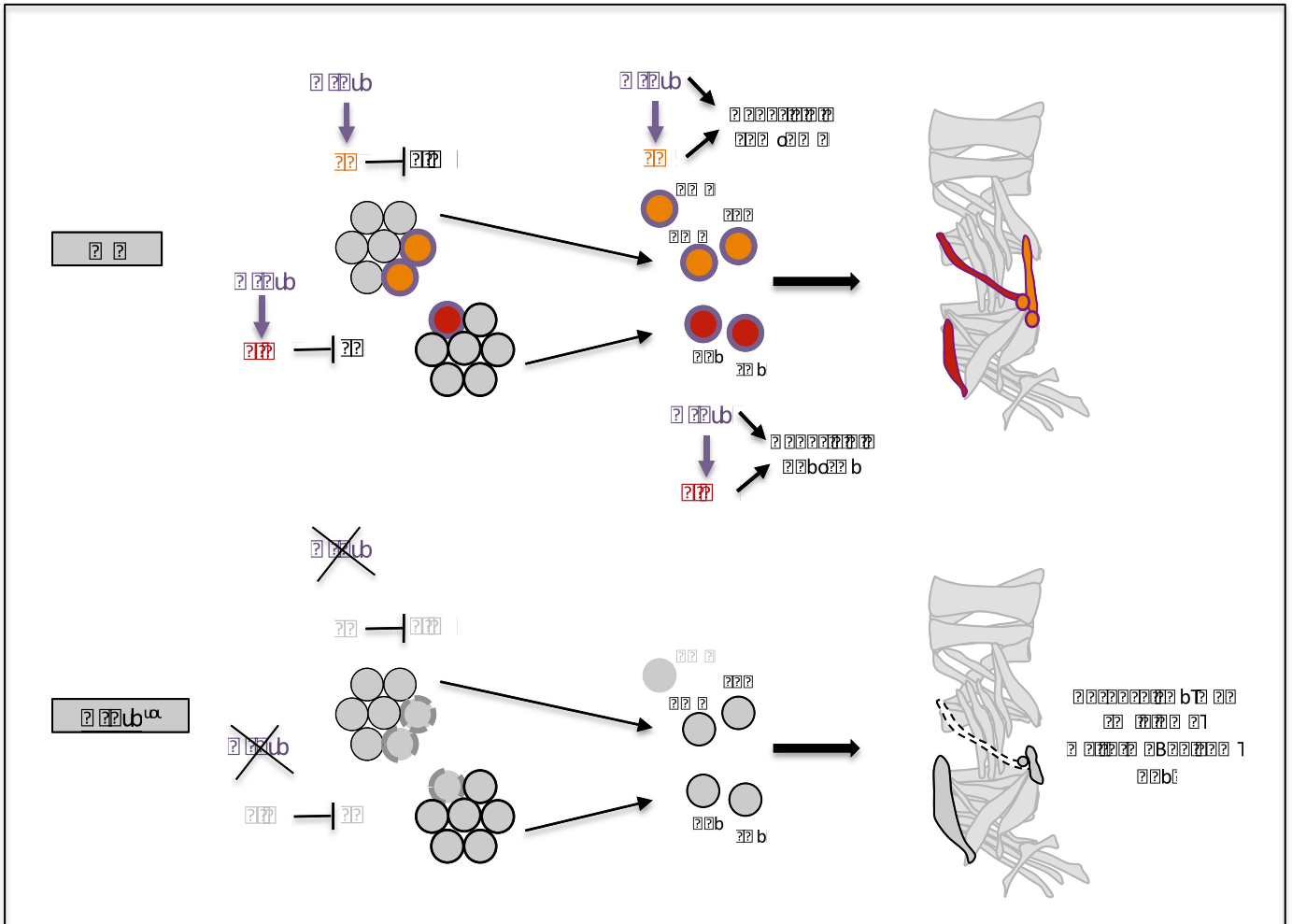
de l'expression de *Slou* ainsi que des défauts morphologiques du VA2 sont observés. Pour finir, la surexpression combinée de *Caup* et de la forme active de *Ras* (*Ras*<sup>V12</sup>) dans ces muscles (VA1-2) permet de lever la répression de *Caup* sur *Slou* dans le VA2 (Carrasco-Rando et al., 2011) (Figure 25).

Pour conclure les formes non phosphorylées de *Ara* et *Caup* sont nécessaires à la mise en place du programme d'identité dans les muscles LTs et cela se fait notamment par la répression de *Slou*. La phosphorylation de *Caup* dépendante de la voie MAPK pourrait jouer un rôle dans l'interaction de *Caup* avec d'autres cofacteurs. Au final le rôle des protéines Iroquois dans la spécification du destin cellulaire semble très conservé entre vertébrés et invertébrés.

### 2.3 Le facteur *Org-1*

Optomotor blind Related Gene 1 (*Org-1*) est un facteur de transcription de la famille T-Box. *Org-1* est exprimé en premier lieu dans le mésoderme viscéral puis il est restreint par la suite aux précurseurs des muscles viscéraux, aux FCs correspondantes aux muscles VT1, LO1 et le SBM et aux muscles alaires (Schaub et al., 2012). Par la suite il est retrouvé au niveau des muscles LO1, VT1 et le SBM. Son expression chevauche celle de *Slou* et *Lb* respectivement dans les muscles VT1, LO1 et le SBM (Jagla et al., 1998; Knirr et al., 1999). Dans les mutants nuls pour *org-1* le muscle LO1 et le SBM ne sont pas identifiables. Concernant le VT1 celui ci semble présenter des formes variables avec des sites d'attachement perturbés. Dans le SBM la perte de *org-1* engendre une absence total d'expression de la protéine *Lb*. Au sein du VT1 et du LO1 la perte de *Slou* est observée ce qui prouve que *Org-1* est un régulateur transcriptionnel en amont de *Lb* et *Slou*. La restauration de *Lb* ou *Slou* dans les muscles *org-1* mutés ne sauvent pas les phénotypes ce qui suggère que *Org-1* joue aussi un rôle important dans l'activation d'autres cibles en parallèle de *Slou* et *Lb* qui sont nécessaires au développement de ces muscles. La surexpression ectopique de *Org-1* dans tout le mésoderme induit le changement d'identité des FCs vers celle du muscle SBM et dans un degré inférieur vers celles du LO1 et du VT1. Des expériences de ChIP au stade embryonnaire 11-12 ont été réalisées et ont mis en évidence les sites de liaison de *Org-1* sur un enhancer de *slou* de 622 pb et sur un enhancer spécifique de *lb* SBM de 925 pb. La mutation de ces différents sites entraîne une perte d'expression de *Lb* dans le FC (SBM) et une inactivation de *Slou* dans les FCs correspondantes aux VT1 et LO1. *Org-1* est donc un



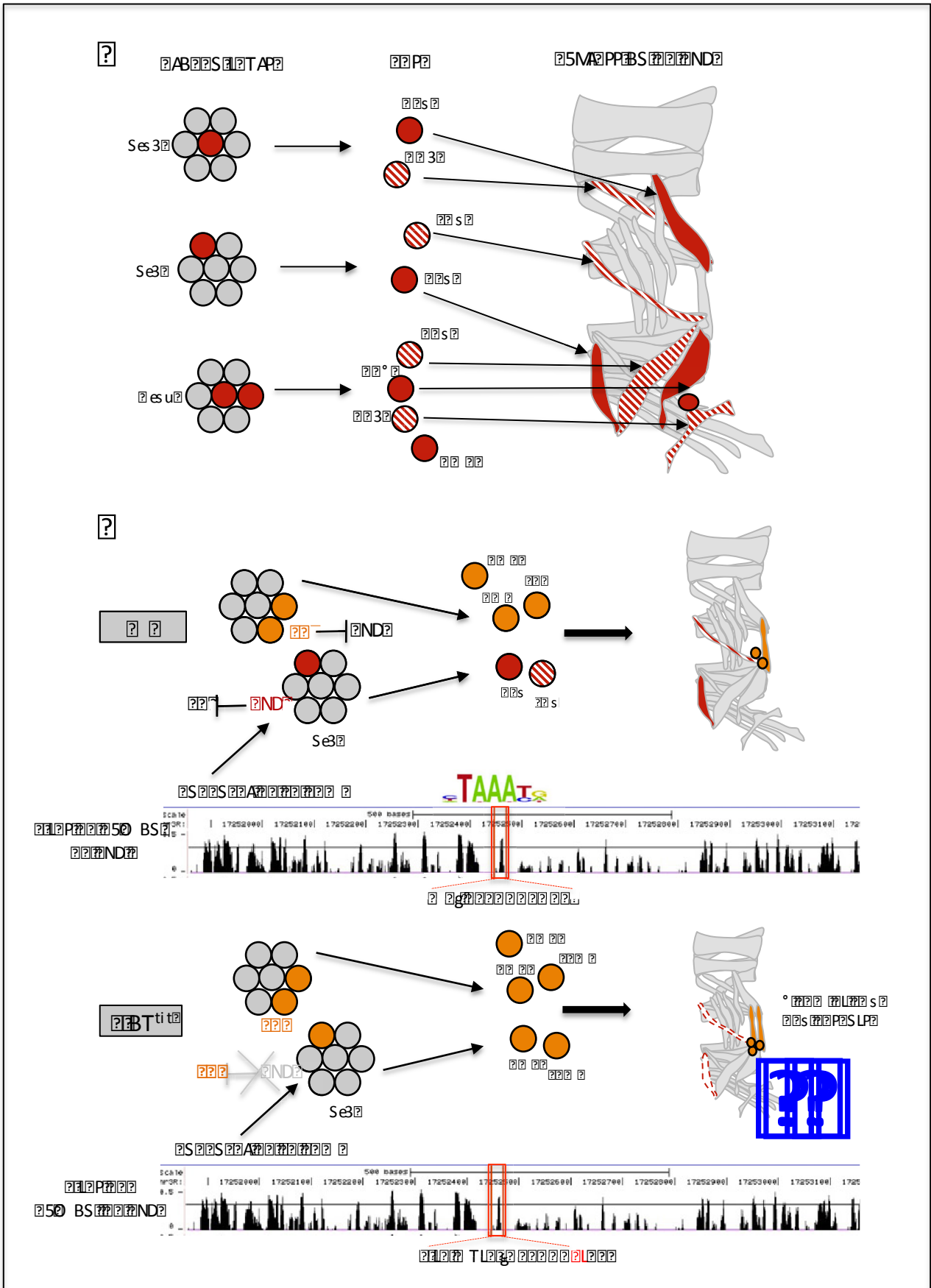


g 1990 2000 2010 2020 2030 2040 2050 2060 2070 2080 2090 2100 2110 2120 2130 2140 2150 2160 2170 2180 2190 2200 2210 2220 2230 2240 2250 2260 2270 2280 2290 2300 2310 2320 2330 2340 2350 2360 2370 2380 2390 2400 2410 2420 2430 2440 2450 2460 2470 2480 2490 2500 2510 2520 2530 2540 2550 2560 2570 2580 2590 2600 2610 2620 2630 2640 2650 2660 2670 2680 2690 2700 2710 2720 2730 2740 2750 2760 2770 2780 2790 2800 2810 2820 2830 2840 2850 2860 2870 2880 2890 2900 2910 2920 2930 2940 2950 2960 2970 2980 2990 3000 3010 3020 3030 3040 3050 3060 3070 3080 3090 3100 3110 3120 3130 3140 3150 3160 3170 3180 3190 3200 3210 3220 3230 3240 3250 3260 3270 3280 3290 3300 3310 3320 3330 3340 3350 3360 3370 3380 3390 3400 3410 3420 3430 3440 3450 3460 3470 3480 3490 3500 3510 3520 3530 3540 3550 3560 3570 3580 3590 3600 3610 3620 3630 3640 3650 3660 3670 3680 3690 3700 3710 3720 3730 3740 3750 3760 3770 3780 3790 3800 3810 3820 3830 3840 3850 3860 3870 3880 3890 3900 3910 3920 3930 3940 3950 3960 3970 3980 3990 4000 4010 4020 4030 4040 4050 4060 4070 4080 4090 4100 4110 4120 4130 4140 4150 4160 4170 4180 4190 4200 4210 4220 4230 4240 4250 4260 4270 4280 4290 4300 4310 4320 4330 4340 4350 4360 4370 4380 4390 4400 4410 4420 4430 4440 4450 4460 4470 4480 4490 4500 4510 4520 4530 4540 4550 4560 4570 4580 4590 4600 4610 4620 4630 4640 4650 4660 4670 4680 4690 4700 4710 4720 4730 4740 4750 4760 4770 4780 4790 4800 4810 4820 4830 4840 4850 4860 4870 4880 4890 4900 4910 4920 4930 4940 4950 4960 4970 4980 4990 5000 5010 5020 5030 5040 5050 5060 5070 5080 5090 5100 5110 5120 5130 5140 5150 5160 5170 5180 5190 5200 5210 5220 5230 5240 5250 5260 5270 5280 5290 5300 5310 5320 5330 5340 5350 5360 5370 5380 5390 5400 5410 5420 5430 5440 5450 5460 5470 5480 5490 5500 5510 5520 5530 5540 5550 5560 5570 5580 5590 5600 5610 5620 5630 5640 5650 5660 5670 5680 5690 5700 5710 5720 5730 5740 5750 5760 5770 5780 5790 5800 5810 5820 5830 5840 5850 5860 5870 5880 5890 5900 5910 5920 5930 5940 5950 5960 5970 5980 5990 6000 6010 6020 6030 6040 6050 6060 6070 6080 6090 6100 6110 6120 6130 6140 6150 6160 6170 6180 6190 6200 6210 6220 6230 6240 6250 6260 6270 6280 6290 6300 6310 6320 6330 6340 6350 6360 6370 6380 6390 6400 6410 6420 6430 6440 6450 6460 6470 6480 6490 6500 6510 6520 6530 6540 6550 6560 6570 6580 6590 6600 6610 6620 6630 6640 6650 6660 6670 6680 6690 6700 6710 6720 6730 6740 6750 6760 6770 6780 6790 6800 6810 6820 6830 6840 6850 6860 6870 6880 6890 6900 6910 6920 6930 6940 6950 6960 6970 6980 6990 7000 7010 7020 7030 7040 7050 7060 7070 7080 7090 7100 7110 7120 7130 7140 7150 7160 7170 7180 7190 7200 7210 7220 7230 7240 7250 7260 7270 7280 7290 7300 7310 7320 7330 7340 7350 7360 7370 7380 7390 7400 7410 7420 7430 7440 7450 7460 7470 7480 7490 7500 7510 7520 7530 7540 7550 7560 7570 7580 7590 7600 7610 7620 7630 7640 7650 7660 7670 7680 7690 7700 7710 7720 7730 7740 7750 7760 7770 7780 7790 7800 7810 7820 7830 7840 7850 7860 7870 7880 7890 7900 7910 7920 7930 7940 7950 7960 7970 7980 7990 8000 8010 8020 8030 8040 8050 8060 8070 8080 8090 8100 8110 8120 8130 8140 8150 8160 8170 8180 8190 8200 8210 8220 8230 8240 8250 8260 8270 8280 8290 8300 8310 8320 8330 8340 8350 8360 8370 8380 8390 8400 8410 8420 8430 8440 8450 8460 8470 8480 8490 8500 8510 8520 8530 8540 8550 8560 8570 8580 8590 8600 8610 8620 8630 8640 8650 8660 8670 8680 8690 8700 8710 8720 8730 8740 8750 8760 8770 8780 8790 8800 8810 8820 8830 8840 8850 8860 8870 8880 8890 8900 8910 8920 8930 8940 8950 8960 8970 8980 8990 9000 9010 9020 9030 9040 9050 9060 9070 9080 9090 9100 9110 9120 9130 9140 9150 9160 9170 9180 9190 9200 9210 9220 9230 9240 9250 9260 9270 9280 9290 9300 9310 9320 9330 9340 9350 9360 9370 9380 9390 9400 9410 9420 9430 9440 9450 9460 9470 9480 9490 9500 9510 9520 9530 9540 9550 9560 9570 9580 9590 9600 9610 9620 9630 9640 9650 9660 9670 9680 9690 9700 9710 9720 9730 9740 9750 9760 9770 9780 9790 9800 9810 9820 9830 9840 9850 9860 9870 9880 9890 9900 9910 9920 9930 9940 9950 9960 9970 9980 9990 10000

facteur clé pour l'identité et la diversification de sous populations musculaires et cela se fait notamment par le contrôle transcriptionnel de *lb* et *slou* (Schaub et al., 2012) (Figure 26). Pour conclure *Org-1* comme son orthologue humain *Tbx1* sont nécessaires à la spécification de sous groupes musculaires et l'identification de leurs cibles ou interacteurs permettrait de mieux comprendre comment ces muscles acquièrent leurs caractéristiques propres.

#### 2.4. Le facteur Slouch /S59

S59, appelé ensuite Slouch a été le premier facteur de transcription à homéodomaine de la famille NK découvert comme étant exprimé dans un sous groupe de cellules neuronales (CNS) et dans seulement quelques cellules précurseurs de muscle et FCs (Dohrmann et al., 1990). Slou est présent au niveau des FCs des muscles VA1-3, DT1, VT1, LO1, DO3 avant d'être restreint aux muscles VT1, DT1 et VA2 et à un AMP ventral (Carmena et al., 1995; Dohrmann et al., 1990) (Figure 27A). Au sein de ces différentes FCs il interagit avec d'autres iTFs tels que Msh, Ap ou encore Kr. Le mutant nul *slou* a été généré à travers une mutagenèse EMS (Ethyl methasulfonate) qui permet de produire des mutations génétiques aléatoires par substitution de nucléotides. En effet, à la position 983 de la séquence codante de *slou* le changement d'un C en T induit un codon stop prématuré à l'acide aminé 219 menant à une protéine tronquée (Knirr et al., 1999). Dans les mutants *slou*, les muscles VT1, VA3 et LO1 sont complètement absents alors que les muscles VA2 et VA1 présentent des morphologies aberrantes et variables. Le SBM semble lui dupliqué dans tous les segments (Jagla et al., 1998). En ce qui concerne les muscles dorsaux DO3 et DT1, un seul syncytium est observé ou peut être deux qui sont morphologiquement affectés. Au final, tous les muscles dérivant des FCs Slou positives sont touchés. De plus, lorsque Slou est exprimé de manière ectopique dans tout le mésoderme le patron des muscles ventraux, latéraux et dorsaux est fortement endommagé. Ces observations suggèrent que l'expression de Slou dans les muscles où il est normalement absent induit un changement d'identité de ces derniers. La perte des muscles LO1 et VT1 provenant du même précurseur semble être corrélée avec la duplication du SBM. Il a été montré que le précurseur du SBM provient de la même région du mésoderme somatique que le précurseur du VT1/LO1. Pour vérifier que la duplication du SBM dans les mutants *slou* est due à une expression ectopique de *Lb*, les auteurs ont suivi l'expression de *slou-LacZ* dans les mutants *slou*. Il s'avère que dans ce contexte *Lb* devient exprimé dans six FCs au lieu de quatre. L'expression supplémentaire de



The figure illustrates the differentiation of cell clusters into various cell types and their corresponding gene expression patterns. The top section shows three cell clusters (Se3, Se3, and es u) differentiating into various cell types (red, striped, and grey) which are mapped to a skeletal diagram. The middle section shows a cell cluster with a red cell and a grey cell, with arrows pointing to different cell types and a skeletal diagram. Below this is a genomic track with a TAAAT motif highlighted in red. The bottom section shows a similar cell cluster with a red cell and a grey cell, with arrows pointing to different cell types and a skeletal diagram. Below this is another genomic track with a TAAAT motif highlighted in red.

*Lb* se retrouve en effet dans les FCs associés aux muscles VT1/LO1. Ces FCs (*Lb* positives) surnuméraires induisent la formation d'un second SBM. Slou agit donc comme un répresseur de *Lb* dans les muscles VT1 et LO1 (Knirr et al., 1999) (Figure 27B).

Pour comprendre comment le facteur Slou participe à la régulation du programme génétique au sein de différents FCs, et notamment à la répression de *Lb* plusieurs expériences ont été réalisées. Tout d'abord le transcriptome au stade où les FCs sont spécifiées a été déterminé. Pour ce faire les cellules du mésoderme ont été isolées par cytométrie en flux à partir d'embryons sur-exprimant Slou (Busser et al., 2012). Environ 1400 gènes différentiellement exprimés de manière statistique (entre 1.5 et 4 fois plus) ont été identifiés. Les ontologies de gènes associés à ces derniers montrent que Slou régule aussi bien des facteurs de transcription que des gènes de différenciation terminale comme des protéines de structure du muscle. Par la suite une expérience de PBM (protein binding microarray) a permis de déterminer les motifs de liaison préférentiels de Slou sur ces cibles (Berger et al., 2006). Comme tous les facteurs à HD ce dernier se lie sur les séquences ayant un cœur TAAT, cependant il peut aussi se fixer sur une séquence TAAAT qui elle est plus spécifique (Berger et al., 2006; Busser et al., 2012; Noyes et al., 2008). En ce qui concerne l'enhancer de *lb* (Philippakis et al., 2006) ce dernier contient plus de 20 sites capables d'être fixés par Slou. Pour valider le rôle de Slou sur la régulation de l'enhancer de *lb*, une expérience de CHIP suivie de qPCR (quantitative real time polymerase chain reactions) a été réalisée. Slou est retrouvée fortement enrichie au niveau du site TAAAT de l'enhancer de *lb* (Figure 27B). Pour savoir si ce site est à l'origine de la répression de *lb* ce dernier a été muté. La mutation réduit l'enrichissement de Slou sur ce site et provoque la réexpression de *lb* dans les FCs LO1 et VT1 (Busser et al., 2012). Ces résultats démontrent que les facteurs HD peuvent se lier aux mêmes séquences ayant un cœur TAAT mais qu'ils ont aussi des sites préférentiels uniques leur permettant de réguler l'expression spécifique de certains gènes.

## 2.5 Le facteur dSix4

La spécification des cellules du mésoderme latéral et ventral semble être contrôlée en partie par le facteur de transcription dSix4. dSix4 fait partie de la famille de protéines à HD Six qui est composée de trois membres chez la drosophile, *Sine oculis*, *Optix* et *dSix4* (Seo et al., 1999). dSix4 est exprimé dans les cellules précurseurs des follicules, dans le corps gras mais aussi dans certaines cellules du muscle somatique en embryon (Kirby et al., 2001). Au



sein du mésoderme *dSix4* est retrouvé tout d'abord dans les domaines d'expression de *Mef2*. Ensuite au stade 10 il est restreint au mésoderme latéral et ventral (domaine d'expression de *Slp*) par *Dpp*. En effet *Dpp* inhibe *dSix4* car la surexpression de ce dernier dans le mésoderme réduit son expression (Clark et al., 2006). La mutation ponctuelle générée par EMS de *dSix4* induit un codon stop au niveau de l'AA 281 (arginine) induisant une perte de fonction de ce dernier (Kirby et al., 2001). Dans les mutants *dSix4* des défauts musculaires sont observés au niveau latéral et ventral. La partie latérale semble plus affectée avec une perte importante de muscles alors que dans la partie ventrale seulement certains muscles sont touchés. En effet le VA3 et le SBM sont absents alors que les VO et les VLs sont quasiment tous présents. Ces résultats suggèrent que l'expression précoce de *dSix4* dans le mésoderme ventro-latérale (avant spécification des FCs) aurait un impact sur la myogenèse dans cette région et sur la formation d'une partie des muscles latéraux. Pour vérifier le rôle myogénique de *dSix4* celui-ci a été surexprimé dans tout le mésoderme mais cela n'a pas affecté le destin cellulaire des FCs. Cela est probablement dû à l'absence de son cofacteur *Eya* dans ces muscles. En effet l'activité d'un autre Six, *Sine oculis* est dépendante de *Eya* pour le développement de l'œil chez la drosophile (Pignoni et al., 1997). L'expression de *Eya* au niveau du mésoderme latéral et ventral est la même que *dSix4*. De plus les phénotypes observés dans les mutants *Eya* sont identiques à ceux de *dSix4* (Boyle et al., 1997). Cela démontre que l'activité de *dSix4* est dépendante de son interaction avec *Eya*. Lorsque les deux sont surexprimés de manière simultanée dans le mésoderme, celui-ci est fortement perturbé. L'ensemble des muscles est affecté cependant les muscles dorsaux et une partie des ventraux qui normalement n'expriment pas *dSix4* sont plus particulièrement touchés. Pour conclure *dSix4* semble jouer un rôle équivalent de *Tin* dans la partie latérale et ventrale du mésoderme favorisant la spécification et la différenciation de certains muscles en accord avec son cofacteur *Eya* (Clark et al., 2006; Liu et al., 2009). La faible pénétrance des phénotypes observés chez les mutants *dSix4* pourrait être due en partie à une redondance avec *pox-meso* également exprimé précocement dans la même région du mésoderme. A ce jour, l'expression de *dSix4* dans les cellules fondatrices et/ou précurseurs musculaires et donc sa contribution au code de gènes d'identité n'ont pas été décrites. *dSix4* est l'orthologue des gènes murins *Six4* et *Six5*. Ces orthologues sont tous les deux exprimés au cours de la myogenèse et il a été montré que *Lbx1* (orthologue de *lb*) est directement régulé par ces facteurs pour la détermination des muscles hypaxiaux (Grifone, 2005). Cela



suggère une conservation des mécanismes de régulation entre dSix4/Six et lb/Lbx1 au cours de la myogenèse entre la drosophile et la souris. L'identification des gènes régulés par dSix4 ainsi que son rôle potentiel dans la diversification musculaire chez la drosophile pourrait apporter une meilleur compréhension des processus qui gouvernent l'acquisition de l'identité musculaire chez les vertébrés.

## Conclusion

La mise en place du programme d'identité au sein de chaque FCs est une étape clé pour l'acquisition des propriétés individuelles (taille, sites d'attachement, nombre d'évènements de fusion) des 30 muscles par hémisegment. Comme démontré précédemment la combinaison et la séquence d'expression des iTFs est très complexe. En effet l'activation spatio-temporelle de ces facteurs est finement régulée par des signaux extrinsèques et intrinsèques. Certains ne sont présents que dans la partie dorsale (Tup, Col) ou latérale (Lms, Lb, Ara/Caup) alors que d'autres sont exprimés plus largement (Slou, Org-1, Kr). De plus, ils n'ont pas tous la même fonction au cours du développement musculaire embryonnaire. Quelques uns sont requis précocement pour la spécification, avec une expression transitoire, alors que d'autres sont maintenus tout le long jusqu'à la différenciation terminale. Les iTFs peuvent aussi se réguler négativement ou positivement entre eux (Slou, Lb, Kr) ou bien s'autoréguler (Col, Tup). Certains sont décisifs alors que d'autres participent au maintien de l'identité.

De manière intéressante la dérégulation des iTFs n'entraîne que rarement des transformations de muscles de manière à ce que l'expression des gènes et les propriétés morphologiques et fonctionnelles soient changées d'un muscle en un autre. Cela renforce l'hypothèse du code combinatoire où la coordination spatio-temporelle de l'expression des gènes d'identité est requise pour spécifier les caractéristiques finales du muscle.

Actuellement seules quelques études à l'échelle du génome ont permis d'identifier les cibles directes de certains iTFs (Busser et al., 2012; Junion et al., 2007; Sandmann et al., 2007). Ces dernières ont validé le fait que les iTFs régulent des gènes d'identité mais contrôlent aussi les étapes d'acquisition des propriétés individuelles en régulant par exemple l'expression de gènes de structures (ex : *if*, *Sls*, *Msp300*).

La majorité des iTFs ont des orthologues vertébrés connus pour participer à la myogenèse crâniofaciale ou des membres (Tzahor, 2009). C'est le cas de *Islet1*, l'orthologue





vertébré de *Tup* qui est nécessaire à la formation de quelques muscles de la mâchoire. Le facteur *Tbx1* (orthologue de *Org-1*) requis pour la spécification de différents muscles branchiomériques. On retrouve aussi *Lbx1* et *Msx1/2* les orthologues respectifs de *Lb* et *Msh* qui sont nécessaires à la spécification des muscles des membres et bien d'autres encore. Cela prouve que les gènes d'identités ainsi que leurs mécanismes d'action semblent bien conservés entre les espèces. Chez la drosophile l'étude de la fonction de ces gènes ainsi que des acteurs en aval reste un enjeu majeur non pas seulement pour la compréhension des processus génétiques et moléculaires qui gouvernent l'identité musculaire mais aussi pour mieux appréhender les pathologies musculaires affectant des types particuliers de muscles.

A ce jour un grand nombre de questions restent encore en suspens. Bien que la combinaison de iTFs soit nécessaire à la formation des 30 muscles individuels y a-t-il une hiérarchie au sein des différentes FCs ? Est-ce que chaque iTF contrôle un ensemble particulier de gènes permettant de gérer les différents processus muscle spécifique ? Bien qu'un certain nombre de iTFs et leurs cibles ont été découverts, il est nécessaire d'avoir une vision cellule spécifique (muscle spécifique) pour découvrir de nouveaux acteurs du code combinatoire ainsi que les effecteurs qui contrôlent les caractéristiques individuelles des muscles *via* les mécanismes clés du développement comme la fusion, l'attachement, ou l'innervation.

Pour obtenir cette vision cellule spécifique, plusieurs méthodes ciblant les différentes étapes clés de la régulation des gènes (transcriptionnelle et traductionnelle) ont été développées. Par exemple le Bits-Chip (Batch isolation of tissue specific chromatin for immunopurification), Le FANS (Fluorescent activated nuclei sorting) ou encore l'INTACT (Isolation of nuclei tagged in specific cell types) permettent d'isoler les noyaux d'une population cellulaire d'intérêt (Bonn et al., 2012; Deal and Henikoff, 2010; Haenni et al., 2012; Henry et al., 2012) rendant possible l'analyse des ARN naissants, des sites de fixations des facteurs de transcription, des modifications post-traductionnelles des histones ou encore l'observation de la conformation tridimensionnelle de la chromatine (Ghavi-Helm et al., 2014).

D'autres techniques permettent de définir le profil d'expression des ARN messagers de manière cellule spécifique. C'est le cas de la méthode PABP (PolyA binding protein tagging), du TRAP (Translating Ribosome Affinity Purification) ou encore du TU-tagging (Heiman et al., 2008; Miller et al., 2009; Yang, 2005). La méthode PABP permet de récupérer



uniquement les ARN polyadénylés, le TRAP lui capture les ARN messagers associés aux ribosomes (en cours de synthèse protéique) et le TU-tagging favorise l'observation de la cinétique de synthèse et de dégradation des ARN messagers.

Pour choisir la méthode la plus adéquate il est nécessaire de se poser différentes questions qui sont: quelle est la question biologique à laquelle on souhaite répondre, les limites du matériel biologique et les ressources en équipement nécessaire à la technique. Pour répondre à notre question biologique principale qui est d'identifier des ensembles de gènes participant à l'acquisition des propriétés individuelles de différentes sous populations musculaires au cours du développement embryonnaire nous avons choisit la technologie TRAP. Celle-ci permet d'avoir une meilleure estimation de l'abondance protéique au sein de ces populations. Je reviendrai en détail sur cette technique ainsi que sur sa mise en place.



?

?

?

?

?

?

?

?

?

?

?

?

# U.S. Autism

?

?

?

?

?

?



A ce jour, découvrir comment les régulateurs de l'identité musculaire sont connectés entre eux à travers un réseau de gènes spécifiant le destin cellulaire et les propriétés fonctionnelles de chaque type de muscle reste un enjeu majeur. En effet malgré de nombreuses découvertes concernant la mise en place d'une combinaison de signaux moléculaires et de facteurs de transcription pour l'acquisition des propriétés musculaires (Busser et al., 2013; Philippakis et al., 2006), toutes les approches mises en place pour comprendre le contrôle génétique de cette diversification au cours de l'embryogenèse sont essentiellement fondées sur des analyses fonctionnelles, génomiques ou bioinformatiques basées sur des données collectées à partir d'études sur l'embryon entier à des temps de développement précis. Ce type d'approche comporte des limitations et ne permet pas d'identifier les réseaux de gènes qui gouvernent l'acquisition des propriétés individuelles des cellules musculaires. Pour comprendre comment la diversité musculaire est mise en place au cours de l'embryogenèse il est nécessaire d'utiliser des approches cellulaires spécifiques sur des fenêtres développementales bien définies.

L'objectif principale de ce projet de thèse était de définir des **signatures d'expression géniques qui spécifient différentes sous populations de muscles** à différents stades de développement embryonnaire afin d'identifier de nouveaux acteurs de la diversification musculaire.

Pour cela j'ai optimisé et utilisé la méthode cellule spécifique TRAP pour capturer les ARN messagers en cours de traduction dans deux populations musculaires restreintes correspondantes aux cellules Slou et Lms positives et dans la population musculaire globale (cellules Duf positives). En utilisant les données d'expression génique obtenues par l'approche TRAP sur la population hétérogène (Slou) et relativement homogène (Lms) je me suis attaché: i) à identifier de nouveaux gènes de régulation de l'identité codant pour des facteurs de transcription, des protéines de liaison à l'ARN ou encore pour des modificateurs de la chromatine et ii) à mettre en évidence et caractériser de nouveaux acteurs qui assurent l'acquisition de propriétés individuelles de muscles au cours des étapes tardives de la myogenèse.





??seG?us?



Dans le but d'identifier de nouveaux acteurs de l'identité musculaire nous avons utilisé l'approche génomique cellule-spécifique TRAP (Translating Ribosome Affinity Purification) pour cibler deux populations de muscles différentes : i) la population Slou qui correspond à 6 muscles par hémisegment (DT1, LO1, VA1-3, VT1) et ii) la population Lms qui regroupe les 4 muscles latéraux transverses (LT1-LT4). Nous avons également analysé la population Duf qui représente l'ensemble de la musculature somatique (30 muscles/hémisegment). Pour identifier des gènes qui assurent l'acquisition de propriétés individuelles des deux sous populations musculaires nous avons effectué nos analyses à travers trois fenêtres de temps différentes correspondant au stade 11-12 (7-10h AEL), au stade intermédiaire 13-15 (10-13h AEL) et au stade tardif 16 (13-16h AEL). Chacune de ces fenêtres temporelles correspond à une ou plusieurs étapes clés de la myogenèse. La spécification des cellules fondatrices et la première vague de fusion prennent place entre 7 et 10h AEL (fenêtre T1). La deuxième étape de fusion, élongation des myotubes et la reconnaissance avec les cellules tendineuses suivent pendant la fenêtre T2 (10-13h AEL) et finalement au cours de la fenêtre T3 (13-16h AEL) on observe la maturation des jonctions myotendineuses, le guidage des motoneurones et la formation des jonctions neuromusculaires. Pour obtenir les données d'expression génique spécifiques aux sous-populations musculaires il a tout d'abord été nécessaire d'adapter et d'optimiser la méthode TRAP pour chaque condition. Cette étape préliminaire aux analyses et essentielle pour l'obtention de données de qualité a occupé une partie importante de ma thèse.

## **A – Mise en place et validation de la méthode TRAP**

### **I. La méthode TRAP**

La méthode TRAP a été décrite en 2008 par Heiman et al., et a été appliquée pour identifier le transcriptome d'une sous population de neurones du cerveau de souris. Cette approche consiste à exprimer de manière cellule spécifique le transgène Rpl10a-EGFP. Celui-ci code pour une protéine de la sous unité 60S du ribosome, Rpl10a (Ribosomal protein L10a), fusionnée à l'EGFP (Enhanced Green Fluorescent Protein). Cette protéine ribosomale a été sélectionnée à travers une douzaine d'autres testées car sa localisation nucléolaire et cytoplasmique était cohérente avec celle des ribosomes et parce qu'elle est retrouvée en microscopie électronique associée aux polysomes. De plus sa localisation périphérique facilite son immunoprécipitation via son étiquette GFP.

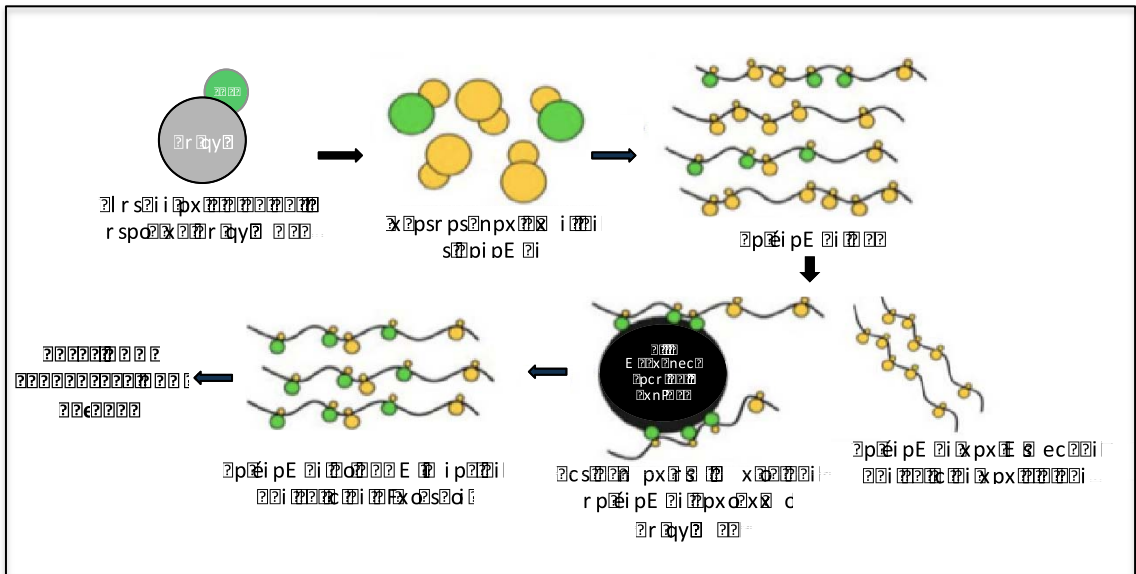
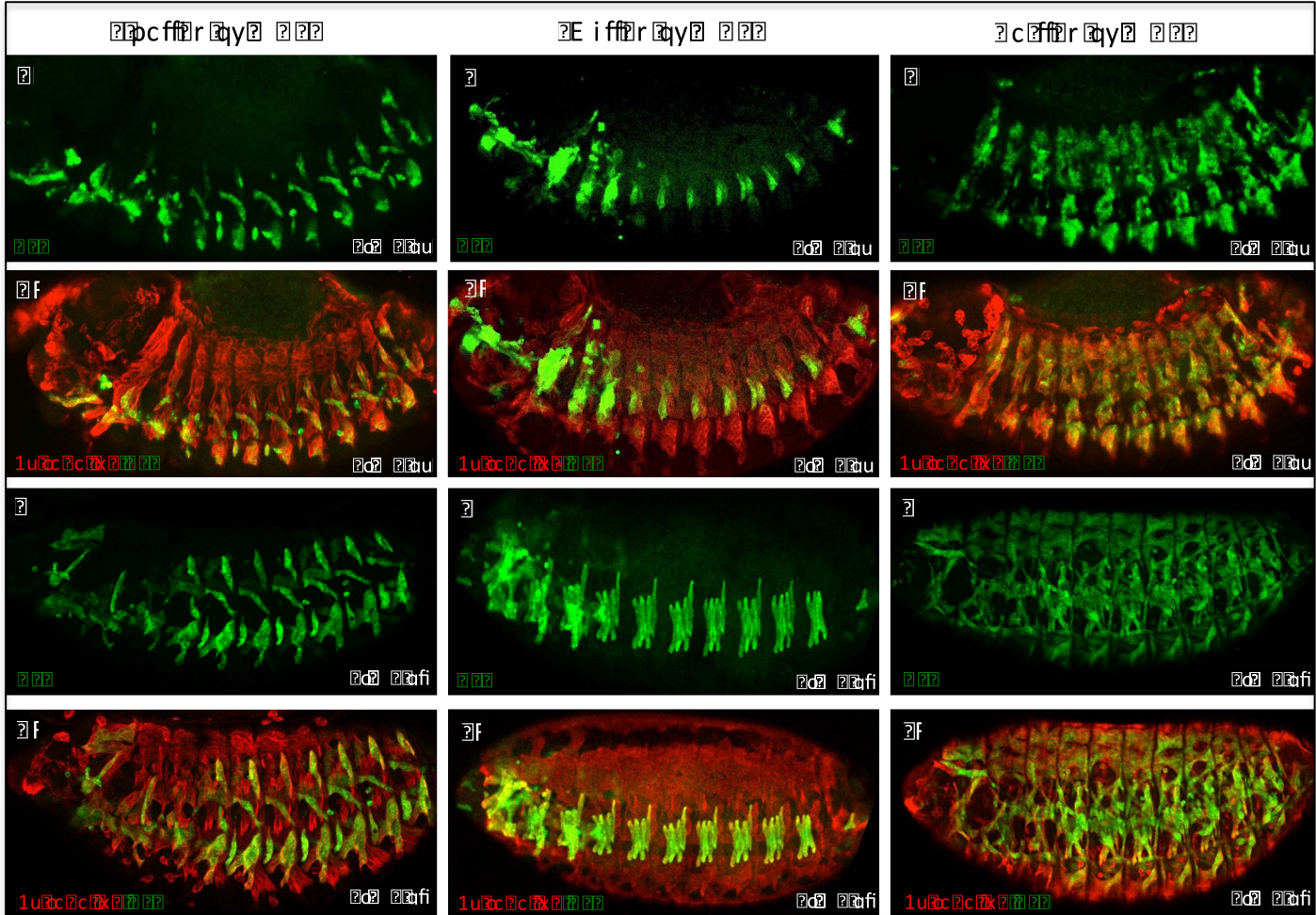


Diagram illustrating the process of DNA replication and transcription. The diagram shows a circular DNA molecule (top left) with a green origin of replication. An arrow points to a replication bubble (top middle) where two replication forks are moving outwards. Another arrow points to a replication fork (top right) where a new DNA strand is being synthesized. A downward arrow points to a replication fork with a replication bubble (middle right) where a replication fork is moving towards a replication bubble. A leftward arrow points to a replication fork (middle left) where a replication fork is moving towards a replication bubble. A final leftward arrow points to a replication bubble (bottom left) where a replication bubble has formed.



Fluorescence microscopy images showing DNA replication and transcription in cells. The grid is organized into three columns: 'pccffir qy? ???', 'E iffir qy? ???', and 'c ffr qy? ???'. Each column has four rows of images. The first row shows green fluorescence. The second row shows red fluorescence. The third row shows green fluorescence. The fourth row shows red fluorescence. The images show the progression of DNA replication and transcription over time.

Une fois que cette protéine est incorporée aux polysomes uniquement dans les cellules d'intérêt, différentes étapes dont l'immunopurification (IP) à l'aide d'un anticorps anti-EGFP sont nécessaires pour capturer les ARN messagers associés aux ribosomes (EGFP) dans la population ciblée (Heiman et al., 2008). Une fois les ARNs isolés ces derniers peuvent être analysés par des méthodes classiques d'étude de l'expression de transcrits telles que le « Northern-blot, la PCR quantitative, les puces à ADNc (ADN complémentaire) ou encore le séquençage ARN (Figure 28). Les avantages du TRAP sont qu'aucune étape de fixation, de dissociation ou encore d'extraction de cellules à partir d'un tissu ne sont nécessaires pour la capture des ARN messagers spécifiques d'un type cellulaire. De plus, par rapport à d'autres méthodes de visualisation de l'expression génique, le TRAP révèle de manière plus précise les ARNs traduits se rapprochant ainsi plus de la composition protéique finale d'une cellule que l'ARN total. Au final cette technique permet d'avoir une meilleure estimation de la production de protéine dans un tissu à un temps donné. Le TRAP a été utilisé avec succès pour étudier le transcriptome spécifique chez les plantes (Jiao and Meyerowitz, 2010), la drosophile (Thomas et al., 2012), le xénope (Watson et al., 2012) ou encore le poisson zèbre (Tryon et al., 2013).

### 1.1 Création des lignées TRAP

Pour adapter cette méthode à l'embryon de drosophile il a fallu tout d'abord créer les lignées de drosophiles associées à chaque population d'intérêt. Pour cibler ces différents ensembles de muscles nous avons utilisé le système d'expression tissu spécifique GAL4/UAS (Upstream Activating Sequence) (Brand and Perrimon, 1993). Les lignées pilotes Slou-Gal4, Lms-Gal4 et Duf-Gal4 ont été croisées à celle portant UAS-Rpl10a-EGFP pour établir les trois lignées stables. Des immunomarquages ont ensuite été réalisés pour vérifier le profil d'expression de la protéine Rpl10a-EGFP dans ces trois populations au cours du temps. Pour Slou, celle-ci est bien retrouvée dans 6 muscles par hémisegment (DT1, LO1, VT1, VA1-3), pour Lms dans les quatre muscles latéraux transverses (LT1-4) et pour Duf dans tous les muscles somatiques du stade 11/12 à 16 (Figure 29). Par la suite chaque lignée a été amplifiée afin de récolter une quantité suffisante d'embryons pour chaque stade de développement ciblé (Stade 11-12, 13-15 et 16).



## 1.2 Optimisation de la technique

Afin d'obtenir un protocole robuste permettant d'isoler des ARNs de bonne qualité à partir de cellules à faible représentativité, avec une très bonne spécificité et sensibilité nous avons optimisé les étapes majeures de la technique. En effet celle-ci repose sur plusieurs facteurs clés, que sont la quantité d'embryons de départ, la méthode de lyse embryonnaire, la densité protéique, la qualité de l'anticorps ainsi que la concentration utilisée pour l'immunopurification et enfin la composition des tampons de lavage. Nous avons réalisé plusieurs tests dans différentes conditions avec les embryons Slou>RpL10a-EGFP (Slou-Gal4 ; UAS-RpL10aEGFP).

La quantité d'embryons de départ est primordiale car en ciblant seulement une sous population musculaire le rendement des ARN associés aux ribosomes peut s'avérer insuffisant pour réaliser des analyses quantitatives globales des transcrits présents. Pour s'assurer un rendement en ARN conséquent nous avons testé plusieurs quantités : 500 mg, 1g et 1,5 g d'embryons. Seule la condition à 1,5 g nous a permis d'isoler suffisamment d'ARNs pour faire des analyses supplémentaires. Le choix de la méthode de lyse est aussi important. Nous avons testé le broyage à l'azote liquide, au piston (Dounce) et le broyeur homogénéiseur Precellys (Bertin Instruments). La méthode Precellys est la plus rapide et la plus efficace. Elle permet de récupérer plus d'ARNs et surtout de meilleure qualité. Cette dernière consiste à placer une suspension d'embryons repris dans un tampon d'extraction de polysomes dans un tube contenant des billes en céramique. Les tubes sont ensuite agités selon un programme spécifique (ici 2x10 secondes à 5000 Rpm (tour par minute) avec 15 secondes de pause entre les deux « runs »). Cette agitation permet aux billes de faire éclater les cellules dans un temps bref limitant ainsi la dégradation des ARNs. Ensuite pour la partie immunopurification (IP), qui est une des étapes sensibles de la technique, les anticorps anti-EGFP Torrey Pines (#TP401), Invitrogen (#G10362), et Neuromab (Clone 86/8, #P42212) ont été utilisés à deux concentrations différentes, 15 et 30 µg. Après chaque IP des « western blot » ont été réalisés afin de déterminer quel anticorps précipitait le mieux la protéine RpL10a-EGFP. L'anticorps Neuromab (30 µg) a révélé la présence d'une plus forte bande pour RpL10a-GFP il a donc été sélectionné comme anticorps de référence pour le reste des expériences. Pour finir, différents tampons de lavage avec des concentrations croissantes en KCl (Chlorure de Potassium) ont été testés. Le tampon le plus stringent a permis d'éliminer le bruit de fond et d'améliorer la spécificité des ARNs isolés. L'établissement de ce nouveau



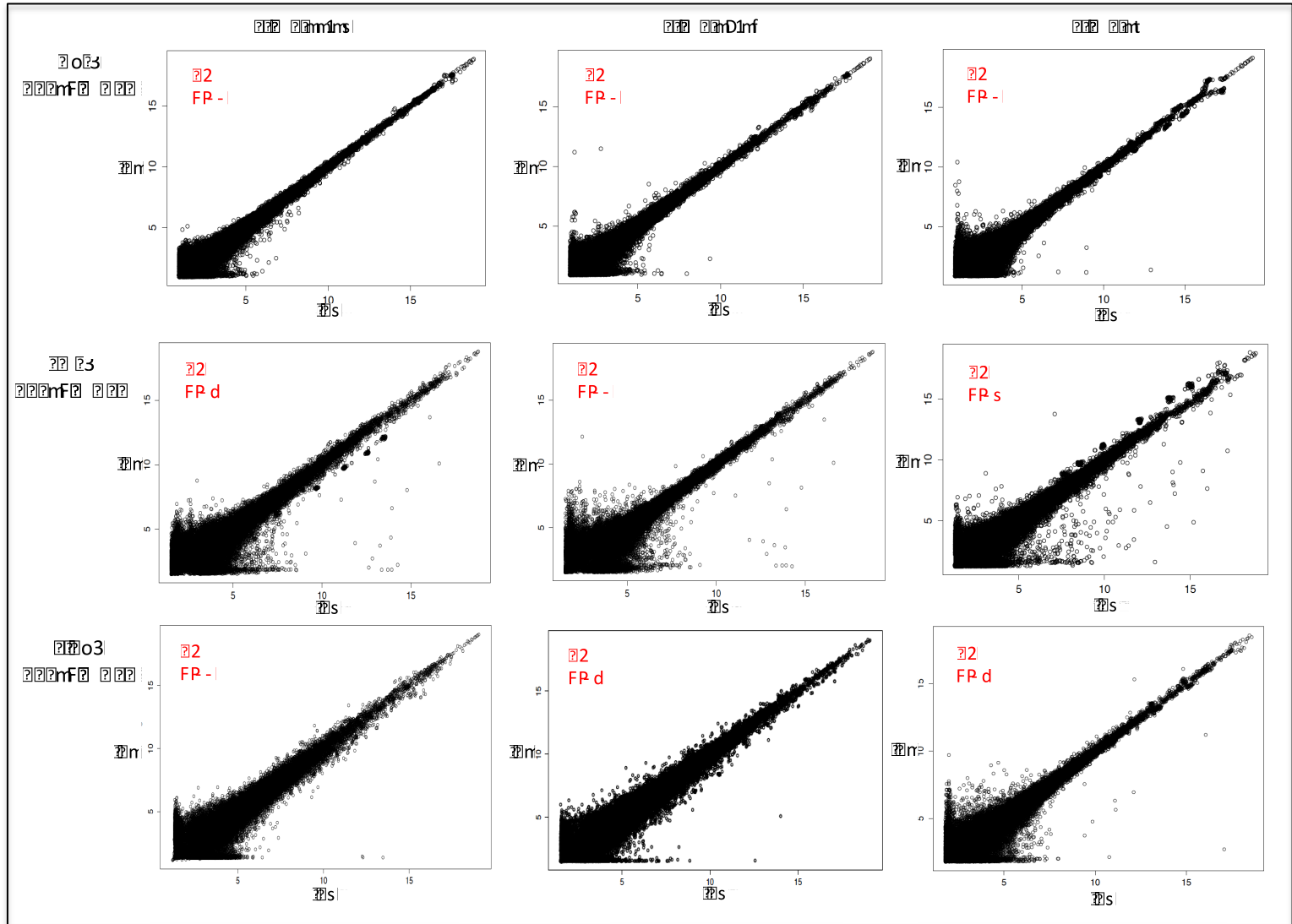


protocole a fait l'objet d'une publication dans le journal JoVe (Journal of visualized experiments) (Bertin et al., 2015)(Voir Publication 1, page 66) .

### *1.3 Qualité, quantité et spécificité des ARNs isolés*

Après chaque expérience de TRAP il a été nécessaire de valider la qualité, la quantité ainsi que la spécificité des ARNs capturés. Pour cela nous avons fait tout d'abord une analyse de nos échantillons au Bioanalyzer. Cet appareil permet grâce à une électrophorèse d'estimer la taille, la quantité et la pureté des échantillons d'ARN. Cette méthode est basée sur la détection des ARNs ribosomiaux 18S et 28S. Sachant que ces ARNs représentent 95% des ARN totaux si ces derniers ne sont pas dégradés l'ensemble des ARNs est considéré de bonne qualité. Dans le cas où ils ne sont pas clairement identifiables la qualité des ARNs n'est pas suffisante pour poursuivre les analyses. Comme le montre la [Figure 30A](#) les ARNs messagers immunopurifiés dans nos conditions d'expérience sont en parfaite intégrité. En ce qui concerne la quantité d'ARNs finale après IP celle-ci est comprise en fonction des stades et des lignées entre 25-100 ng. L'efficacité de ce rendement dépend en effet fortement de la force du pilote Gal4.

Pour évaluer le bruit de fond et la spécificité des ARNs des différentes populations des RT-qPCR (reverse transcription – quantitative polymerase chain reaction) ont été réalisées sur les trois réplicats de chaque stade en prenant comme contrôles positifs deux gènes connus pour être exprimés dans le muscle et deux gènes contrôles négatifs exprimés seulement dans le système nerveux. Les résultats de qPCR sont normalisés grâce à la mesure des transcrits du gène de ménage Rp49 et par comparaison avec l'« input » qui correspond au lysat embryonnaire avant IP. L'expression relative pour chaque gène testé est ensuite calculée. Les qPCR réalisées pour chaque population montrent un enrichissement des transcrits de gènes musculaires et un appauvrissement des transcrits de gènes du système nerveux ([Figure 30B](#)). Ces résultats confirment la spécificité et la robustesse de cette technique dans la capture d'ARN messagers de sous populations de muscles embryonnaires. Bien que la qPCR soit une technique fiable elle ne permet pas d'avoir une vue d'ensemble de l'expression des gènes dans ces trois groupes de muscles. C'est pourquoi les ARN isolés ont été envoyés sur la plateforme génomique de l'EMBL (Heidelberg, Allemagne) pour être hybridés sur des puces ADNc Agilent 8x60000 sondes (60-mers par sonde). Pour chaque



FP - | FP d FP s  
 S FP - | FP d FP s

population trois réplicats ont été préparés et analysés par stade ce qui correspond au final à 54 échantillons (27 IP et les 27 inputs correspondants)

## II. Validation des résultats obtenus par puce ADNc

Une fois les puces ADNc hybridées et les données obtenues et normalisées, les coefficients de corrélation ont été calculés pour les trois réplicats IP et « input » d'un même stade. L'ensemble des graphiques de corrélation présentent un coefficient supérieur à 0,9 ce qui signifie que les échantillons sont très similaires et peuvent donc être utilisés pour des analyses complémentaires (figure 31). Pour identifier les gènes enrichis et appauvris dans les populations ciblées les signaux provenant de l'IP ont été normalisés sur leurs correspondants en input. Un seuil supérieur à 2 et une p-value inférieure à 0,05 ont été choisis pour déterminer les gènes enrichis.

Pour vérifier la spécificité des gènes enrichis j'ai contrôlé l'enrichissement de certains gènes (contrôles internes) connus pour être exprimés respectivement dans les populations Duf, Lms et Slou et inversement j'ai vérifié que les gènes connus pour en être absents étaient appauvris par rapport au seuil fixé. Par exemple, toutes les sondes associées aux gènes normalement présents dans la population Slou sont bien enrichies aux différents stades de développement (Tableau 1).

Pour valider les résultats j'ai également analysé l'ontologie des gènes (GO) enrichis et déplétés au sein des différentes populations grâce au logiciel GO Term Finder (Boyle et al., 2004). Pour la population Slou la grande majorité des GO des gènes enrichis sont liées au développement musculaire avec notamment des processus comme la fusion retrouvée au stade 11-12 (T1), la différenciation des myotubes au stade 13-15 (T2) et l'organisation sarcomérique au stade 16 (T3) ce qui correspond parfaitement à leur temps d'apparition au cours du développement musculaire embryonnaire (Figure 32). Concernant les gènes appauvris les GOs sont associés au développement des neurones, à la communication cellulaire ou encore au transport.

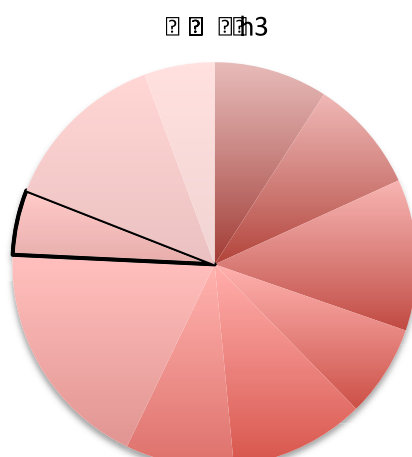
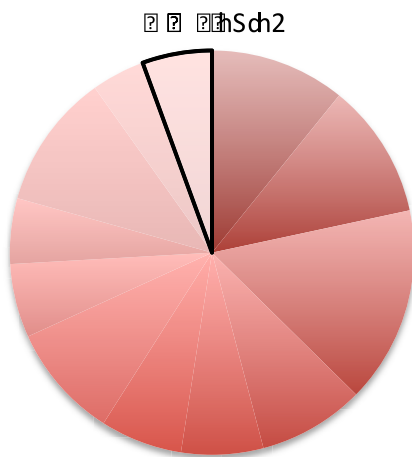
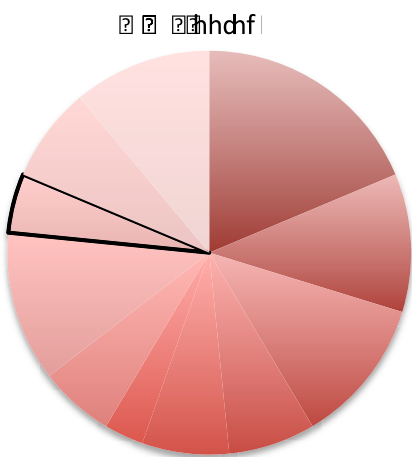
A travers tous ces résultats nous pouvons facilement confirmer la spécificité de la méthode TRAP. Les mêmes résultats ont été obtenus pour la population Lms et Duf attestant d'une robustesse de la méthode TRAP après son optimisation.

Category	Group 1	Group 2	Group 3
u	u	s	
h	hf	f	f
ca	3f	2f	p
o	Sf	2f	1f
s	Sf	2f	2f
a	f	f	Sf
e	p	p	f
n	2f	Sf	f
k	hfff	3f	Sf

Category	Group 1	Group 2	Group 3
t	hf	hf	hfff
f	df	df	d
d	df	df	d
a	df	df	df
e	df	hf	d
n	p	hf	df

Source: Author's calculation

The results of the analysis show that the most common errors in the use of the letter 'h' are related to the omission of the letter 'h' in the initial position of the word (e.g., 'u' instead of 'uh'). The most common errors in the use of the letter 'f' are related to the omission of the letter 'f' in the initial position of the word (e.g., 's' instead of 'sf'). The most common errors in the use of the letter 'p' are related to the omission of the letter 'p' in the initial position of the word (e.g., 'ca' instead of 'cap').



- omission of the letter 'h' in the initial position of the word
- omission of the letter 'h' in the middle position of the word
- omission of the letter 'h' in the final position of the word
- omission of the letter 'h' in the initial position of the word
- omission of the letter 'h' in the middle position of the word
- omission of the letter 'h' in the final position of the word
- omission of the letter 'h' in the initial position of the word
- omission of the letter 'h' in the middle position of the word
- omission of the letter 'h' in the final position of the word
- omission of the letter 'h' in the initial position of the word
- omission of the letter 'h' in the middle position of the word
- omission of the letter 'h' in the final position of the word

- omission of the letter 'f' in the initial position of the word
- omission of the letter 'f' in the middle position of the word
- omission of the letter 'f' in the final position of the word
- omission of the letter 'f' in the initial position of the word
- omission of the letter 'f' in the middle position of the word
- omission of the letter 'f' in the final position of the word
- omission of the letter 'f' in the initial position of the word
- omission of the letter 'f' in the middle position of the word
- omission of the letter 'f' in the final position of the word
- omission of the letter 'f' in the initial position of the word
- omission of the letter 'f' in the middle position of the word
- omission of the letter 'f' in the final position of the word

- omission of the letter 'p' in the initial position of the word
- omission of the letter 'p' in the middle position of the word
- omission of the letter 'p' in the final position of the word
- omission of the letter 'p' in the initial position of the word
- omission of the letter 'p' in the middle position of the word
- omission of the letter 'p' in the final position of the word
- omission of the letter 'p' in the initial position of the word
- omission of the letter 'p' in the middle position of the word
- omission of the letter 'p' in the final position of the word
- omission of the letter 'p' in the initial position of the word
- omission of the letter 'p' in the middle position of the word
- omission of the letter 'p' in the final position of the word

The results of the analysis show that the most common errors in the use of the letter 'h' are related to the omission of the letter 'h' in the initial position of the word (e.g., 'u' instead of 'uh'). The most common errors in the use of the letter 'f' are related to the omission of the letter 'f' in the initial position of the word (e.g., 's' instead of 'sf'). The most common errors in the use of the letter 'p' are related to the omission of the letter 'p' in the initial position of the word (e.g., 'ca' instead of 'cap').

### **III. Publication 1 « TRAP-rc, Translating Ribosome Affinity Purification from Rare Cell Populations of Drosophila Embryos »**

Dans cet article nous avons décrit toutes les étapes nécessaires à l'isolation des ARN messagers en cours de traduction par la méthode TRAP pour la sous population musculaire Slou positive.

## Video Article

# TRAP-rc, Translating Ribosome Affinity Purification from Rare Cell Populations of *Drosophila* Embryos

Benjamin Bertin<sup>1</sup>, Yoan Renaud<sup>1</sup>, Rajaguru Aradhya<sup>2</sup>, Krzysztof Jagla<sup>1</sup>, Guillaume Junion<sup>1</sup>

<sup>1</sup>University of Clermont-Ferrand

<sup>2</sup>Memorial Sloan Kettering Cancer Center

Correspondence to: Krzysztof Jagla at [christophe.jagla@udamail.fr](mailto:christophe.jagla@udamail.fr)

URL: <http://www.jove.com/video/52985>

DOI: [doi:10.3791/52985](https://doi.org/10.3791/52985)

Keywords: Molecular Biology, Issue 103, TRAP, translome, *Drosophila*, embryonic stage, muscle subtypes

Date Published: 9/10/2015

Citation: Bertin, B., Renaud, Y., Aradhya, R., Jagla, K., Junion, G. TRAP-rc, Translating Ribosome Affinity Purification from Rare Cell Populations of *Drosophila* Embryos. *J. Vis. Exp.* (103), e52985, doi:10.3791/52985 (2015).

## Abstract

Measuring levels of mRNAs in the process of translation in individual cells provides information on the proteins involved in cellular functions at a given point in time. The protocol dubbed Translating Ribosome Affinity Purification (TRAP) is able to capture this mRNA translation process in a cell-type-specific manner. Based on the affinity purification of polysomes carrying a tagged ribosomal subunit, TRAP can be applied to translome analyses in individual cells, making it possible to compare cell types during the course of developmental processes or to track disease development progress and the impact of potential therapies at molecular level. Here we report an optimized version of the TRAP protocol, called TRAP-rc (rare cells), dedicated to identifying engaged-in-translation RNAs from rare cell populations. TRAP-rc was validated using the Gal4/UAS targeting system in a restricted population of muscle cells in *Drosophila* embryos. This novel protocol allows the recovery of cell-type-specific RNA in sufficient quantities for global gene expression analytics such as microarrays or RNA-seq. The robustness of the protocol and the large collections of Gal4 drivers make TRAP-rc a highly versatile approach with potential applications in cell-specific genome-wide studies.

## Video Link

The video component of this article can be found at <http://www.jove.com/video/52985/>

## Introduction

Understanding how cells acquire specific properties during development is crucial in order to appreciate the complexity of organs and their potential evolution towards pathological conditions. There is a big research push to decipher the gene regulatory networks that underly cell specification and differentiation by unrevealing global gene expression profiles, Pol II binding status, transcription-factor occupancy on regulatory sequences or post-translational modifications of histones.

To assess gene expression at a global level microarrays and RNA-seq approaches are used. Microarrays allow to detect mRNA abundance, whereas RNA-seq is better geared to informing on the different types of RNAs including coding and noncoding RNAs like microRNAs, lncRNAs or snRNAs. However, these techniques cannot differentiate actively-transcribed, steady-state mRNA, actively-translating mRNA, and mRNA entering the degradation process. Measuring steady-state mRNA levels is known to be a poor estimation of proteome composition<sup>1-3</sup>. In contrary, identifying actively-translating mRNA gives a more accurate picture of protein production.

However, transcriptomic studies have mainly been led at whole-organism level by comparing gain or loss of function of a specific factor to wild-type condition<sup>4-7</sup> or on dissected tissue, making gene expression profiles difficult to interpret. Recent advances in cell targeting tools, affinity purification and sensitivity improvements now allow to perform genome-wide analyses on small cell populations or even single cells in a living organism.

In *Drosophila*, large collections of transgenic lines enable specific temporal and spatial targeting of different tissues and cell subpopulations via the GAL4/UAS system<sup>8-10</sup> yielding more accurate quantification of the molecular mechanisms required for cell function.

Among newly developed approaches polysome profiling by fractionation was described as a way to capture actively-translating RNA<sup>11</sup>. This method based on sucrose gradient centrifugation allows the selection of polysome fraction and determination of mRNA translated genome-wide when analyzed with microarrays or deep sequencing. Polysome isolation can be coupled with nuclease digestion to identify ribosomal footprints corresponding to RNA fragments protected by the ribosomes during the translation process<sup>12</sup>. Ribosomal footprinting increases the accuracy of quantification of RNA translation and RNA sequence-level resolution and ribosome positioning, makes it possible to measure rate of translation. However, to date it has not been applied to cell-specific isolation of translomes, mainly due to the strong decrease in RNA yield obtained at the end of the ribosome footprinting process. Given that the size selection of RNA may generate potential false-positives from different RNPs that

could also be protecting RNA in a similar manner, further developments are needed to improve ribosomal footprinting and adapt it to cell-specific approaches.

One of such methods, called the Translating Ribosome Affinity Purification (TRAP) has been described in 2008 by Heiman *et al.* and applied to isolate the translome from a subset of neurons in mice. By epitope-tagging a ribosomal protein in a cell-type-specific manner, polysomes can be selectively purified without the laborious step of cell dissociation and sorting. In this case, the 60S ribosomal protein L10a at the surface of the ribosome was tagged with eGFP<sup>13</sup>. Since 2008, several studies have used this method in different species, from mice<sup>14-19</sup> to *X. laevis*<sup>20,21</sup>, zebrafish<sup>22,23</sup> and *Arabidopsis thaliana*<sup>24</sup>. The TRAP method has also been adapted to the *Drosophila* model and used to purify cell-type-specific mRNAs from neuronal cells<sup>25</sup> and astrocytes<sup>26</sup>. The versatile binary GAL4/UAS system was used to express GFP-tagged RpL10A in a tissue-specific manner. Heads of adult flies were dissected to perform polysome selection from neuronal cells (targeted by pan-neural Elav-Gal4 driver) and isolated RNAs were sequenced. The large number of up-regulated genes corresponded to those known to be expressed in the nervous system, indicating that the approach has good specificity and prompting us to adapt it to rare cell populations (less than 1% of cells) present in developing *Drosophila* embryos.

Within the *Drosophila* model, the embryonic stage is a model of choice for the study of developmental processes, as the molecular actors and morphogenetic movements are evolutionarily conserved. The only tissue-specific transcriptomic approaches conducted so far in this model have been performed using cell or nuclear sorting and have only been able to study steady-state transcriptome<sup>27-31</sup>, creating a need for methods dedicated to tissue/cell-specific translome profiling in the fly embryo. Here we report the first TRAP protocol dedicated to *Drosophila* embryos. With this method, engaged-in-translation mRNA in a very restricted cell population of around 100 muscle cells per embryo was successfully isolated with high quality and specificity. The binary GAL4/UAS system is used to drive the expression of GFP-tagged RpL10A in subpopulation of muscles without any toxicity, no apparent phenotypes or developmental delay as shown previously<sup>25</sup>. *Drosophila* embryos possess 30 muscles per hemisegment that will give rise to the musculature of the larva. By using a regulatory region discovered in the vicinity of the slouch identity gene, 6 out of the 30 multi-nucleated muscles are targeted. In this assay, ribosomes are immobilized on the mRNA using the translation elongation inhibitor cycloheximide. Cytosolic extracts are then used for affinity purification with GFP antibody-coated magnetic beads. Quality of purified RNA was validated using bioanalyzer. Quantitative reverse transcription PCR (RT-qPCR) was used to determine the specificity and sensitivity of mRNA isolation and demonstrated that our optimized protocol is highly efficient.

## Protocol

### 1. Fly Line Generation

1. Generate two different constructs. In the first construct, RpL10A cDNA (ribosomal protein subunit L10a) is cloned downstream of GFP into pUASP-PL-GFP-Nter plasmid (modified version from<sup>32</sup>). The use of a germline promoter allows a more polyvalent usage of the transgenic line.  
NOTE: Obtain the second construct by cloning the regulatory sequence driving tissue-specific expression of the slouch gene upstream of GAL4 (TF) using pPTGAL4 plasmid.
2. Inject plasmids separately into *w*<sup>1118</sup> flies and insert constructs into the fly genome by P-element insertion (according to standard procedure)<sup>33</sup>.
3. Select transformants in order to recover fly lines with constructs on two different chromosomes. The final TRAP line is created by combining these two lines (genetic crosses) in order to get one double-transgenic stable line. F0: *w*;<sup>-</sup>; Cyo, UAS EGFP::RpL10A; MKRS/TM6B *x w*;<sup>-</sup>;Cyo/Scu;Slouch Gal4/TM6B. F1: *w*;<sup>-</sup>; Cyo, UAS EGFP::RpL10A; Slouch Gal4/TM6B. Massively amplify this line and collect the desired-staged embryos.

### 2. Embryo Collection

1. Prepare 8 large cylindrical population cages containing around 40 g of young flies per cage (around 180,000 flies/cage).
2. Proceed with so called pre-lays that allow flies to clear developing embryos from their oviducts. For that, lay the flies for 1 hr on two 11 cm-diameter petri dishes containing a mixture of solidified agar and grape juice and cover 1/3 of the surface with freshly-made yeast paste. After 3 exchanges of grape juice plates collect the real embryos on similar freshly made plates.  
NOTE: Incubation time will vary according to developmental time-window studied.
3. Remove plates from cages and leave them at the same temperature for the time needed to obtain correct developmental stage (e.g., 3 hr of egg-laying + 10 hr of additional incubation gives embryos from stage 10-13 hr After Egg Laying (AEL)). Resuspend the plate content (embryos, yeast paste, dead flies) with 50 ml of water using a brush.
4. Go through a series of sieves (700 μm, 355 μm, 112 μm) to separate embryos from dead flies and remaining body parts and collect embryos retained on the smaller-diameter sieve, then wash briefly in deionized water. Do not use more than 18 plates per collection as it can clog the sieves.
5. Dechorionate embryos in 4.5% bleach in deionized water for 2 min and rinse thoroughly with deionized water for 30-60 sec. Incubate embryos in PBS 0.01% Tween 20, 100 μg/ml cycloheximide, for 5-10 min at RT under agitation.
6. Dry embryos on cellulose absorbent sheets and transfer them to microcentrifuge tubes. Weigh the tubes and flash-freeze the embryos by immersion in liquid nitrogen. At this stage, the dried embryos can be stored for several months at -80 °C.

### 3. Preparation of Magnetic Beads Coupled to GFP Antibody

1. Thoroughly resuspend the Protein G magnetic beads in the original bottle by gentle agitation.
2. Transfer 90 μl of beads to an RNase-free tube and collect them on a magnet (30-60 sec). 90 μl of beads is necessary to perform immunoprecipitation (IP) with 1 ml of lysate containing 60-80 mg/mL proteins. Respect ratios to prevent bead saturation. Use several tubes according to protein amount.



3. Wash beads with 1× PBS 0.01% Tween 20 (500 µl) and collect beads on the magnet to discard supernatant. Add 30 µg of GFP antibody in 350 µl of PBS 0.01% Tween 20 to the beads and incubate with slow end over end mixing for 30 min at RT.
4. Collect beads on the magnet (30-60 sec), discard supernatant and rinse in 500 µl of polysome extraction buffer (Hepes 10 mM, KCl 150 mM, MgCl<sub>2</sub> 5 mM, DTT 0.5 mM, Triton 1%, Cycloheximide 100 µg/ml, Protease inhibitor 1×, RNase inhibitor 100 U).
5. Collect beads on the magnet and add 500 µl of blocking buffer (BSA 0.1 µg/µl, yeast tRNA 0.1 µg/µl, glycogen 0.1 µg/µl in polysome extraction buffer).
6. Incubate for 30 min at RT and repeat this step once with fresh blocking buffer.
7. Collect beads on the magnet and rinse in 500 µl of polysome extraction buffer, then recollect beads and proceed immediately to the immunopurification step (section 6).

## 4. Lysate Preparation

1. Before starting, set up the program on the multi-directional fast speed beads grinder: 2×10 sec at 5,000 rpm, 15 sec pause. Transfer dried embryos (1.5 g) to 15 ml tubes prechilled on ice containing polysome extraction buffer, and homogenize immediately. Homogenize 1.5 g of embryos in 4 ml of polysome extraction buffer.
2. Spread homogenized embryos in two prechilled 2 mL tubes and grind the embryos by running the homogenizer program. Transfer the lysate into fresh prechilled microcentrifuge tubes on ice and prepare a postnuclear supernatant by centrifugation at 4 °C for 10 min at 2,000 g.
3. Transfer supernatant to a fresh prechilled microcentrifuge tube on ice, add 1/100 sample volume of 10% Nonidet P-40 to the supernatant (final concentration = 0.1%), and mix it gently by inverting the tube.
4. Pulse-centrifuge the sample in a minifuge to collect the liquid at the bottom of the tube, add 1/9 sample volume of 300 mM 1,2-Diheptanoyl-sn-Glycero-3-Phosphocholine (DHPC) (final concentration = 30 mM), mix it gently by inverting the tube, and incubate the mixture on ice for 5 min (mix by inverting several times during incubation).
5. Prepare the post-mitochondrial supernatant by centrifugation at 4 °C for 10 min at 20,000 g. Measure protein concentration by the Bradford method: concentration in lysate should be between 60-80 mg/ml. Take the supernatant to a fresh new prechilled microcentrifuge tube and proceed immediately to the pre-absorption step (section 5).

## 5. Pre-absorption

1. Thoroughly resuspend the magnetic beads by gentle agitation and transfer 30 µl of beads into a microcentrifuge tube at RT (30 µl of beads/1 ml embryonic lysate).
2. Collect beads on the magnet, pipette off the supernatant, and resuspend the beads in polysome extraction buffer (500 µl). Collect beads on the magnet, add embryonic lysate, and incubate for 1 hr at 4 °C on a rotator.
  1. Remove beads and keep supernatant on ice for the immunopurification step. Before immunopurification, keep 100 µl of supernatant (input) to compare global RNA sample with the RNA sample collected after immunopurification, by RT-qPCR for example.

## 6. Immunopurification

1. Add 1 ml (corresponding to 60-80 mg/ml protein) of pre-absorbed lysate to an RNase-free tube containing 90 µl of blocked beads coupled to GFP antibody. Incubate the samples at 4 °C for 2 hr with gentle end-over-end mixing in a tube rotator. Alternatively, perform the incubation O/N at 4 °C.
2. After incubation, collect beads on the magnet. Use a minifuge to spin down the beads, then resuspend them in 500 µl of wash buffer (Hepes 10mM, KCl 350mM, MgCl<sub>2</sub> 5mM, Nonidet P-40 1%, DTT 0.5 mM, cycloheximide 100 µg/ml, RNase inhibitor 100 U) and collect them on the magnet. Repeat this step two more times.
3. Change the beads to a new prechilled RNase-free tube and wash twice. Collect beads on a magnet and proceed to RNA extraction by adding 1 ml TRIzol directly to the beads and follow the manufacturer's instructions.

## 7. RNA Clean-up and Quality Assessment

1. Use an RNeasy Micro Kit as per the manufacturer's instructions. At completion, elute RNA with (x) µl RNase-free water. Warm for 2 min at 60 °C, and store snap-frozen samples at -80 °C. Check RNA quality using bioanalyzer (follow the manufacturer's instructions).
2. To test the specificity of the TRAPed RNA, perform reverse transcription on 3 ng of purified RNA (input and IP) according to the manufacturer's instructions (superscript III). Use the cDNA obtained for qPCR using gene-specific primer sets.

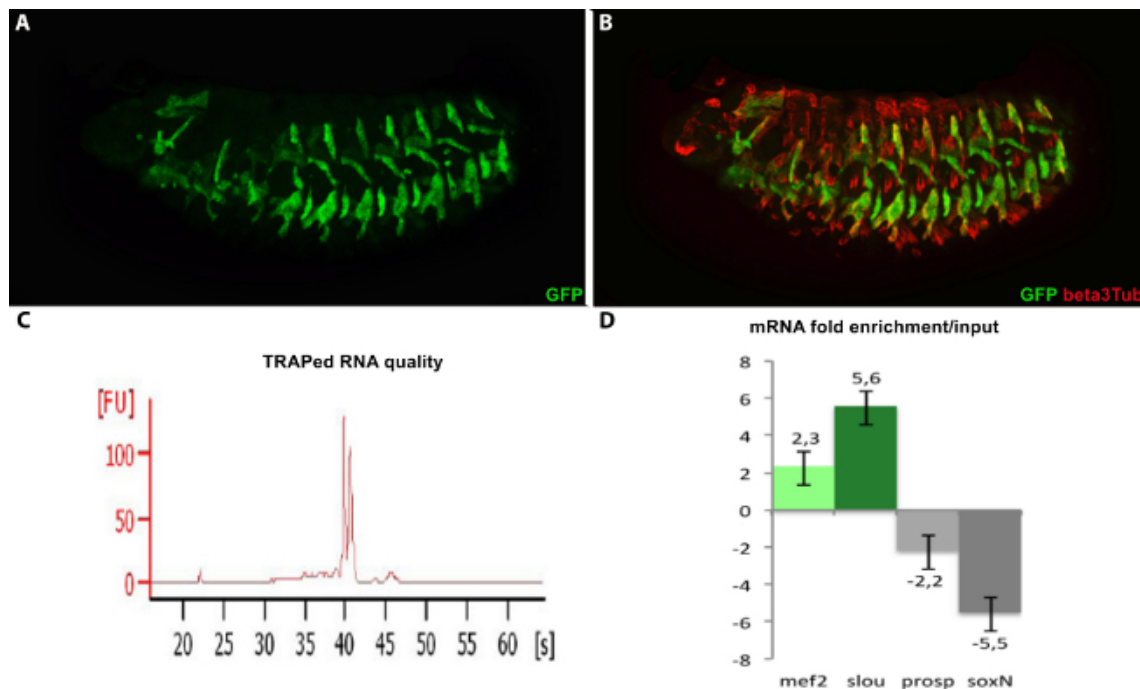
## Representative Results

The *Drosophila* embryonic somatic muscle system is composed of 30 muscles per hemisegment, and each muscle has a specific set of properties: code of identity genes, position, number of fusion events, attachment sites and innervation. Identifying translated mRNA in a specific muscle subset will give information on the proteins required to form these specific muscle characteristics. Using the TRAP-based method, RNA from a subpopulation of muscles expressing the identity gene *slouch* was purified (**Figure 1A-B**). A major concern of this approach is the quality and specificity of the isolated RNA. The optimized protocol reported here enabled systematic recovery of high-quality RNA assessed with bioanalyzer analysis. The profile obtained shows no degradation of RNA (**Figure 1C**).

In other studies developed around the TRAP method, questions were raised over the specificity of the data. Here we improve the method to suppress background linked to non-specific binding of RNA to the beads or the tubes and by the optimization of the wash buffer. In order to assess the background level and efficiency of isolating *slouch*-expressing cells, we led RT-qPCR trials on 3 replicates of the same embryonic stage. Fold-enrichments of 4 different genes (*mef2*, *slouch*, *prospero*, *soxNeuro*) were calculated compared to input and normalized against the

*RpL32* gene (Figure 1D). These results showed 2.3-fold enrichment of pan-muscular gene *mef2* and 5.6-fold enrichment of the *slouch* gene. In contrast, the two genes expressed in the neural system were depleted compared to input. Very similar fold change values were observed on 3 biological replicates, demonstrating that the protocol is robust. With *Slouch-Gal4* driver and starting with 1.5 g, around  $1.5 \times 10^6$  embryos were obtained that contain 100 GFP cells/embryo (total of  $150 \times 10^6$  positive cells).

After running the TRAP experiment, the yield is around 25–45 ng of specific RNA depending on stage-of-interest. This material is enough to run microarray analysis or RNA-seq using an amplification protocol to build an RNA-seq library. Efficiency will depend strongly on the strength of the driver used to express Rpl10A-EGFP.



**Figure 1. Quality and specificity assessment of TRAP-isolated RNA from *Slouch-GAL4*>*UASRpl10A-EGFP* embryos.**

(A-B) Confocal images of a stage-16 embryo showing Rpl10A-EGFP expression specifically in the six *Slouch* muscle cells per hemisegment (A). Co-localization of Rpl10A-EGFP with the general muscle marker Beta3tubulin is observed on merge picture (B). (C) Quality control of TRAPed RNA ran on bioanalyzer showing perfect integrity of rRNA 18S and 28S. (D) RT-qPCR analysis showing high specificity of TRAP-isolated mRNA experiments on 3 biological replicates of stage-16 embryos. Fold change is calculated compared to input and normalized against the *RpL32* gene. *Mef2* transcripts present in all muscle lineages are 2.3-fold-enriched compared to input whereas more restricted *slouch* transcripts are 5.6-fold-enriched. Neural cells expressing *prospero* and *soxN* genes are depleted 2.2-fold and 5.5-fold respectively. Error bars represent standard deviation (n=3). [Please click here to view a larger version of this figure.](#)

## Discussion

This paper describes a modified Translating Ribosome Affinity Purification protocol (dubbed 'TRAP-rc') dedicated to the study of rare cell populations in *Drosophila* embryos. Information is provided on the key steps for successful isolation of specific RNA with a yield suitable for microarray or RNA-seq analysis, i.e., 1) quantity of biological material required and optimized procedure for embryo lysis and polysome extraction; 2) beads/antibody-to-lysate ratios for optimal immunoprecipitation; 3) steps allowing reduction of background including wash buffer composition so as to run TRAP-rc experiments with high specificity and sensitivity.

This protocol yields reproducible data making it readily applied to any other cell types. This technique makes it possible to analyze differential gene expression at the level of actively-translating mRNA and thus pave the way to understanding protein expression in a specific cell type at a specific time window. Note however that protein abundance will depend on the rate of the translation and degradation processes. A major limitation of the TRAP method is its inability to measure protein content in an accurate quantitative manner or to detect post-translational modification. Improving the quantification by measuring ribosome density along the mRNA body and, in doing so, in a cell-type-specific manner can make TRAP combined with ribosome footprinting a very powerful tool. In a recent paper<sup>34</sup>, this combination was performed in human embryonic kidney 293 cells. The authors ran nuclease footprinting followed by affinity purification of an inducible biotinylated form of Rpl10A using streptavidin beads. This is a proof of principle that it is possible to perform ribosome footprinting in a whole organism while targeting a specific cell type, the limitation being the amount of biological material required for the purpose.

Performing TRAP experiments in parallel with global transcriptomic analyses will make it possible to track proportions between newly-transcribed and actively-translating RNA. This will be deeply informative of the potential post-transcriptional mechanisms taking place in specific developmental contexts or specific cell populations—by microRNAs for example.

In conclusion, TRAP is a highly efficient, specific and sensitive method for identifying RNAs bound to actively-translating ribosomes in a cell-specific manner. This method can be used in a wide variety of organisms and tissue types. The new TRAP-rc protocol described here was

optimized for rare cell populations (less than 1% of total cell number) and for a quantity of final material sufficient for subsequent analyses at whole-genome level.

A possible limitation of this method is the requirement of a strong driver to produce tagged ribosomes in sufficient amount to compete with endogenous untagged ones in the targeted cell type. In a recent study<sup>25</sup>, authors estimated to 10-30% the ratio of tagged versus untagged ribosomes.

To overcome this problem increasing UAS-RpL10A-EGFP copy number should considerably improve this balance. Alternatively, adding to the described genetic background a UAS-GAL4 transgene will amplify the production of GAL4 protein and indirectly increase expression of RpL10A-EGFP. This should favor a better occupancy of tagged ribosomes on mRNA.

Note that this already efficient approach can be made even more powerful by adapting complementary methods to bring a more quantitative aspect or to discover and unravel molecular mechanisms that are essential for gene expression control.

## Disclosures

The authors have nothing to disclose.

## Acknowledgements

The authors thank Nicolas Allegre, Maud Peyny and Jean-Philippe Da Ponte for their excellent technical assistance. This work was supported by the Agence Nationale de la Recherche (ANR-JC-CARDIAC-SPE), the INSERM starting grant, ANR ID-CELL-SPE grant, Equipe FRM and the AFM 15845 grant.

## References

1. Sousa Abreu, R., Penalva, L. O., Marcotte, E. M., Vogel, C. Global signatures of protein and mRNA expression levels. *Molecular bioSystems*. **5**, 1512-1526 (2009).
2. Maier, T., Guell, M., Serrano, L. Correlation of mRNA and protein in complex biological samples. *FEBS letters*. **583**, 3966-3973 (2009).
3. Schwanhaussner, B., et al. Global quantification of mammalian gene expression control. *Nature*. **473**, 337-342 (1038).
4. Junion, G., et al. Genome-wide view of cell fate specification: ladybird acts at multiple levels during diversification of muscle and heart precursors. *Genes*. **21**, 3163-3180 (2007).
5. Jakobsen, J. S., et al. Temporal ChIP-on-chip reveals Biniou as a universal regulator of the visceral muscle transcriptional network. *Genes*. **21**, 2448-2460 (2007).
6. Cunha, P. M., et al. Combinatorial binding leads to diverse regulatory responses: Lmd is a tissue-specific modulator of Mef2 activity. *PLoS genetics*. **6**, e1001014 (2010).
7. Sandmann, T., et al. A core transcriptional network for early mesoderm development in *Drosophila melanogaster*. *Genes, & development*. **21**, 436-449 (2007).
8. Kvon, E. Z., et al. Genome-scale functional characterization of *Drosophila* developmental enhancers in vivo. *Nature*. **512**, 91-95 (2014).
9. Pfeiffer, B. D., et al. Tools for neuroanatomy and neurogenetics in *Drosophila*. *Proceedings of the National Academy of Sciences of the United States of America*. **105**, 9715-9720 (2008).
10. Gallo, S. M., et al. REDfly v3.0: toward a comprehensive database of transcriptional regulatory elements in *Drosophila*. *Nucleic acids research*. **39**, D118-D123 (2011).
11. Gandin, V., et al. Polysome fractionation and analysis of mammalian translatoemes on a genome-wide scale. *Journal of visualized experiments : JoVE*. (2014).
12. Ingolia, N. T., Brar, G. A., Rouskin, S., McGeachy, A. M., Weissman, J. S. The ribosome profiling strategy for monitoring translation in vivo by deep sequencing of ribosome-protected mRNA fragments. *Nature protocols*. **7**, 1534-1550 (2012).
13. Heiman, M., et al. A translational profiling approach for the molecular characterization of CNS cell types. *Cell*. **135**, 738-748 (2008).
14. Doyle, J. P., et al. Application of a translational profiling approach for the comparative analysis of CNS cell types. *Cell*. **135**, 749-762 (2008).
15. Dougherty, J. D., Schmidt, E. F., Nakajima, M., Heintz, N. Analytical approaches to RNA profiling data for the identification of genes enriched in specific cells. *Nucleic acids research*. **38**, 4218-4230 (1093).
16. Fomchenko, E. I., et al. Recruited cells can become transformed and overtake PDGF-induced murine gliomas in vivo during tumor progression. *PLoS one*. **6**, e20605 (2011).
17. Helmy, K., et al. Identification of global alteration of translational regulation glioma in vivo. *PLoS one*. **7**, e46965 (2012).
18. Zhou, P., et al. Interrogating translational efficiency and lineage-specific transcriptomes using ribosome affinity purification. *Proceedings of the National Academy of Sciences of the United States of America*. **110**, 15395-15400 (2013).
19. Drane, L., Ainsley, J. A., Mayford, M. R., Reijmers, L. G. A transgenic mouse line for collecting ribosome-bound mRNA using the tetracycline transactivator system. *Frontiers in molecular neuroscience*. **7**, 82 (2014).
20. Yoon, B. C., et al. Local translation of extranuclear lamin B promotes axon maintenance. *Cell*. **148**, 752-764 (2012).
21. Watson, F. L., et al. Cell type-specific translational profiling in the *Xenopus laevis* retina. *Developmental dynamics : an official publication of the American Association of Anatomists*. **241**, 1960-1972 (2012).
22. Fang, Y., et al. Translational profiling of cardiomyocytes identifies an early Jak1/Stat3 injury response required for zebrafish heart regeneration. *Proceedings of the National Academy of Sciences of the United States of America*. **110**, 13416-13421 (2013).
23. Tryon, R. C., Pisat, N., Johnson, S. L., Dougherty, J. D. Development of translating ribosome affinity purification for zebrafish. *Genesis*. **51**, 187-192 (2013).
24. Jiao, Y., Meyerowitz, E. M. Cell-type specific analysis of translating RNAs in developing flowers reveals new levels of control. *Molecular systems biology*. **6**, 419 (2010).

25. Thomas, A., *et al.* A versatile method for cell-specific profiling of translated mRNAs in *Drosophila*. *PLoS one*. **7**, e40276 (2012).
26. Huang, Y., Ng, F. S., Jackson, F. R. Comparison of Larval and Adult *Drosophila* Astrocytes Reveals Stage-Specific Gene Expression Profiles. *G3*. (2015).
27. Steiner, F. A., Talbert, P. B., Kasinathan, S., Deal, R. B., Henikoff, S. Cell-type-specific nuclei purification from whole animals for genome-wide expression and chromatin profiling. *Genome Res.* **22**, 766-777 (2012).
28. Henry, G. L., Davis, F. P., Picard, S., Eddy, S. R. Cell type-specific genomics of *Drosophila* neurons. *Nucleic acids research*. **40**, 9691-9704 (2012).
29. Dobi, K. C., Halfon, M. S., Baylies, M. K. Whole-genome analysis of muscle founder cells implicates the chromatin regulator Sin3A in muscle identity. *Cell reports*. **8**, 858-870 (2014).
30. Salmand, P. A., Iche-Torres, M., Perrin, L. Tissue-specific cell sorting from *Drosophila* embryos: application to gene expression analysis. *Fly*. **5**, 261-265 (2011).
31. Busser, B. W., *et al.* Molecular mechanism underlying the regulatory specificity of a *Drosophila* homeodomain protein that specifies myoblast identity. *Development*. **139**, 1164-1174 (2012).
32. Rorth, P. Gal4 in the *Drosophila* female germline. *Mechanisms of development*. **78**, 113-118 (1998).
33. Rubin, G. M., Spradling, A. C. Genetic transformation of *Drosophila* with transposable element vectors. *Science*. **218**, 348-353 (1982).
34. Ingolia, N. T., *et al.* Ribosome profiling reveals pervasive translation outside of annotated protein-coding genes. *Cell reports*. **8**, 1365-1379 (2014).



## B – Identification de nouveaux gènes impliqués dans l'identité musculaire par la méthode TRAP

### I. Publication 2 « Translational profiling of muscle subsets in *Drosophila* identifies muscle type specific growth tuners controlling multiple muscle properties »

Dans ce manuscrit (en cours de préparation) une analyse translatomique, réalisée de manière cellule spécifique grâce à la méthode TRAP, a été utilisée pour identifier des signatures d'expression génique qui déterminent la spécification et l'acquisition de l'identité musculaire. Pour cela les ARN messagers engagés en traduction ont été isolés à partir de deux populations restreintes de muscles (Lms- et Slou-positifs) et à trois fenêtres de temps au cours du développement musculaire. Cette approche globale muscle spécifique (analyse spatiale) nous a permis d'identifier de nouveaux facteurs, conservés chez l'homme, comme étant impliqués potentiellement dans le code d'identité des muscles Slou-positifs. Des analyses spatio-temporelles de l'expression ont également mené à l'identification de nouveaux gènes réalisateurs du programme d'identité des muscles Lms-positifs qui codent pour des protéines de liaison au cytosquelette d'actine.

Une première étude spatiale des données a révélé des gènes codant pour des protéines de régulation de l'expression des gènes tels que des facteurs de transcription, des protéines de liaison à l'ARN ou encore des modificateurs de la chromatine comme différemment exprimés entre ces deux populations. L'analyse du profil d'expression des gènes retrouvés enrichis dans la population Slou a validé la présence des transcrits *dSix4*, *EndoGl*, *CG10565*, *CG10654* et *Ptx1* au sein de ces muscles. Les fonctions de ces facteurs restent à déterminer, néanmoins leur expression muscle-spécifique laisse envisager l'implication de ces derniers dans le code combinatoire des iTFs qui contrôlent la diversification de la population Slou.

Par la suite l'analyse spatio-temporelle de nos données a permis de déterminer des signatures de traduction spécifiques des populations Lms et Slou dans les trois fenêtres de temps ciblées au cours de la myogenèse. Il s'avère qu'un ensemble de gènes codant pour des protéines de régulation de l'actine présentent un enrichissement croissant à travers les stades de développement dans la population Lms alors que ces derniers sont très peu



représentés dans la population Slou. L'analyse de leur profil d'expression a montré que parmi les gènes candidats sélectionnés deux en particulier, *gel* et *dCryab* présentent une expression plus spécifique des muscles Lms-positifs (LTs). *Gel* et *dCryAB* s'expriment dans les muscles LTs à partir du début du stade 15 (fenêtre temporaire intermédiaire T2) et sont maintenus dans les stades plus tardifs (fenêtre T3). Nous avons ensuite analysé les effets de la perte de fonction de *gel* et *dCryAB* à l'aide de mutants conditionnels générés par la méthode de CRISPR/Cas9. Dans les contextes mutants pour ces deux gènes, des phénotypes similaires ont été observés. Ces derniers correspondent à une augmentation du nombre de noyaux, souvent lié à la séparation des extrémités des muscles LTs et à la spécification d'une cellule tendon supplémentaire. Ces deux gènes semblent donc agir de manière coordonnée dans cette population restreinte de muscles pour contrôler de manière fine les étapes de fusion et d'élongation vers les cellules de tendon. De ce fait, *gel* et *dCryAB* jouent un rôle multiple dans l'acquisition de l'identité des muscles Lms-positifs. Les deux protéines ayant des fonctions dans la dépolymérisation ou dans la limitation de polymérisation (capping) de l'actine, nous proposons que *Gel* et *dCryAB* fonctionnent comme réalisateurs de l'identité musculaire en régulant le cytosquelette d'actine.





## **Translational profiling of muscle subsets in *Drosophila* identifies muscle type specific growth tuners controlling multiple muscle properties**

Benjamin Bertin, Yoan Renaud#, Teresa Jagla, Guillaume Lavergne, Cristiana Dondi, Jean-Philippe Da Ponte, Guillaume Junion\* and Krzysztof Jagla\*

GReD - INSERM U1103, CNRS UMR6293, Clermont Université,  
28, Place Henri Dunant, 63000 Clermont-Ferrand, France

# Actually: [www.byonet.fr](http://www.byonet.fr)

\* Corresponding authors  
[christophe.jagla@uca.fr](mailto:christophe.jagla@uca.fr)  
[guillaume.junion@uca.fr](mailto:guillaume.junion@uca.fr)

Keywords: *Drosophila*, muscle diversification, TRAP, identity code, *Gelsoline*, *dCryAB*

## Abstract

Diversification of cell types is an essential developmental process requiring timely and spatially coordinated regulatory mechanisms at gene expression levels. *Drosophila* embryonic muscle system emerged as an accurate model to identify mechanisms governing diversification of cell subsets. Here, we applied TRAP approach to identify translational gene expression profiles underlying individual characteristics of two subpopulations of embryonic muscles that express identity genes *slouch* and *lms*. Comparative analyses of genes differentially expressed in these two muscle subsets revealed new components of identity code including the evolutionarily conserved homeobox genes, *Ptx1* and *dSix4*, with restricted expression patterns overlapping Slouch-positive muscle progenitors. To understand how the acquisition of individual muscle properties is acquired we generated temporal transition profiles of TRAP-enriched genes. This has led to the identification of a cluster of genes encoding regulators of actin cytoskeleton enriched in *lms*-positive lateral transverse (LT) muscles. Among them, *Gelsolin* and *dCryAB* code for conserved actin-interacting proteins and display LT muscle specific expression in mid and late stage embryos. Conditional muscle-specific CRISPR-Cas9 knockout of either *Gelsolin* or *dCryAB* leads to aberrant LT muscle characteristics including the fusion program, polarized myotube extension and attachment to tendon cells revealing that single realisor gene could control multiple muscle properties. Altogether, we provide the first dynamic translational profiling of muscle subsets and identify conserved components of muscle identity code including a new class of realisor genes required for fine-tuning of growth related LT muscle properties.

## Introduction

Diversification of cell types is a fundamental process that occurs during development of multicellular organisms and is essential to build functional organs. Skeletal muscle, the major component of animal body, diversifies into muscle types showing different individual properties that support distinct functions such as locomotion or breathing. However, understanding genetic and molecular mechanisms conferring these individual characteristics to each muscle during development remains a major challenge. Muscle network in *Drosophila* embryos represents an attractive system for studying cell diversification. It is composed of 30 muscle fibers per abdominal hemisegment, which are easy to detect and to follow during development. Despite common characteristics such as formation by myoblast fusion and capacity to contract each muscle displays a specific size, orientation, number of nuclei, attachment and innervation (Baylies et al., 1998). All mesodermal cells that are at the origin of somatic muscles are specified by the myogenic regulatory gene *Twist*, and then subdivided into dorsal, lateral and ventral domains by the activity of homeodomain transcription factors *Tnman*, *dSix4* and *Pox-meso* (Bodmer, 1993; Duan et al., 2007; Ip et al., 1992; Riechmann et al., 1997; Clark et al., 2006). Within the somatic mesoderm three different types of cells are specified, the founder cells (FC), the fusion competent myoblast (FCMs) and the adult muscle precursors (AMP) that keep their non-differentiated fate until their reactivation during the larval stage (Bate, 1990; Bate et al., 1991, Aradhya et al., 2015). Fusion of a defined number of FCMs with a FC leads to the formation of a syncytial myotube (Beckett and Baylies, 2007), which elongates and selects its correct attachment sites. Muscle guidance toward tendon cells is regulated by generic signalling molecules such as *Silt* and its receptor *Robo* (Ypsilanti et al., 2010) but also fine-tuned at individual muscle level by

Derailed in lateral transverse muscles (Callahan et al., 1996) and Kon tiki, Grip and Echinoid in ventral longitudinal muscles (Estrada et al., 2007; Schnorrer et al., 2007; Swan et al., 2006). By the end of embryogenesis each muscle fibre gets innervated by motoneurons and acquire its contractile abilities (Landgraf et al., 1997). Motoneurons select their target muscle through an active recognition process involving cell adhesion molecules (CAMs) such as Capricious, Tartan, Hattifatener, Connectin, Toll, Fasciclin II-III (Kurusu et al., 2008; Nose et al., 1997; Rose et al., 1997; Davis et al., 1997; Kose et al., 1997) and secreted proteins like Semaphorin2a, NetrinB and Wnt4 (Inaki et al., 2007; Winberg et al., 1998). They are all expressed in different subsets of muscle cells and required for the formation of neuromuscular specificity (Nose, 2012). However, how these biological processes are regulated at the muscle specific level remains elusive even if it is largely accepted (Rushton et al., 1995; Landgraf et al., 1999) that the FCs harbour all the information required for the individual muscle identity. This information is thought to be provided by the combinatorial expression of genes encoding transcription factors called the identity genes or iTFs (Carmena et al., 1998; Tixier et al., 2010). To date, twenty iTFs that control muscle identity have been discovered in *Drosophila*. They are all expressed in subsets of muscle progenitors or founders and share one important feature, their loss of function leads to the loss or the aberrant formation of a muscle in which they are expressed (Baylies et al., 1998). Some identity genes such as *ladybird (lb)* (Jagla et al., 1998) or *lateral muscle scarcer (lms)* (Müller et al., 2010) show a restricted expression in lateral muscles only whereas others like *slouch (slou)* or *Krüppel (Kr)* display a broader expression pattern in ventral, lateral and dorsal domains (D'Alessio and Frasch, 1996; Knirr et al., 1999). Among recently identified iTFs Optomotor-blind-related-gene (*Org-1*) is required for the specification of SBM, LO1 and VT1 muscles (Schaub et al., 2012), Araucan/Caupolican for lateral muscles (Carrasco-Rando et al., 2011) and Tailup for four dorsal muscles (Boukhatmi et al., 2012). Interestingly, all these factors play conserved roles during cranial musculature specification in vertebrates.

Thus, knowledge gained on iTF gene function in *Drosophila* could help in the understanding of how diversification processes are regulated in general. One important aspect that remain to be investigated is the gene network that operates downstream of iTFs to set individual muscle properties. Whole genome approaches based on ChIP experiments have shown that iTFs regulate directly not only other identity genes but also downstream regulators of muscle identity that control cell size, shape, attachment or innervation (Busser et al., 2012; Junion et al., 2007; Schaub et al., 2012). Indeed *Eve*, *Lb* and *Slou* control the size of individual muscles by regulating the number of fusion events. They achieve this by setting expression levels of three genes (*mp20*, *Pax* and *mspo*) that regulate actin dynamics and cell adhesion in muscle specific manner (Bataillé et al., 2010, De Jossineau et al., 2012). However our knowledge on genetic control of cell diversification during development is essentially based on functional analyses of individual genes and on genomic analyses performed on whole embryos at a given developmental point. This type of strategy carries important limitations and is not adapted to identify gene networks governing the acquisition of individual cell properties. Moreover the genome wide approaches performed on whole embryos without selecting particular cells are not adapted to identify genes expressed in rare cell populations such as subsets of muscles.

To overcome this technological issue we optimised the Translational Ribosome Affinity Purification (TRAP) approach to small subsets of FCs and developing muscle precursors. This method allows the identification of messenger RNA being translated in cell specific manner (Heiman et al., 2008). We applied TRAP method to two restricted muscle

populations corresponding to Slou- and Lms-positive cells and to the entire muscle cell population expressing Dumbfounded (*Duf*). To identify genes involved in the acquisition of individual muscle properties we performed TRAP in three developmental time windows targeting myoblast specification (T1 - 7-10 AEL), myoblast fusion, myotube elongation and recognition of tendon cell (T2 - 10-13 AEL) and innervation by motoneurons and stabilisation of attachment (T3 - 13-16h AEL). Transcriptional profiling of TRAP-purified mRNA followed by bioinformatics analyses has led to identification of new candidate genes with potential role in setting the identity of Lms- and Slou-positive muscles. Temporal transition profiles have also been generated for each muscle subset revealing a cluster of genes encoding actin binding/interacting proteins enriched in the Lms-muscle subset. This suggested potential role of identified genes in one of actin dynamics involving processes such as myotube elongation, myoblast fusion, attachment and/or innervation. Indeed, two genes from the identified cluster, *Gelsolin (Gel)* and *dCryAB* display LT muscle specific expression starting from T2 and when knocked down via conditional CRISPRs led to affected LT muscle growth with increased number of myonuclei, myotube splitting and fork-like muscle extremities with ectopic attachments. This suggests that *Gel* and *dCryAB* by tuning LT muscle growth are involved in the acquisition of multiple muscle properties. We hypothesise that these unsuspected multi-task roles are related to actin severing and/or capping functions of *Gel* and *dCryAB* which could be important for stopping fusion process and limiting aberrant extensions of myotube extremities.

Thus, our genome wide transcriptional profiling of muscle subsets has led to the identification of gene expression signatures underlying muscle diversity and new class of realisor genes ensuring fine tuning of LT muscle growth.

## Results and Discussion

### Genome wide translational profiling of rare populations of embryonic muscle cells.

To better understand the molecular mechanisms conferring individual muscle characteristics, we performed translational profiling of subsets of muscle cells using TRAP approach. TRAP method was first developed to isolate polysome-associated mRNA from a subset of neurons in mice (Heiman et al., 2008) and then adapted to other model organisms including *Xenopus* (Watson et al., 2012), zebrafish (Tryon et al., 2013) and *Drosophila* (Thomas et al., 2012; Bertin et al., 2015). It is based on affinity purification of the GFP-tagged ribosomes with their associated mRNAs. Tagging of ribosomes in the cell-type of interest occurs *in vivo* by incorporation of the ribosomal Rpl10A subunit fused to GFP. Here, we took advantage of the versatile GAL4/UAS system to drive expression of Rpl10A-GFP in two sub-populations of muscle cells, using Slou-GAL4 and Lms-GAL4 drivers. We also performed control TRAP experiment by targeting all embryonic muscles with *Duf*-GAL4. For each muscle subset, translational profiling has been performed on embryos collected from three developmental time windows (T1 : 7-10 h AEL, T2 : 10-13 h AEL and T3 : 13-16 h AEL), that cover together all steps of the acquisition of muscle identity (Fig. 1, arrow). We found that both the quality and quantity of purified mRNA were adapted for performing genome-wide analyses. Also as revealed by RT-qPCR (Fig. 1C and data not shown), TRAP-ed transcripts from the Slou-positive muscles were enriched for *Slou* and *Mef2* RNAs (9,3 and 2,2 times, respectively) and depleted for non-muscle *Prospero* and *SoxNeuro* transcripts thus showing the efficacy of muscle targeting. Quality-checked mRNAs purified from the Slou-, Lms- and *Duf*-positive muscle cells at three time windows (T1, T2, T3) were used for microarray

analysis (Materials and Methods). Three microarray replicates for each condition have been analysed revealing reproducibility and robustness of our approach (Fig. 1D and data not shown). We then used bioinformatics tools to identify the enriched and depleted transcripts for each muscle population and each time point. Gene expression profiling with total embryonic RNA (input fractions) has been used as a reference for identifying differential gene expression. A significant portion of upregulated genes turns out to be common between the restricted muscle populations (Slou- and Lms-positive) and the global (Duf-positive) musculature (Fig. 1E). The percentage of Slou-specific transcripts remains constant at different time points while the proportion of Lms-specific transcripts increases (Fig. 1E). This difference could be due to the fact that Slou-expressing muscles are more heterogeneous than Lms-positive ones (LT muscles only). To validate our muscle type specific expression lists, we tested whether genes with previously characterized expression patterns in muscle subsets display expected up or down regulation in our datasets (Fig. S1 A,C). Indeed as shown in Fig. S1C, genes known to be expressed in Slou- or Lms-positive muscles are enriched in the corresponding TRAP-ed samples. We also tested whether TRAP method allows detection of “low expression” genes. To do so we crossed our lists of enriched and depleted transcripts with modEncode datasets (The modENCODE Consortium et al., 2010) and found that more than 30% of upregulated TRAP-ed genes corresponds to “low expression” modEncode category whereas the majority of depleted transcripts fits into “high expression” modEncode category (Fig. S1B). Moreover, to assess specificity of TRAP-based muscle targeting we analysed GO terms associated with the lists of enriched and depleted genes. As expected the GO terms for the enriched genes are muscle-related including muscle cell differentiation and myofibril assembly (Fig. 1F). In contrast, GO categories associated with the global list of downregulated genes are not related to muscle processes but rather to general development terms.

Thus, TRAP-based translational profiling of muscle subsets turns out to be well suited for identifying new genes involved in muscle identity.

### **New potential muscle identity factors showing differential expression in Slou- and Lms-positive muscle subsets**

Muscle diversification is tightly regulated by the combinatorial expression of transcription factors known as muscle identity code (Baylies et al., 1998; Tixier et al., 2010; De Jossineau et al., 2012). To identify new components of muscle identity code we first defined lists of genes enriched in Slou- and Lms- muscle populations with respect to Duf-positive global musculature. For each muscle subset three time points have been combined to obtain unique spatial list of enriched genes for Slou and Lms muscle population. Gene clustering according to the level of enrichment was then performed (Fig. 2A) revealing two distinct gene clusters, one with genes that are exclusively upregulated in Slou or in Lms population (orange and blue colours in the heatmap) and a second with genes displaying differential expressions in these two populations of muscles. Within the clusters of exclusively upregulated genes we found 170 candidates enriched in Slou and 99 enriched in Lms muscles only. These include genes that display Slou- or Lms-specific enrichment at least at one of the three time windows analysed. The clusters with candidates displaying differential expression comprise 156 genes with higher expression in Slou population and 51 with differential enrichment in Lms muscle cells. Given the heterogeneity of Slou population that includes dorsal, lateral and ventral muscles we focused on genes enriched in this muscle subset and belonging to the spatial cluster of type one.

Genes exclusively enriched in Slou-positive muscles include mainly the evolutionarily conserved DNA or RNA binding proteins (Fig. 2B) and among them Slou itself but also *Org-1* whose expression and key role in specification of Slou-positive muscle precursors have been extensively analysed (Schaub et al., 2012). Moreover, we identified *dSix4* (Clark et al., 2006) and *Ptx1* (Vorbrüggen et al., 1997) known to be expressed in ventral somatic mesoderm (*dSix4*) or in restricted population of ventral muscles (*Ptx1*) but with muscle identity functions that remain so far not fully understood. Importantly, our data reveal that *dSix4* in addition to the previously described early mesodermal pattern (Clark et al., 2006) also shows specific expression in VA1 and VT1 muscles in later embryonic stages (Fig. 2F,F'). This suggests that *dSix4* could contribute to the identity code of these two Slou-positive muscles. We also found that *EndoG1*, *CG10565* and *CG10654* expressions overlap that of *Slou* (Fig. 2) and could thus contribute to setting Slou-muscle identity. Interestingly, all identified Slou-enriched genes have human orthologues (Fig. 2B) and three of them, *Org1/Tbx1*, *dSix4/Six4* and *Ptx1/Pitx2* are known to play key functions during vertebrate myogenesis suggesting they are part of universal muscle identity code. Detailed developmental expression and functional analyses of identified candidates are necessary to determine their contribution to the specification of Slou-expressing muscles. As stated before, the heterogeneity of Slou-positive muscles could explain complexity of their identity code. Several genes encoding DNA or RNA binding proteins are also enriched in Lms-positive muscles (Fig. S2). They represent a pool of candidates with potential role in diversification of lateral transverse muscles.

### **Temporal transition profiles lead to identification of clusters of genes with potential roles in the acquisition of individual muscle properties.**

Embryonic myogenesis consists of a succession of developmental steps such as specification of founder cells, myoblast fusion, attachment to tendon cells and innervation by motoneurons, which are controlled at muscle specific level. General myogenic factors that control these different steps are well known, however only few genes have been discovered to be required for these processes in particular set of muscles. To identify new realisor genes that control the acquisition of individual properties of Lms- and Slou-positive muscles we generated temporal transition profiles of differentially expressed genes for both muscle subsets. Two transition profiles have been created by comparing expression levels of genes from each muscle population at temporal windows T2 versus T1 and T3 versus T2. This has led to the generation of a temporal transition heatmap (Fig. S3) with clusters of genes showing similar or distinct expression profiles overtime in Slou and in Lms muscle cells (Fig. 3A and Fig. S3).

We observed several gene expression behaviours including the “up-up”, “down-down”, “up-down” and “down-up” transition profiles, which could be common for both muscle subsets or specific for one of them (Fig. 3A). These profiles could underlie acquisition of particular muscle properties. For example we expect that genes belonging to the “down-up” cluster (mainly activated at temporal window T3) (Fig. S3) could contribute to the late events of muscle development such as innervation. Here we focused on gene clusters showing two contrasting transition behaviours, the “down-down” (cluster 1) and “up-up” (cluster 2) clusters. To the “up-up” cluster belong genes whose expression progressively increases from early to late developmental time points. These genes could potentially contribute to all major identity acquisition steps including muscle-specific fusion, growth, attachment and innervation. In contrast, the “down-down” cluster is expected to gather genes whose

expression needs to be turned down while muscle differentiation progresses. Cluster 2 is more heterogeneous than cluster 1 and comprises genes fitting into “up-stable” and “stable-up” sub-clusters and with behaviour that could be different for Slou- versus Lms- muscle population. To characterise genes from these two clusters we first performed comparative GO analyses using Cluster Profiler (see also Cluster Profiler GO analysis of all transition profiles – Fig. S3B). Cluster 2 genes show a significant enrichment of several GO terms related to muscle development whereas cluster 1 is characterized by enrichment of GOs associated with cell cycle and DNA replication, which occur during early steps of myogenesis but are not specific to muscle development (Fig. 3B). We then generated graphs of transition profiles for each muscle subset for selected representative genes from both clusters (Fig. 3C). Among the genes progressively down regulated (cluster 1) in both muscle subsets is *twist* encoding the myogenic bHLH factor whose expression is known to decrease during myogenic differentiation. Cluster 1 also comprises genes involved in Notch signalling pathway including *bib*, *phyl*, *E(spl)m4-BFM* and *E(spl)mgamma-HLH* whose expression needs to be down regulated to promote cell differentiation. Thus transition profiles of cluster 1 genes confirm that analysing temporal transition profiles of TRAP-ed genes could help in assessing their myogenic functions. As stated before, the cluster 2 genes are more heterogeneous and fit into sub-clusters with profiles that differ between muscle subsets. Among them, a group of genes encoding actin-binding proteins displays “up-up” profile in Lms-muscle subset and behaves “up-stable” in Slou-muscle population (Fig. 3C). We thought, that this gene cluster is of particular interest as it gathers genes with similar biological functions and with convergent temporal expression profiles that differ between the Lms- and Slou-muscle subsets. The fact that all of them are evolutionarily conserved prompted us to understand their muscle-type specific functions.

### ***Gelsolin* and *dCryAB* are specifically expressed in lateral transverse muscles and control acquisition of growth-related LT muscle properties**

Among the genes encoding actin binding proteins and belonging to cluster 2 (Fig. 3C) *sals*, *dCryAB* and *Tpn41C* have already been found to be involved in sarcomere organisation of larval and adult muscles (Bai et al., 2007, Chechenova et al., 2017, Wojtowicz et al., 2015). However, function of none of them has been analysed during embryonic muscle formation. We first characterised expression patterns of selected cluster 2 genes and found that *Gelsolin* (*Gel*) and *dCryAB* are expressed majoritary in lateral transverse (LT) muscles starting from the embryonic stage 15 (Fig. 4A-A’). In stage 16 embryos, in addition to high LT muscle expression, a lower *Gel* transcript and protein levels are also detected in DT1 muscle (Fig. 4B-C’). Moreover, *Gel* is highly expressed in the developing fat body and in the visceral musculature (Fig. S4). Like *Gel*, the *dCryAB* transcripts are also detected in LT muscles from stage 15 and accumulate in parallel in DT1 (Fig. 4D). In stage 16 embryos *dCryAB* is also detected in additional set of muscles including the SBM, VA1, LO1 and VL1 (Fig. 4E), however its predominant expression remains in LTs and DT1. This restricted muscular expression of *Gel* and *dCryAB* suggested potential functions in the acquisition of LT muscle properties.

To test *Gel* and *dCryAB* functions in embryonic muscles we created conditional knock out for both genes using CRISPR-Cas9 technology (Jinek et al, 2012). We opted for conditional mutant strategy to avoid mutant effects of *Gel* in the fat body and visceral muscle (Fig. S4) and applied the same strategy for *dCryAB* to generate similar mutant contexts for both genes. First, to test efficacy of the conditional CRISPR knockout in syncytial muscles we generated analogous conditional mutant context for the LT muscle identity gene



*apterous*. We found that conditional CRISPR system phenocopies LT muscle defects observed in null *apterous* mutant embryos (Bourgouin et al., 1992) including the occasional loss of lateral muscles (Fig. S5). This observation suggests that conditional CRISPRs-based gene targeting in syncytial muscle is efficient enough and could serve for testing gene functions

We then created transgenic lines that ubiquitously express the guide RNA directed against all the isoforms encoded by the *Gel* and *dCryAB/l(2)efl* loci (Fig. 5A). We also generated double transgenics carrying muscle-specific Mef-GAL4 or Lms-GAL4 drivers combined with UAS-Cas9. Crossing lines with muscle targeted Cas9 expression to ubiquitous gRNA lines generated conditional, muscle specific knockouts for *Gel* and *dCryAB*. We first confirmed specific expression of Cas9 in targeted muscles (Fig. 5B, D-F) and found that the majority of embryos co-expressing Cas9 and *Gel* gRNA show aberrant enlarged shapes of LT muscles often associated with splitted extremities (Fig. 5C, C', arrows). Interestingly, LT-muscle targeted CRISPR knockout of *dCryAB* leads to highly reminiscent muscle defects (Fig. 5D, D' arrows). Enlarged myotube phenotypes suggested potential role of *Gel* and *dCryAB* in limiting/stopping myoblast fusion in LT muscles, whereas splitted fork-like myotube ends indicated that both genes function to limit unpolarized LT myotube growth. To test whether loss of *Gel* and *dCryAB* leads to additional fusion events we counted number of nuclei in the knockout conditions using Cas9 nuclear dots as a marker. Indeed, the enlarged, LT muscles contain 3-4 nuclei more compared to the wild type (10-11 instead of 6-7 nuclei) (Fig. 5E-G). Moreover, we found that fork-like extremities of splitted muscles contact ectopic tendon cells (Fig. 5E'-G'). Altogether, this suggests that *Gel* and *dCryAB* are activated in Lms-positive muscles in stage 15 embryos to fine-tune the growth related LT muscle properties including fusion arrest, proper LT muscle extension and attachment. We hypothesise that splitted LT muscle morphology in *Gel* and *dCryAB* loss of function contexts could be related to the additional fusion events that are incomplete leading to the partial membrane breakdown with the cytoplasm exchange but without endocytosis of nuclei. As all these phenotypes are not fully penetrant (not all segments and not all LT muscles within a segment are affected) we propose that it could be due to the incomplete efficacy of conditional CRISPR knockout system in syncytial muscle cells and/or potential functional redundancy of *Gel* and *dCryAB*. Redundancy of phenotypes observed in *Gel* and *dCryAB* mutant embryos also suggests that realisor genes with multiple identity acquisition roles form multi-gene modules to ensure spatial and temporal precision and functional robustness.

Altogether, our genome wide translational profiling of muscle subsets has led to identification of spatial and temporal gene expression signatures, which represent a unique resource for better understanding of genetic control of diversification of muscle types. Our data reveal that, in addition to *Org1*, the *dSix4* and *Ptx1* transcripts are enriched in Slou-positive muscles suggesting they could contribute to a conserved Tbx/Six/Ptx muscle identity module. Temporal transition profiles of enriched genes have led to defining clusters of genes with similar expression behaviour overtime and to identification of a new class of muscle identity realisor genes that control acquisition of several growth-related muscle properties. In particular we found that *Gel* and *dCryAB* encoding two actin-interacting proteins with potential actin capping/depolymerisation functions act as growth tuners of LT muscles. *Drosophila* *Gel* belongs to conserved Gelsolin/villin family of actin interactors (Nag et al., 2013) with actin depolymerisation activity (Chellaiah et al., 2000; Lu et al., 1997), whereas *dCryAB* codes for a small heat shock protein (sHSP) carrying acting binding domain

and known to ensure chaperon functions (Wojtowicz et al., 2015). The *in vivo* actin capping capacity of dCryAB remains hypothetical, however it is well known that sHSP proteins could inhibit actin polymerisation and act as capping proteins *in vitro* (Mounier and Arrigo, 2002). Thus, our data suggest that *Gel* and *dCryAB* by acting as muscle-type specific regulators of actin cytoskeleton, contribute to the acquisition of LT muscle identity program. Their functional redundancy and precise roles during setting growth related LT muscle properties remain to be investigated.

### **Acknowledgements**

We thank A. Irhashava and A. Joly for their technical help in analyses of candidate gene expression.

## **Materials and methods**

### **Fly Stocks**

All *D. melanogaster* stocks and crosses were grown on standard medium at 25°C. To generate the TRAP lines, the *UAS-RpL10a-EGFP* line was crossed with the *Slou-Gal4* (gift of M. Frasch, Universität Erlangen-Nürnberg, Germany), *Lms-Gal4* (Janelia Farm collection N°46861) and *Duf-Gal4* (gift of K. Vijayraghavan, TIFR, India) driver lines to specifically target polysomes in the respective muscle populations. For inducible CrispR-Cas9 gene knockout experiments the generated UAS-Cas9-P2A-BFP line was combined with *Mef2-Gal4* or *Lms-Gal4* drivers and crossed with gRNA lines generated for *gel* and *dCryAB* candidate genes.

### **Generation of *gel* and *dCryAB* gRNA lines**

The two gRNA for *gel* were designed using CRISPR optimal target finder (Gratz et al., 2014). To generate *gel* guide RNA1 we used 5'-GTCGAGACCTCGACCGATGAGGC-3' and 5'-AAACGCCTCATCGGTTCGAGGTCTC-3' and for *gel* guide RNA2 5'-GTCGAAATAGCTCAGAAACAGCT-3' and 5'-AAACAGCTGTTTCTGAGCTATTTTC-3' pair of primers. For *dCryAB* the following pair of primers was designed: 5'-CTTCGACCGAAGTGCTGGTCCAAG-3' and 5'-AAACCTTGACCAGCACTTCGGTC-3'. For *apterous* guide we used 5'-TGCAGACGTGAAAAACAAGATTTTC-3' and 5'-AAACGAAATCTTGTTTTTTCACGTC-3' pair of primers. Primers were annealed, digested by Bbs1 and cloned into pCFD3:U6:3 plasmid (plasmid #49410, Addgene) for *gel* and *dCryAB* and into pCFD5 (Plasmid #73914, Addgene) for *apterous*. Plasmids were injected in attP40 embryos (BestGene). gRNA carrying lines were established from the progeny of the injected flies.

### **Generation of *UAS-Cas9-P2A-BFP*.**

5X UAS sequence from pValium20 (Plasmid #2786918, Harvard Medical School) was amplified by using primers containing gateway sites.

(Primers used : 5'-GGGGACAAGTTTGTACAAAAAAGCAGGCTTTCGAGGTCGGAGTACTGTCCT-3' ; 5'-GGGGACCACTTTGTACAAGAAAGCTGGGTCGAGTCTCCGCTCGGAGGACA-3').

The UAS sequence was inserted in the pDEST-HemmarR vector (Plasmid #31222, Addgene) (Han et al., 2011) by gateway reaction (Gateway kit, ThermoFisher scientific). Cas9 and P2A-BFP was amplified respectively from pBS-Hsp70-Cas9 (Plasmid #46294, addgene) and pHRdSV40-dCas9-10xGCN4\_v4-P2A-BFP (plasmid #60903, Addgene) (Tanenbaum et al., 2014) using the following primers:

Hemmar : 5'-tttcgttagccaagactcgagATGGACTATAAGGACCACG-3'

P2A-BFP : 5'- ccgcttccCTTTTTCTTTTTGCCTGGC-3'

Cas9 : 5'-aaagaaaagGGAAGCGGAGCTACTAAC-3'

Hemmar : 5'ctggcttagttaattaattctagaTTAGACCTTCCTTTCTTTTTG-3'

Amplified Cas9 and P2A-BFP fragments were cloned in frame downstream of the UAS sequence in the XhoI-XbaI digested pDEST-hemmarR vector, by one Gibson reaction (Gibson assembly kit, New England Biolabs).

### **RNA extraction and RT-qPCR**

mRNA was extracted from Lms, Slou and Duf muscle populations using TRIzol reagent (Invitrogen) following manufacturer's instructions. RNA quality and quantity were then assessed using Agilent RNA 6000 Pico kit on Agilent 2100 Bioanalyzer (Agilent Technologies). RT-qPCR was performed to test the enrichment of positive muscle genes in TRAP RNA from muscle subsets compared to whole RNA fraction and using Rp49 gene as reference.

### **Microarray Analysis**

The Agilent 8x60K probes (60-mer) gene expression microarrays were used. We assessed the quality of triplicates using Pearson's correlation test. Correlations for all conditions were 85% or higher. The microarrays data were quantile quantile normalised. Gene expression data from Slou, Lms and Duf positive cells were compared to the whole embryo datasets to generate lists of genes differentially expressed, fold change  $\geq 2$ , p value  $< 0.05$ . GO Princeton software was used to assign GO classification. Then we compared Lms and Slou positive cells to Duf expressing cells to determine 2 lists of differentially regulated muscle specific genes, fold change  $\geq 2$ , p value  $< 0.05$ . We computed and compared GO biological processes from these two lists using R package Cluster profiler (Yu et al., 2012).

**Spatial heatmap.** From lists generated by the comparison of Slou and Lms positive cells to Duf population, we compared fold change of every selected gene of both population overtime and we kept only those that have a significant difference between them (ratio  $> 1.5$ ). We distinguished two states for each population : in one hand we considered genes that are only up-regulated in one population and on the second genes that are up-regulated in the two populations but displaying significant difference from one population to another. Heatmap was generated by the union of these genes from the different time points and by keeping genes proportions for each case. Note that we did not consider genes present in different states overtime.

**Temporal heatmap.** Translational temporal profiles from Slou and Lms at the three time points were converted to "transition values" defined as log ratios between T2 and T1, T3 and T2. Transition values were considered as three discrete classes: upregulated ( $> 1$ ), stable (between -1 and 1), downregulated ( $< 1$ ). Thus, expression profiles from Lms and Slou muscles, which contain three temporal windows, were converted into vectors of two transitions (Tr1, Tr2), which allow the determination of correct genes behaviour. For example the profile "red-red" regroups genes whose RNA level increased between T2 and T1 (Tr1), and then continue to increase between T3 and T2 (Tr2).

### **In situ hybridization and immunostaining.**

Embryos were dechorionated and fixed in 4% paraformaldehyde/heptane for all immunohistochemistry. Fluorescent in situ hybridization with TSA amplification system (Perkin-Elmer) and immunohistochemistry was done as described previously (Junion et al., 2007). To generate RNA probes for *gel* (primers used: 5'-5'AATCGACTCCGTGGTACTC-3' and

5'-GGGAGGCCAAAGATGAGCTGTC-3'), for *CG10565* (5'-CGCTCGTCGATTTAGCCTAC-3' and 5'-GGTCCAACCTCGGCAGCTATCTT-3'), for *CG10654* (5'-ACGAGTCAGCCAAACAGGAT-3' and 5'-AAAGCACTTGCCACACTCCT-3'), for *EndoGl* (5'-AGTTGTGCTGGTGAAGAAGCAAC-3' and 5'-GGAGTGAAACCAGCATCACCTTGT-3'), for *Ptx1* (5'-CTCGGTGAACTGCTTCAACA-3' and 5'-AACCGAGATCGGAGTGGTC-3') and for *dSix4* (5'-ATGTTTGACAAGAATTTGGA-3' and 5'-ATCCATTGCCATCCAGTTGG-3') the corresponding DNA sequences were cloned by PCR in pCR II topo vector. The corresponding anti-sense RNAs were transcribed *in vitro* using T7 or SP6 RNA polymerase. For *dCryAB*, Gold collection clone *GH01960* was used to generate RNA probe. For fluorescent staining, the following antibodies were used: rabbit anti- $\beta$ 3 tubulin (1:5000; R. Renkawitz-Pohl, Philipps University, Marburg, Germany), rabbit anti-Ptx1 (1:500, G. Vorbrüggen, Max-Planck-Institut, Germany) guinea pig anti-Stripe (1:1000; T. Volk, Weizmann Institute of Science, Isreal), rat anti-Actin (1:300, MAC 237; Babraham Bioscience Technologies), anti-GFP (1:1000 Developmental Studies Hybridoma Bank (DSHB)), anti-Cas9 (1/2000; Abcam). Cy3, Cy5, and 488 conjugated secondary antibodies were used (1:300; Jackson Immuno-Research). Embryos were mounted in anti-fade Fluoromount-G reagent (Southern Biotech). Labeled embryos were analyzed using an SP8 (Leica) equipped with HyD detector, with a 40X objective. Images were processed with ImageJ.

### TRAP experiment

RPL10aGFP-tagged embryos were collected and messenger RNAs from the different muscle populations isolated as described in Bertin et al., 2015.

### References

- Bai, J., Hartwig, J.H., and Perrimon, N. (2007). SALS, a WH2-Domain-Containing Protein, Promotes Sarcomeric Actin Filament Elongation from Pointed Ends during *Drosophila* Muscle Growth. *Dev. Cell* *13*, 828–842.
- Bataillé, L., Delon, I., Da Ponte, J.P., Brown, N.H., and Jagla, K. (2010). Downstream of Identity Genes: Muscle-Type-Specific Regulation of the Fusion Process. *Dev. Cell* *19*, 317–328.
- Bate, M. (1990). The embryonic development of larval muscles in *Drosophila*. *Development* *110*, 791–804.
- Bate, M., Rushton, E., and Currie, D.A. (1991). Cells with persistent twist expression are the embryonic precursors of adult muscles in *Drosophila*. *Development* *113*, 79–89.
- Baylies, M.K., Bate, M., and Gomez, M.R. (1998). Myogenesis: a view from *Drosophila*. *Cell* *93*, 921–927.
- Beckett, K., and Baylies, M.K. (2007). 3D analysis of founder cell and fusion competent myoblast arrangements outlines a new model of myoblast fusion. *Dev. Biol.* *309*, 113–125.
- Bertin, B., Renaud, Y., Aradhya, R., Jagla, K., and Junion, G. (2015). TRAP-rc, Translating Ribosome Affinity Purification from Rare Cell Populations of *Drosophila* Embryos. *J. Vis. Exp.*
- Bodmer, R. (1993). The gene *tinman* is required for specification of the heart and visceral muscles in *Drosophila*. *Development* *118*, 719–729.
- Boukhatmi, H., Frendo, J.L., Enriquez, J., Crozatier, M., Dubois, L., and Vincent, A. (2012). *Tup/Islet1* integrates time and position to specify muscle identity in *Drosophila*. *Development* *139*, 3572–3582.
- Bourgouin, C., Lundgren, S, and Thomas, JB. (1992). *apterous* is a *Drosophila* LIM Domain Gene Required for the Development of a Subset of Embryonic Muscles. *Neuron* *9*, 549–

561.

Busser, B.W., Shokri, L., Jaeger, S.A., Gisselbrecht, S.S., Singhania, A., Berger, M.F., Zhou, B., Bulyk, M.L., and Michelson, A.M. (2012). Molecular mechanism underlying the regulatory specificity of a *Drosophila* homeodomain protein that specifies myoblast identity. *Development* *139*, 1164–1174.

Callahan, C.A., Bonkovsky, J.L., Scully, A.L., and Thomas, J.B. (1996). *derailed* is required for muscle attachment site selection in *Drosophila*. *Development* *122*, 2761–2767.

Carmena, A., Gisselbrecht, S., Harrison, J., Jiménez, F., and Michelson, A.M. (1998). Combinatorial signaling codes for the progressive determination of cell fates in the *Drosophila* embryonic mesoderm. *Genes Dev.* *12*, 3910–3922.

Carrasco-Rando, M., Tutor, A.S., Prieto-Sánchez, S., González-Pérez, E., Barrios, N., Letizia, A., Martín, P., Campuzano, S., and Ruiz-Gómez, M. (2011). *Drosophila* Araucan and Caupolican Integrate Intrinsic and Signalling Inputs for the Acquisition by Muscle Progenitors of the Lateral Transverse Fate. *PLoS Genet.* *7*, e1002186.

Chechenova, M.B., Maes, S., Oas, S.T., Nelson, C., Kiani, K.G., Bryantsev, A.L., and Cripps, R.M. (2017). Functional redundancy and nonredundancy between two Troponin C isoforms in *Drosophila* adult muscles. *Mol. Biol. Cell* *28*, 760–770.

Chellaiah, M., Kizer, N., Silva, M., Alvarez, U., Kwiatkowski, D., and Hruska, K.A. (2000). Gelsolin deficiency blocks podosome assembly and produces increased bone mass and strength. *J. Cell Biol.* *148*, 665–678.

Clark, I.B.N., Boyd, J., Hamilton, G., Finnegan, D.J., and Jarman, A.P. (2006). *D-six4* plays a key role in patterning cell identities deriving from the *Drosophila* mesoderm. *Dev. Biol.* *294*, 220–231.

D'Alessio, M., and Frasch, M. (1996). *Msh* may play a conserved role in dorsoventral patterning of the neuroectoderm and mesoderm. *Mech. Dev.* *58*, 217220–231.

Davis, G.W., Schuster, C.M., and Goodman, C.S. (1997). Genetic analysis of the mechanisms controlling target selection: target-derived Fasciclin II regulates the pattern of synapse formation. *Neuron* *19*, 561–573.

Dobi, K.C., Halfon, M.S., and Baylies, M.K. (2014). Whole-Genome Analysis of Muscle Founder Cells Implicates the Chromatin Regulator Sin3A in Muscle Identity. *Cell Rep.*

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

Estrada, B., Gisselbrecht, S.S., and Michelson, A.M. (2007). The transmembrane protein *Perdido* interacts with *Grip* and integrins to mediate myotube projection and attachment in the *Drosophila* embryo. *Development* *134*, 4469–4478.

Heiman, M., Schaefer, A., Gong, S., Peterson, J.D., Day, M., Ramsey, K.E., Suárez-Fariñas, M., Schwarz, C., Stephan, D.A., Surmeier, D.J., et al. (2008). A Translational Profiling Approach for the Molecular Characterization of CNS Cell Types. *Cell* *135*, 738–748.

Inaki, M., Yoshikawa, S., Thomas, J.B., Aburatani, H., and Nose, A. (2007). *Wnt4* Is a Local Repulsive Cue that Determines Synaptic Target Specificity. *Curr. Biol.* *17*, 1574–1579.

Ip, Y.T., Park, R.E., Kosman, D., Yazdanbakhsh, K., and Levine, M. (1992). *dorsal-twist* interactions establish *snail* expression in the presumptive mesoderm of the *Drosophila* embryo. *Genes Dev.* *6*, 1518–1530.

Junion, G., Bataille, L., Jagla, T., Da Ponte, J.P., Tapin, R., and Jagla, K. (2007). Genome-wide view of cell fate specification: *ladybird* acts at multiple levels during diversification of muscle and heart precursors. *Genes Dev.* *21*, 3163–3180.

Knirr, S., Azpiazu, N., and Frasch, M. (1999). The role of the NK-homeobox gene *slouch* (*S59*)

in somatic muscle patterning. *Development* 126, 4525–4535.

Kose, H., Rose, D., Zhu, X., and Chiba, A. (1997). Homophilic synaptic target recognition mediated by immunoglobulin-like cell adhesion molecule Fasciclin III. *Development* 124, 4143–4152.

Kurusu, M., Cording, A., Taniguchi, M., Menon, K., Suzuki, E., and Zinn, K. (2008). A Screen of Cell-Surface Molecules Identifies Leucine-Rich Repeat Proteins as Key Mediators of Synaptic Target Selection. *Neuron* 59, 972–985.

Landgraf, M., Bossing, T., Technau, G.M., and Bate, M. (1997). The origin, location, and projections of the embryonic abdominal motoneurons of *Drosophila*. *J. Neurosci.* 17, 9642–9655.

Lu, M., Witke, W., Kwiatkowski, D.J., and Kosik, K.S. (1997). Delayed retraction of filopodia in gelsolin null mice. *J. Cell Biol.* 138, 1279–1287.

Müller, D., Jagla, T., Bodart, L.M., Jährling, N., Dodt, H.-U., Jagla, K., and Frasch, M. (2010). Regulation and Functions of the *lms* Homeobox Gene during Development of Embryonic Lateral Transverse Muscles and Direct Flight Muscles in *Drosophila*. *PLoS ONE* 5, e14323.

Nag, S., Larsson, M., arten, Robinson, R.C., and Burtnick, L.D. (2013). Gelsolin: the tail of a molecular gymnast. *Cytoskeleton* 70, 360–384.

Nose, A. (2012). Generation of neuromuscular specificity in *Drosophila*: novel mechanisms revealed by new technologies. *Front. Mol. Neurosci.* 5.

Nose, A., Umeda, T., and Takeichi, M. (1997). Neuromuscular target recognition by a homophilic interaction of connectin cell adhesion molecules in *Drosophila*. *Development* 124, 1433–1441.

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

Rose, D., Zhu, X., Kose, H., Hoang, B., Cho, J., and 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.

Rushton, E., Drysdale, R., Abmayr, S.M., Michelson, A.M., and Bate, M. (1995). Mutations in a novel gene, *myoblast city*, provide evidence in support of the founder cell hypothesis for *Drosophila* muscle development. *Development* 121, 1979–1988.

Schaub, C., Nagaso, H., Jin, H., and Frasch, M. (2012). *Org-1*, the *Drosophila* ortholog of *Tbx1*, is a direct activator of known identity genes during muscle specification. *Development* 139, 1001–1012.

Schnorrer, F., Kalchauer, I., and Dickson, B.J. (2007). The Transmembrane Protein *Kon-tiki* Couples to *Dgrip* to Mediate Myotube Targeting in *Drosophila*. *Dev. Cell* 12, 751–766.

Swan, L.E., Schmidt, M., Schwarz, T., Ponimaskin, E., Prange, U., Boeckers, T., Thomas, U., and Sigrist, S.J. (2006). Complex interaction of *Drosophila* GRIP PDZ domains and Echinoid during muscle morphogenesis. *EMBO J.* 25, 3640–3651.

The modENCODE Consortium, Roy, S., Ernst, J., Kharchenko, P.V., Kheradpour, P., Negre, N., Eaton, M.L., Landolin, J.M., Bristow, C.A., Ma, L., et al. (2010). Identification of Functional Elements and Regulatory Circuits by *Drosophila* modENCODE. *Science* 330, 1787–1797.

Thomas, A., Lee, P.-J., Dalton, J.E., Nomie, K.J., Stoica, L., Costa-Mattioli, M., Chang, P., Nuzhdin, S., Arbeitman, M.N., and Dierick, H.A. (2012). A Versatile Method for Cell-Specific Profiling of Translated mRNAs in *Drosophila*. *PLoS ONE* 7, e40276.

Tixier, V., Bataillé, L., and Jagla, K. (2010). Diversification of muscle types: Recent insights from *Drosophila*. *Exp. Cell Res.* 316, 3019–3027.

Tryon, R.C., Pisat, N., Johnson, S.L., and Dougherty, J.D. (2013). Development of translating

ribosome affinity purification for zebrafish: Zebrafish Translating Ribosome Purification. *Genesis* 51, 187–192.

Vorbrüggen, G., Constien, R., Zilian, O., Wimmer, E.A., Dowe, G., Taubert, H., Noll, M., and Jäckle, H. (1997). Embryonic expression and characterization of a Ptx1 homolog in *Drosophila*. *Mech. Dev.* 68, 139–147.

Watson, F.L., Mills, E.A., Wang, X., Guo, C., Chen, D.F., and Marsh-Armstrong, N. (2012). Cell type-specific translational profiling in the *Xenopus laevis* retina. *Dev. Dyn.* 241, 1960–1972.

Winberg, M.L., Mitchell, K.J., and 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.

Wojtowicz, I., Jab onska, J., Zmojdian, M., Taghli-Lamalle, O., Renaud, Y., Junion, G., Daczewska, M., Huelsmann, S., Jagla, K., and Jagla, T. (2015). *Drosophila* small heat shock protein CryAB ensures structural integrity of developing muscles, and proper muscle and heart performance. *Development* 142, 994–1005.

Ypsilanti, A.R., Zagar, Y., and Chedotal, A. (2010). Moving away from the midline: new developments for Slit and Robo. *Development* 137, 1939–1952.

## Figure Legends

**Figure 1: Translational profiling of embryonic muscle subsets during acquisition of individual muscle properties using TRAP approach.** **A.** TRAP-targeted muscle populations revealed by anti-GFP immunostaining of stage 16 embryos showing expression of Rpl10a-eGFP in six muscles per hemisegment (Slou>Rpl10a-eGFP), in four lateral muscles (Lms>Rpl10a-eGFP) and in the global musculature (Duf>Rpl10a-eGFP). **B.** TRAP pipeline. Arrow depicts the extents of muscle developmental processes with respect to the TRAP time windows (T1-T3). T1 corresponds to the specification of founder cells and the first wave of myoblast fusion (7-10 h AEL); T2 encompasses second wave of myoblast fusion, myotube elongation and attachment to tendon cells (10-13 AEL); T3 window covers stabilisation of muscle attachment and their innervation by motoneurons. Below is shown a scheme of the main steps of the TRAP method. GFP-tagged polysomes from transgenic embryos expressing Rpl10A-eGFP in muscle subsets are affinity purified using magnetic beads coated with anti-GFP antibodies. Associated mRNA engaged in translation (pink) is extracted, purified and used for global gene expression analyses. **C.** qRT-PCR validation of TRAP approach. An example of gene expression analysis using TRAP-isolated mRNA from Slou-positive muscles showing enrichment of muscle gene expression (Mef2 and Slou) and depletion of non-muscle gene expression (Prosp, SoxN) compared to the input RNA from the whole embryo. **D.** Correlation graph based on two replicates from the same TRAP condition shows high quality of the two replicates ( $R= 0.97$ ) **E.** Bar plots showing proportions of genes whose expression is commonly up (FC>2) in both individual muscle subsets and in all muscles (dark red and dark blue) versus genes up regulated exclusively in Slou (light red) and Lms (light blue) populations. **F.** Pie chart showing distribution of GO terms (biological processes) for genes commonly up-regulated (orange) or down-regulated (grey) in Slou and Lms positive muscles and at all time windows.

**Figure S1: Specificity and sensitivity assessment of translational profiling by the TRAP method.** **A.** Volcano plot depicting the distribution of enriched (red) and depleted (blue) genes identified in Slou and Lms muscle subsets. Black dots represent examples of genes

enriched or depleted in Lms and Slou muscle population. **B.** Distribution of Slou and Lms up-regulated and down-regulated genes from T2 time window with respect to modEncode RNA expression level based on whole embryo RNA sequencing (stage 12-16). **C.** Table summarising the fold change enrichment of muscle specific genes expressed (red part) or absent (blue part) in Slou and Lms populations. Numbers in red highlight fold change of identity genes required for specification of Slou or Lms positive muscles.

**Figure 2: Spatial comparison of Lms and Slou up-regulated genes reveals new potential identity players expressed in Slou muscle population.** **A.** heatmap of genes encoding regulatory proteins with potential role in specification of muscle identity (DNA, RNA and chromatin binding proteins) and isolated from the cluster of genes up regulated in Slou population (A3, see proportional heatmap Figure S2). **B.** Database information (Symbol, Flybase gene ID, Protein Type and human orthologue) on the selected genes. **C-H.** RNA and/or protein expression profiles of EndoGl, CG10565, CG10654, Six4 and Ptx1 displaying enriched or specific expression in subsets of ventral muscle precursors corresponding to Slou positive cells (arrows). For example EndoGl is enriched in Slou-positive VA1 whereas dSix4 expression is specific to VA1 and VT1.

**Figure S2: Spatial behaviour of potential new muscle identity factors upregulated in Lms and/or Slou population overtime.** A simplified proportional heatmap showing four clusters of genes enriched at least at one time window in Lms or in Slou muscle subsets and showing differential expression ( $>1,5$ ) compared with expression in all muscles (Duf population). **A1-4.** Heatmap zoom showing candidate regulatory factors with higher expression levels in Lms (A1-2) or in Slou (A4) positive cells (A3 presented in Figure 2) B. Database information (Symbol, Flybase gene ID, Protein Type and human orthologue) are shown for the selected candidate genes.

**Figure 3: Temporal gene expression signatures of Slou and Lms positive muscle subsets identify new actin regulators enriched in Lms muscle cells.** **A.** Heatmap showing the transition profiles Tr1 and Tr2 in Lms and Slou population. Tr1 represents transition T2 versus T1 and Tr2 the transition T3 versus T2. **B.** GO comparison of genes from the two opposite transition clusters: Cluster 1 (green) and Cluster 2 (red). An R package “Cluster profiler” has been applied to identify differential GO terms. **C.** Plots of transition profiles for selected candidate genes reveal Lms-synexpression group of genes encoding actin interacting proteins (from cluster 2) and Notch components group with mirroring Lms/Slou profiles (from cluster 1). x-axis indicates the temporal transition points, y-axis indicates the log2 signal value and each line corresponds to one specific gene.

**Figure S3: Global representation of temporal transition profiles and associated GOs.** **A.** temporal transition heatmap corresponding to a detailed view of the heatmap from figure 3 showing gene names. **B** Cluster profiler plot representing GOs (biological process) comparison of upregulated genes from the two populations during the two transitions (T2/T1 and T3/T2). **C.** Transition profiles of genes potentially involved in neuromuscular junction formation.

**Figure 4: Gelsolin and dCryAB display muscle type specific expression pattern.** *Gelsolin* RNA expression of stage 15 (A) and 16 (B) embryos stained with *gel* RNA probe (A-A', B-B') and  $\beta 3$

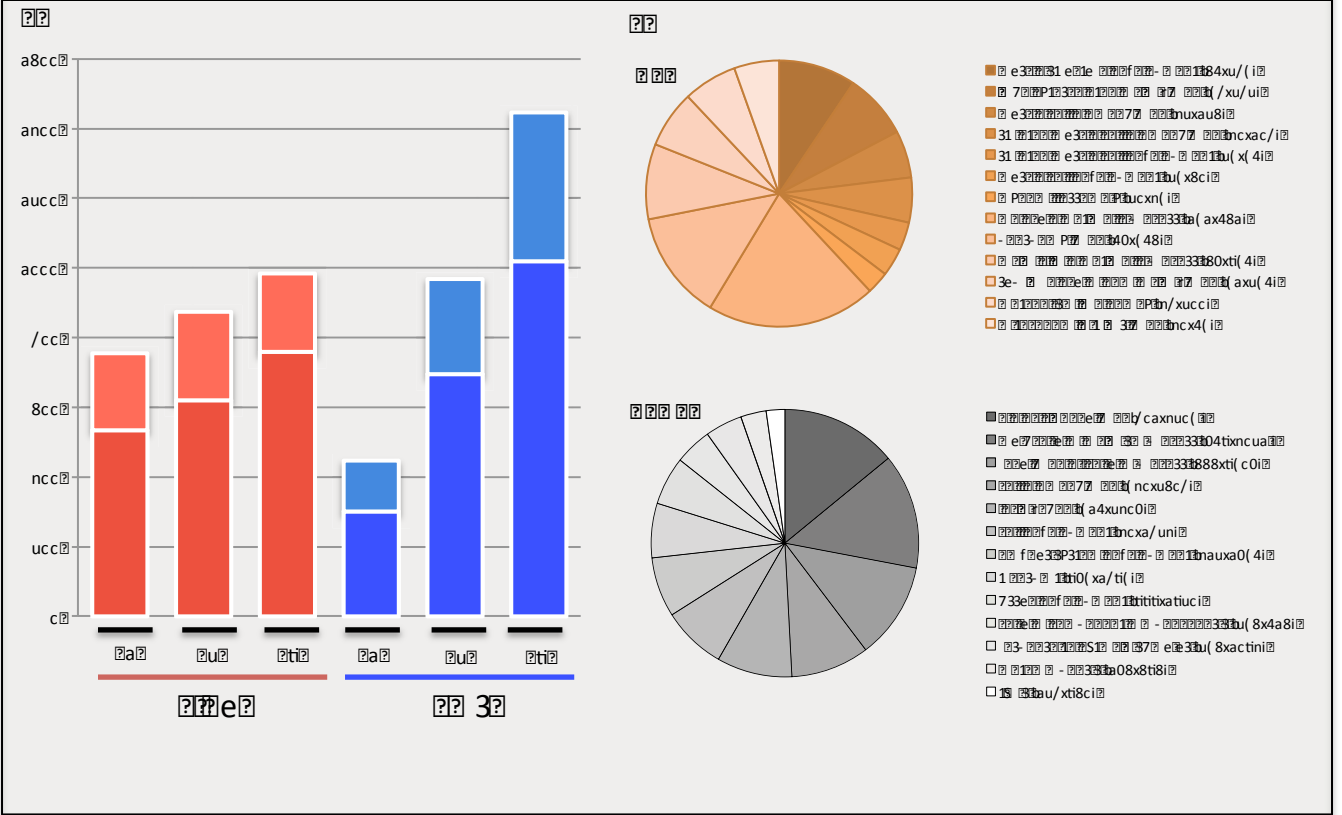
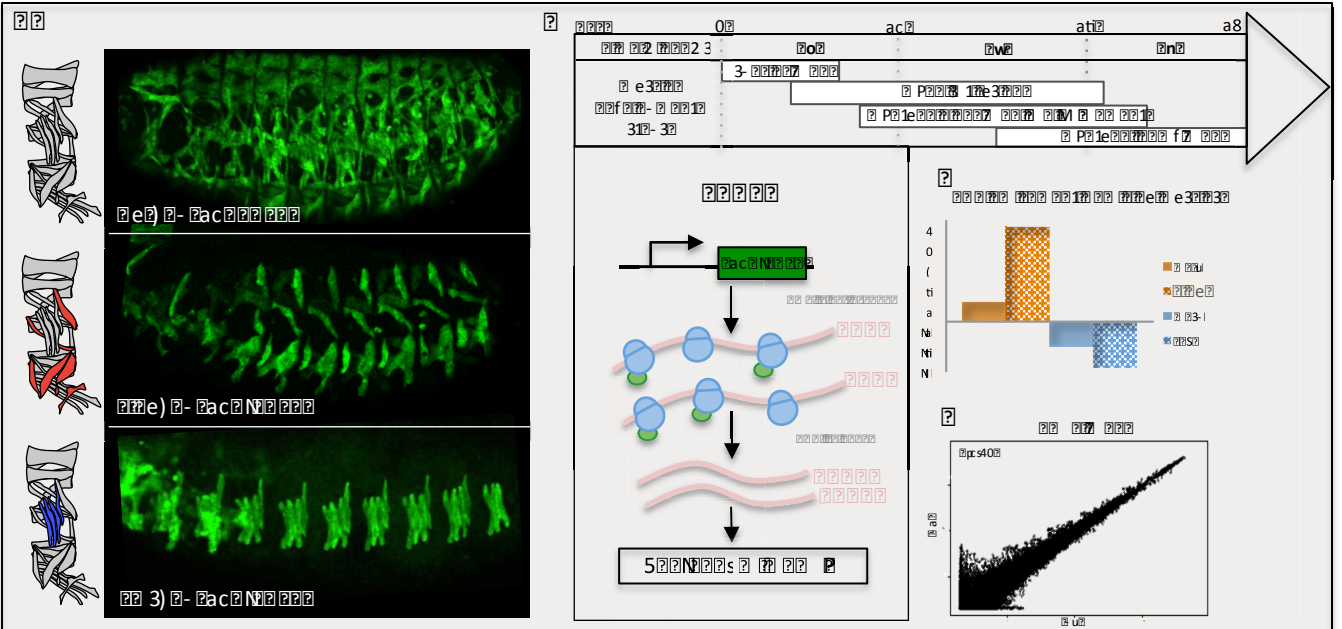


tubulin antibody (A'',B''). *Gel* transcripts are first detected in four LTs muscles at early stage 15 and then appear in DT1 muscle at stage 16. The protein pattern observed with *Gel* antibody (C,C',C'') reflects RNA expression. *dCryAB* RNA pattern at early stage 15 (D) and stage 16 (E) revealed with *dCryAb* RNA probe. *dCryAB* like *Gel* starts to be expressed in early stage 15 and is detected in LTs and in DT1 muscle. Its expression extends at stage 16 to the SBM, VA2 and DA1 muscles with predominant expression in LTs. *dCryAB* protein profile follows the RNA pattern and shows a dotted aspects and accumulations at muscle extremities and at lateral muscle sides (F,F'). G. Schematic representation of *Gelsolin* and *dCryAB* expression profile. Dark red and light red refer to the intensity and timing of expression (light red, low and late stage expression).

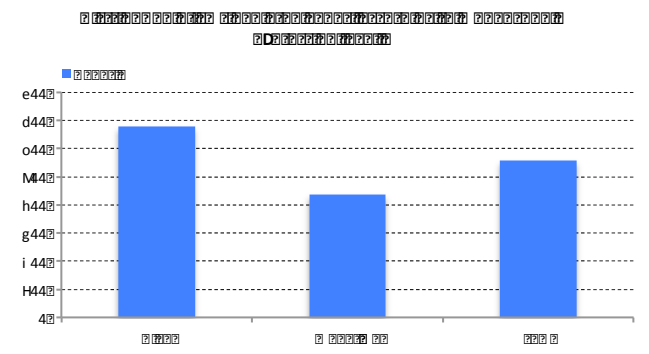
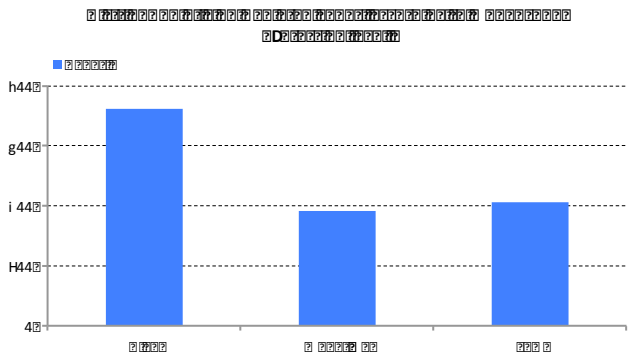
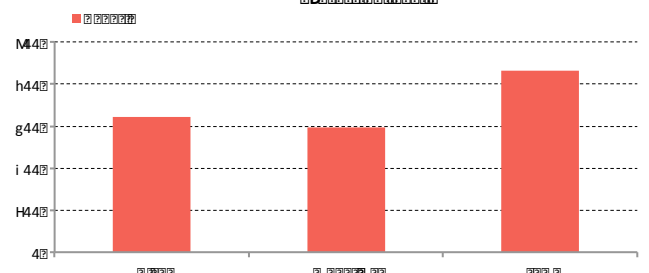
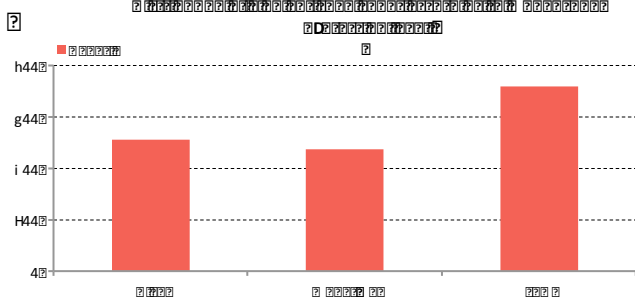
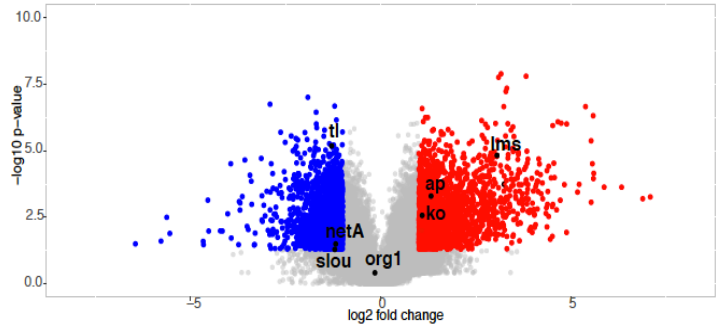
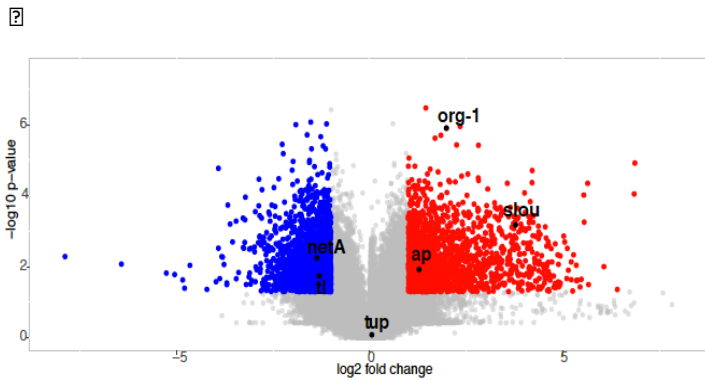
**Figure S4: Gelsolin is also expressed in other mesodermal tissue.** *Gelsolin* RNA expression of stage 15 embryo stained with *Gel* RNA probe (A-C) and  $\beta$ 3 tubulin antibody (B-D). *Gel* transcripts are found in visceral muscle and fat body (arrows).

**Figure 5: Conditional CRISPR/Cas9 Knock out of *Gelsolin* and *dCryAB* lead to LT muscle defects.** **A.** *Gel* and *dCryAB* guide RNAs localisation is depicted by red dashed rectangles within the *gel* and *dCryAB* loci. **B-C** Muscle pattern in *Mef2>Cas9-BFP* (control), *Mef>Cas9-BFP*; *gel* guide1 and *Lms>Cas9-BFP*; *dCryAB* guide 1 mutant embryos stained with Actin and Cas9 antibodies. A wild type muscle pattern is detected in control embryo, in contrast to LT muscle splitting defects (arrows) found in conditional KO embryo for *Gel* and *dCryAB*. Note that loss of *Gel* also leads to occasional splitting of DT1 muscle and that the majority of defects observed in *dCryAB* CRISPRs mutants is seen in LT3. **D-F.** Number of nuclei and muscle attachment to tendon cells in *Lms>Cas9* (control), *lms>Cas9*; *gel* guide 1 and *Lms>Cas9* mutant embryos stained with Actin, Cas9 and Stripe antibody. Seven nuclei and four tendon cells are detected in control embryo, in contrast LT muscle in conditional KO for *gel* and *dCryAB* display 10-11 nuclei (white dots) and 5-6 tendon cells (arrows).

**Figure S5: Conditional CRISPR/Cas9 Knockout of LT muscle identity gene *apterous* leads to severe LT muscle phenotypes.** **A** Muscle pattern in *Mef2>Cas9-BFP* *apterous* guide1 mutant embryos stained with Actin and Cas9 antibodies. Muscle targeted conditional CRISPR mutant embryos for *apterous* display phenotypes restricted to LT muscles that resemble phenotypes observed in null *apterous* mutants (loss or aberrant shapes (asterisks) of some of LT muscles (A-A')).



e3: epidermis (1.84x), cortex (7.0x), endodermis (3.1x)  
 3: epidermis (1.1x), cortex (3.1x), endodermis (3.1x)



??-H?

		??	??	??	??	??	??	??	??	??	??	??	??	??	
??	??	i	h e	i h	i d	e d	g H	H r	utito	g d	Hh H	4 h	He	4 M	H
??	??	i o	4 o	Ho	i H	d H	i H	oto	4 e	H	H4	4 h	H d	4 M	4 e

ur

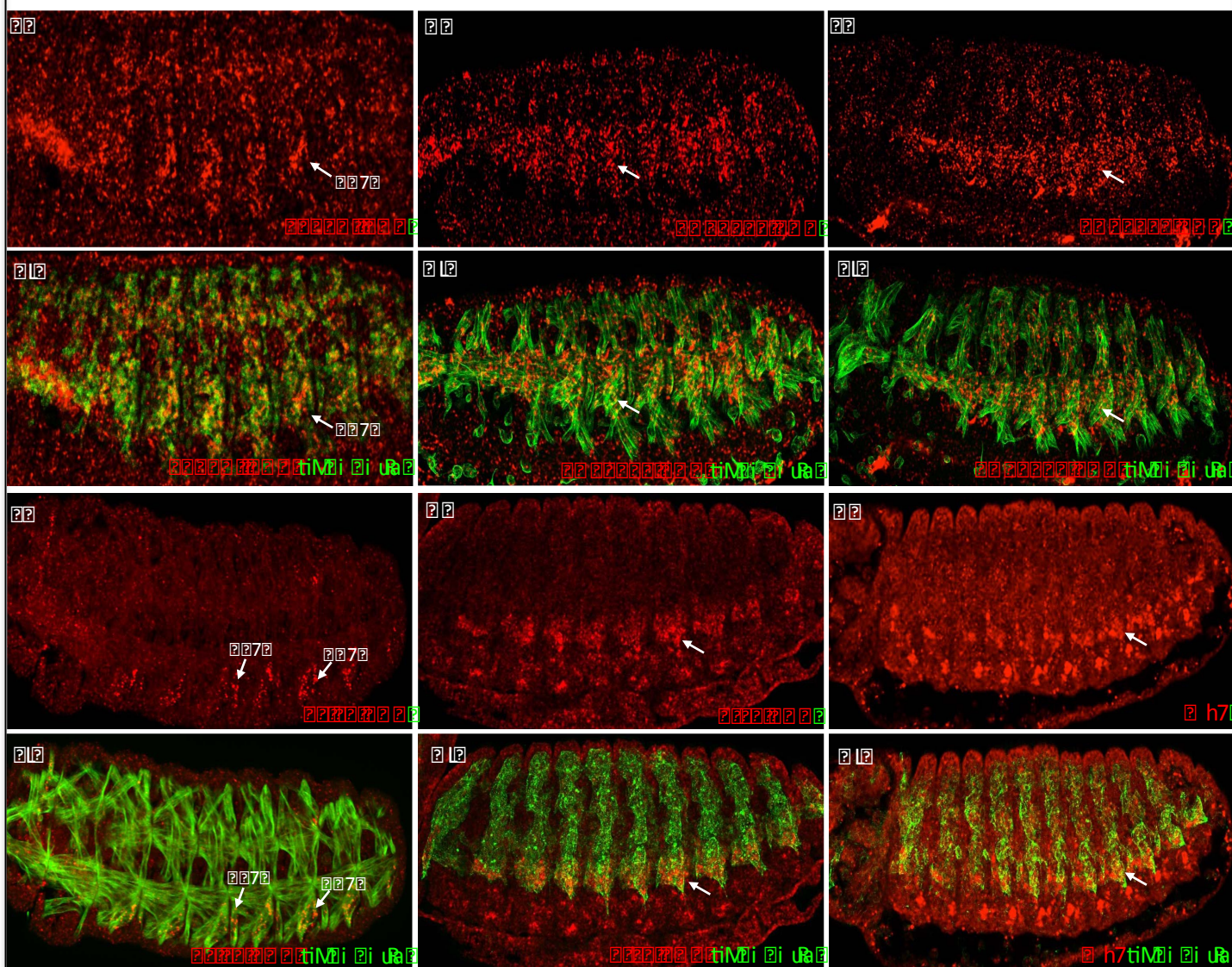
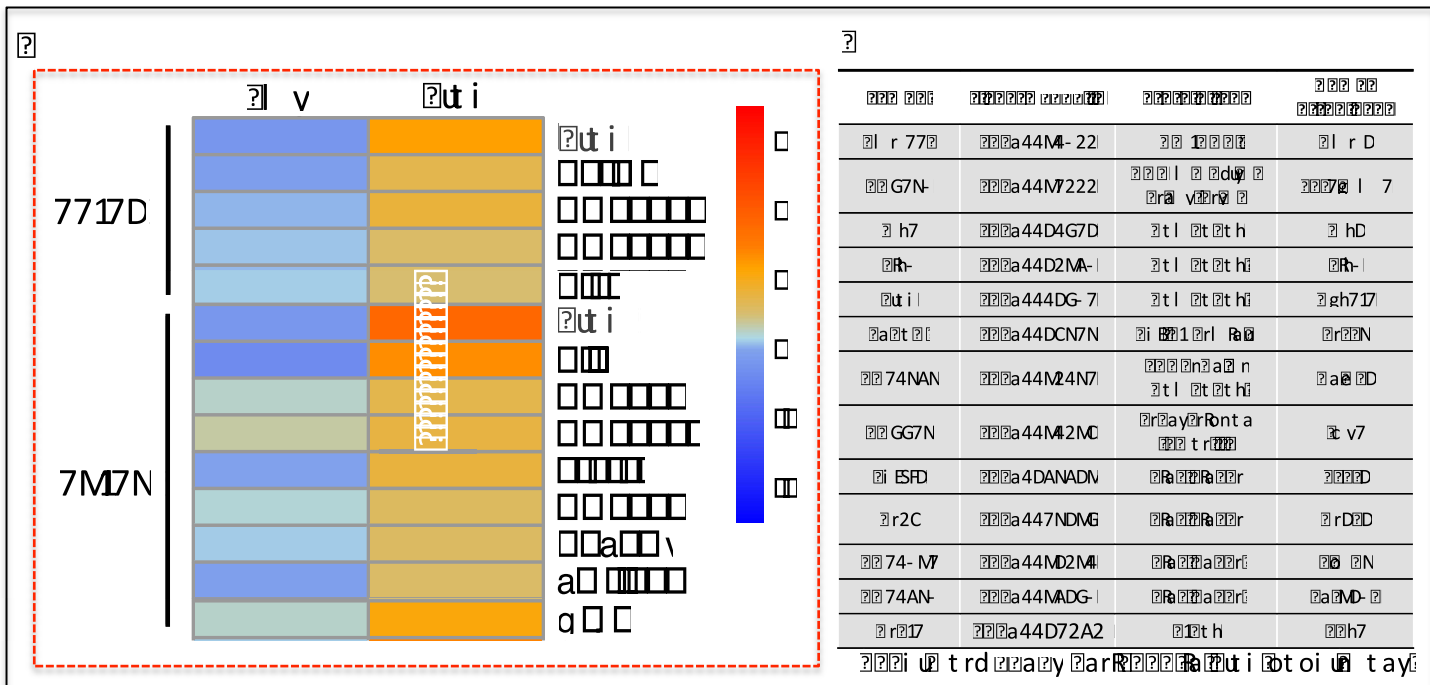
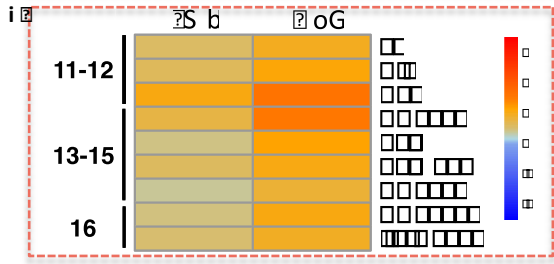
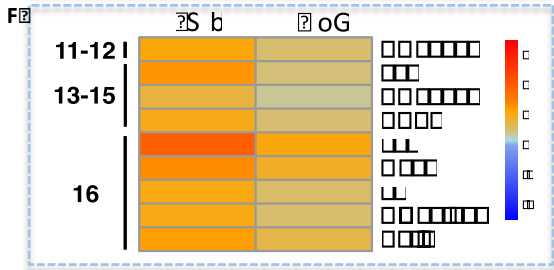
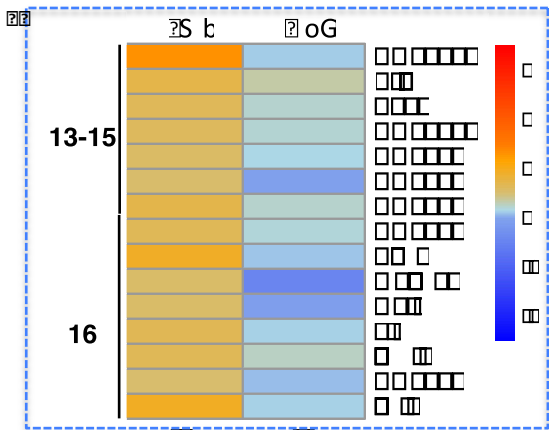
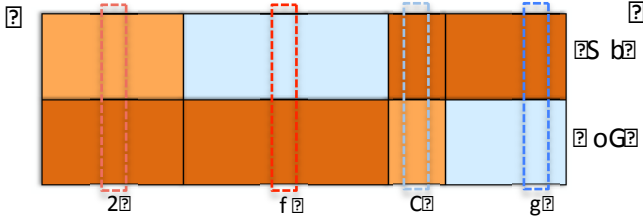


Figure 1: Gene expression patterns and localization. The heatmap and table show expression levels for various genes across different conditions. The fluorescence microscopy images show the localization of these genes in various tissues, with white arrows indicating specific features.

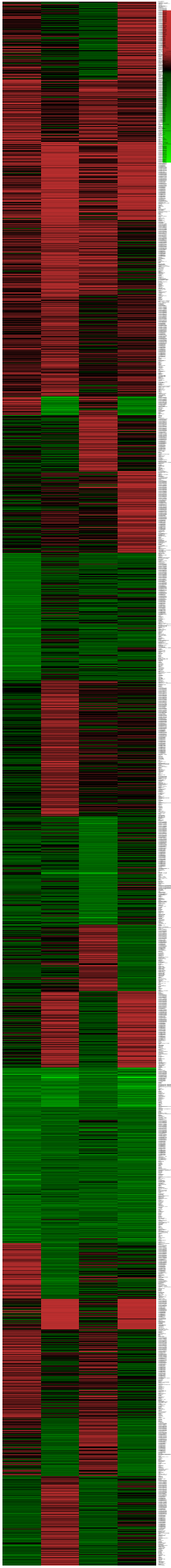


Region	Label	Gene	RefSeq	Ensembl
1	S	hhHc60	NC_023836.2	ENSG00000223098
1	b	hhC6Zg	NC_023836.2	ENSG00000223098
2	f	hhCZCC	NC_023836.2	ENSG00000223098
3	o	hhfgcZ	NC_023836.2	ENSG00000223098
4	g	hh2ghg	NC_023836.2	ENSG00000223098
5		hh26fC	NC_023836.2	ENSG00000223098
6		hh260C	NC_023836.2	ENSG00000223098
7		hh2Zgf	NC_023836.2	ENSG00000223098
8		hhghf gf	NC_023836.2	ENSG00000223098
9		hhgg26g	NC_023836.2	ENSG00000223098
10		hhCg6ZB	NC_023836.2	ENSG00000223098
11		hhC8f f 2	NC_023836.2	ENSG00000223098
12		hhC62f 2	NC_023836.2	ENSG00000223098
13		hhC606f	NC_023836.2	ENSG00000223098
14		hhCZg26	NC_023836.2	ENSG00000223098
15		hhf CggB	NC_023836.2	ENSG00000223098
16		hhf CC80	NC_023836.2	ENSG00000223098
17		hhf Cf 66	NC_023836.2	ENSG00000223098
18		hhf 2hOC	NC_023836.2	ENSG00000223098
19		hhf 26Cg	NC_023836.2	ENSG00000223098
20		hhf Of ZO	NC_023836.2	ENSG00000223098
21		hhf 06gC	NC_023836.2	ENSG00000223098
22		hhf 06f 6	NC_023836.2	ENSG00000223098
23		hhf 8C8B	NC_023836.2	ENSG00000223098
24		hhf 80gh	NC_023836.2	ENSG00000223098
25		hhf 8ZCf	NC_023836.2	ENSG00000223098
26		hhf 6hOf	NC_023836.2	ENSG00000223098
27		hhf 6ZBg	NC_023836.2	ENSG00000223098
28		hhf ZCCO	NC_023836.2	ENSG00000223098
29		hhf ZZf 6	NC_023836.2	ENSG00000223098
30		hh2COZO	NC_023836.2	ENSG00000223098
31		hhBh22f	NC_023836.2	ENSG00000223098
32		hhBg6gC	NC_023836.2	ENSG00000223098
33		hhBCbf C	NC_023836.2	ENSG00000223098
34		hc0g68C	NC_023836.2	ENSG00000223098
35		hc08Z86	NC_023836.2	ENSG00000223098

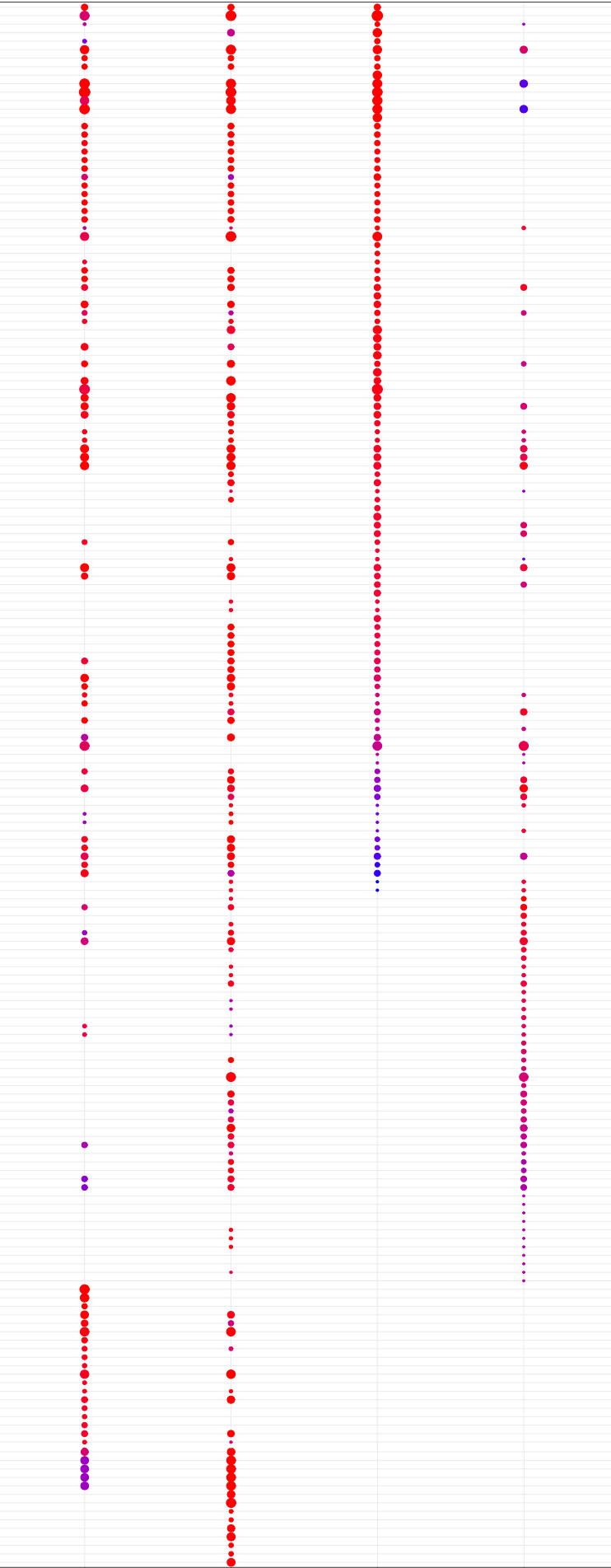


??

??



Gene identifiers and their corresponding expression values are listed in this column. The text is partially illegible due to the small font size and overlapping characters, appearing as a dense list of symbols and numbers.



p.adjust  
0.075  
0.050  
0.025

GeneRatio  
0.05  
0.10  
0.15

2?2?2?2?2?2? 2? 2?2?2? 2?2?2?2?2?2? 2?2?2?2?2?2?

2? 2?2?2?2?

2? 2?2?2?2?

2?2?2?2?2?2?

2?2?2?2?2?2?





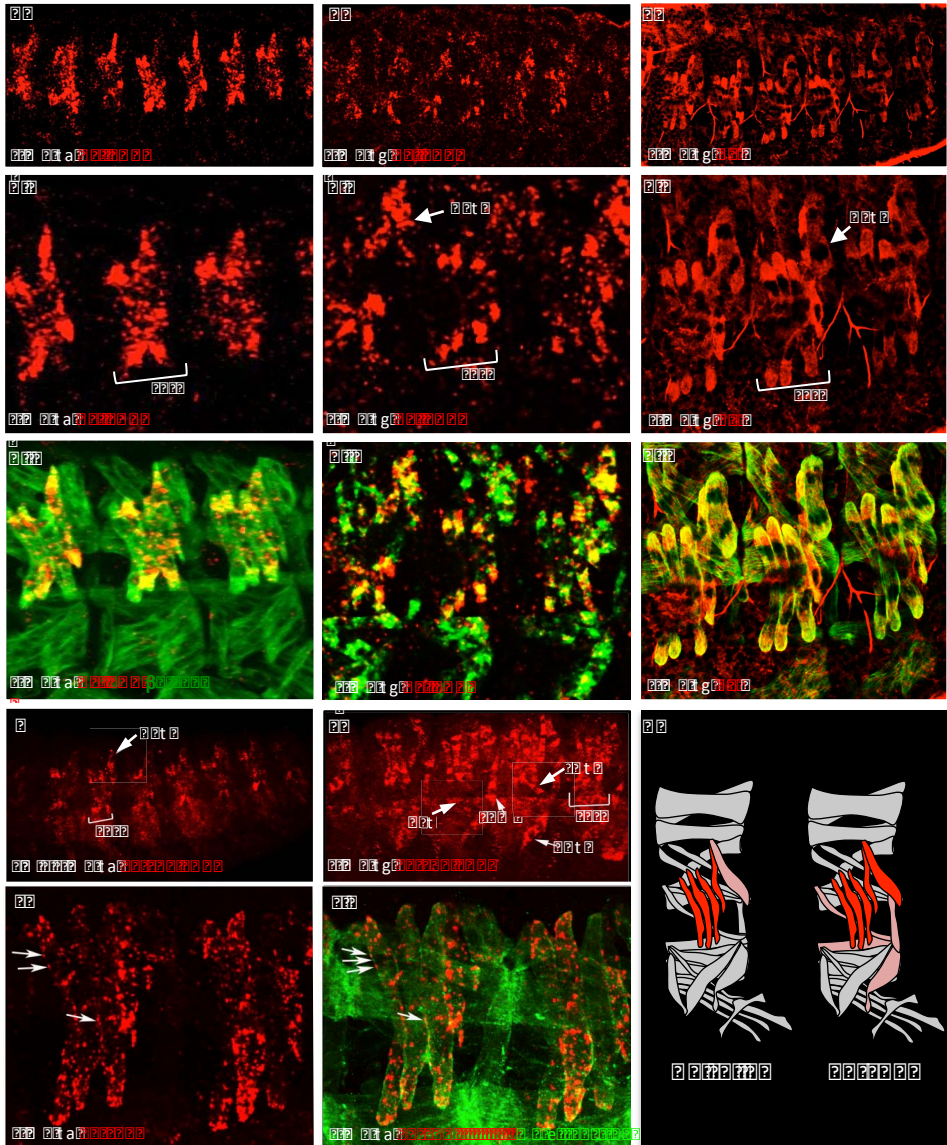
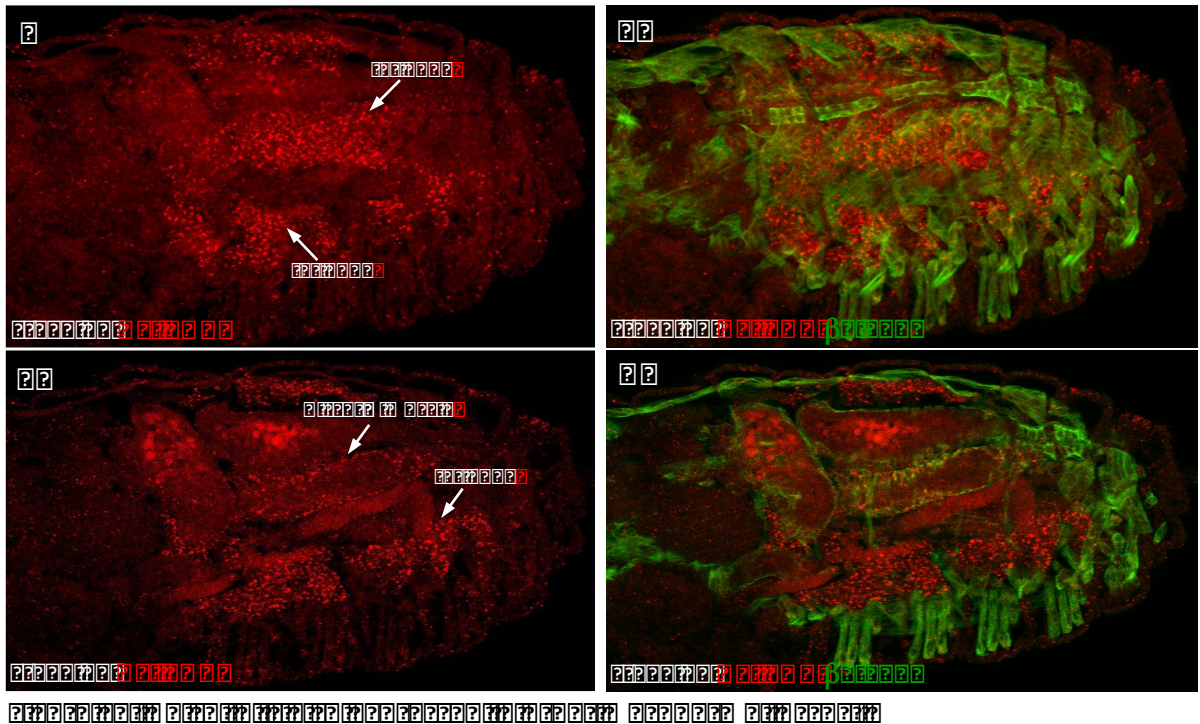


Figure 1. Fluorescence microscopy images of the structure shown in the schematic diagrams. The images show the localization of the red and green signals in the structure. The schematic diagrams illustrate the structure and the localization of the red and green signals. The images are arranged in a grid with some cells containing white arrows and brackets.



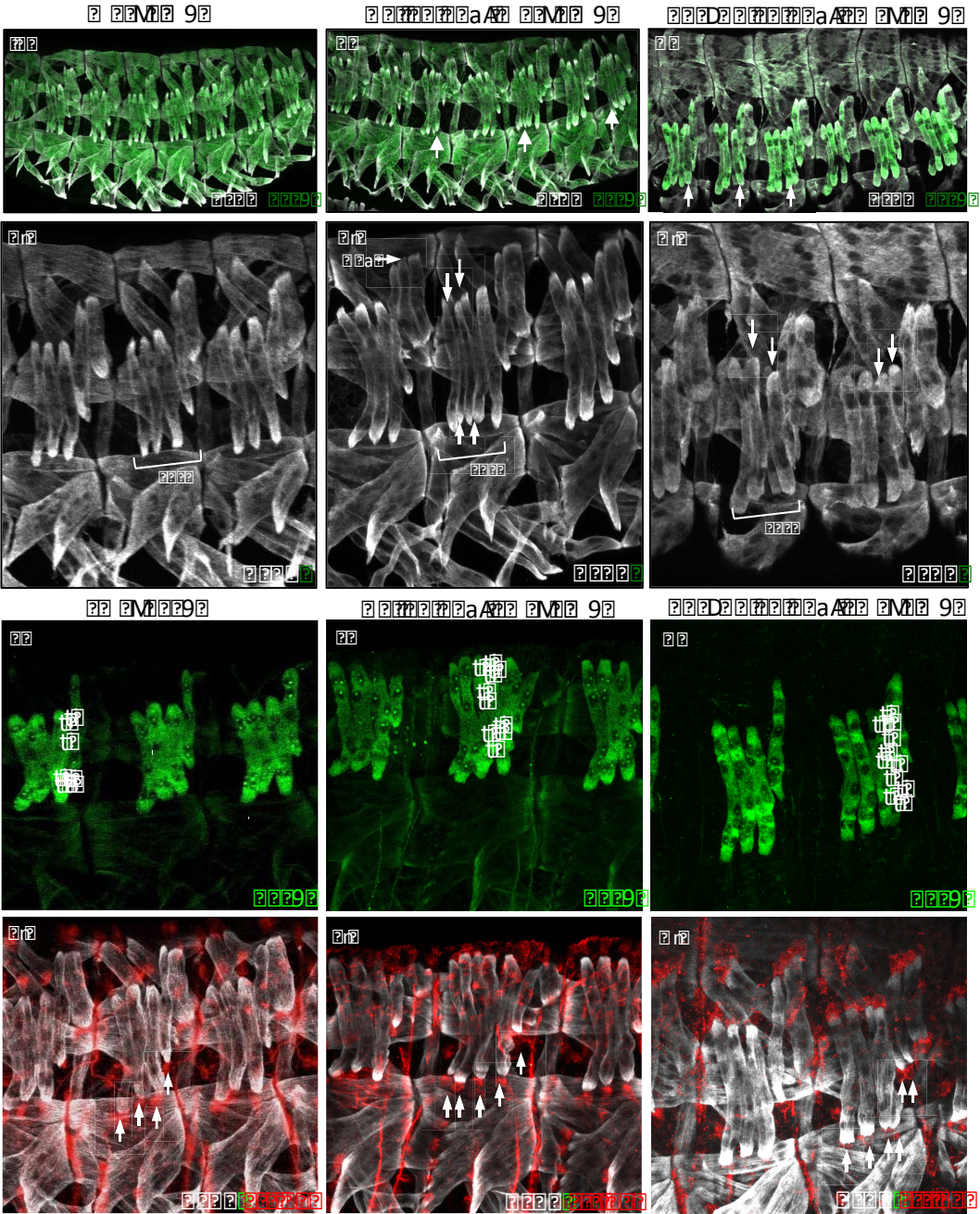
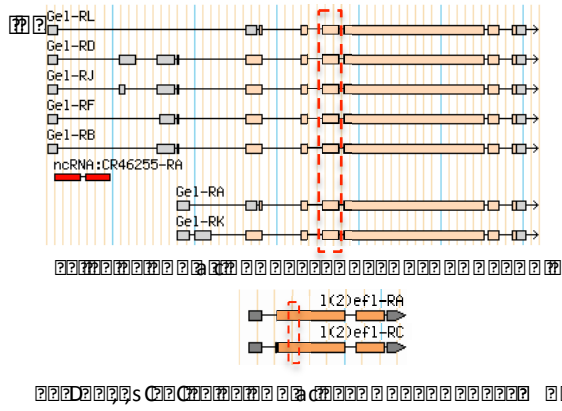
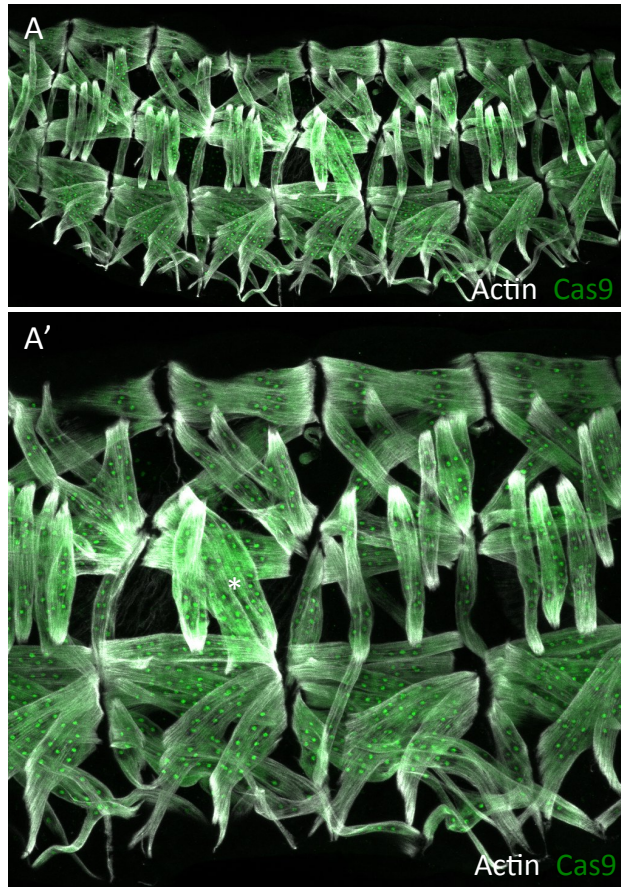


Figure 1. Ge1 gene family structure and expression patterns. (a) Schematic of Ge1 family members and ncRNAs. (b) Whole plant fluorescence images. (c) Leaf epidermal cell fluorescence images. (d) Leaf epidermal cell fluorescence images. Scale bars: (b) 10 cm, (c) 1 mm, (d) 100 μm.

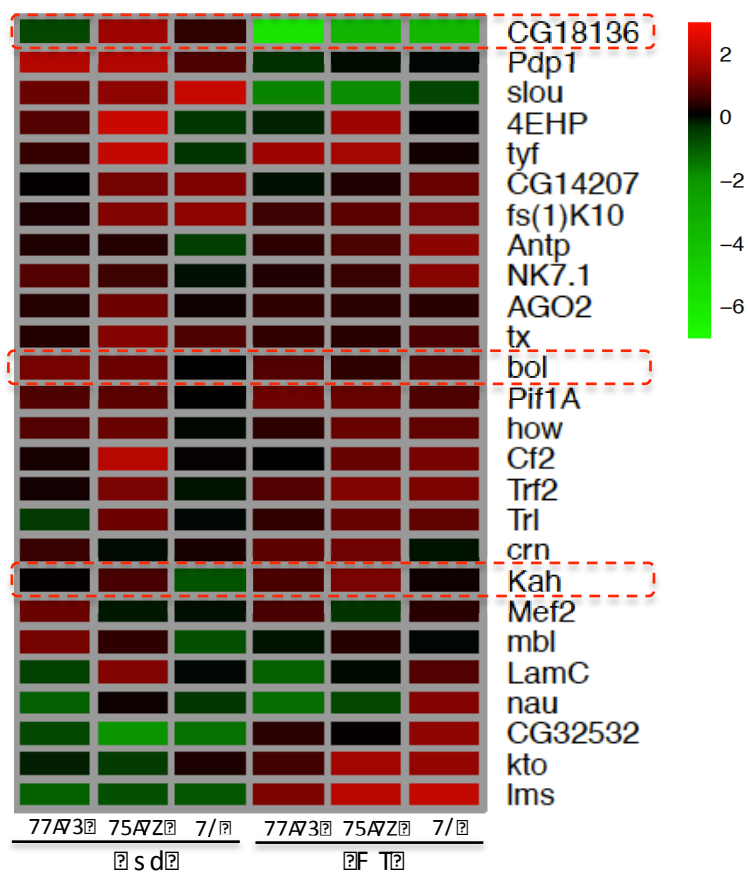


**Figure S5: Conditional CRISPR/Cas9 Knockout of LT muscle identity gene *apterous* leads to severe LT muscle phenotypes**

Genotype	Strain	Condition	Measurement
3	na88N85Z	Hdt	H
ape	na83/8/P3	s F s s é	w
s	na887738/	H	
S3	na88883N/	ayan	A
7P38w	na885785w	eb T p a H p Tbs e D p	A
53Z53	na8823Z53	s F s s é	7
Da	na88885ww	S A	7
STG 2778	na8888N78	P8	H
st	na83/PP67	s F s smi H	
B	na885Z7PP	ayan	3
l ps	na888753P	s	73
F T	na885PZ38	s F s s é	A
	na83/ZPNw	ayan	3
S3	na8877/Z/	A s é	3
B	na8883633		
wg	na883P537	s F s s é	A
e7	na887//6P	P8	
67	na83/787Z	s o A s o	A
s d	na88836P7	s F s s é	7A
7N75/	na885/N56	s F a h a yan	5
S3	na83/7w65	A s é	7
D	na88753/5	7 H H H a yan	A
é	na83/577N	G 2	A
p S	na883/8N5	A	A

Genotype: 77A3 75A3 7/ 77A3 75A3 7/

Condition: s d F T



Genotype: 77A3 75A3 7/ 77A3 75A3 7/

Condition: s d F T

Genotype: uu

Condition: p e e

Genotype: na

Condition: s e e

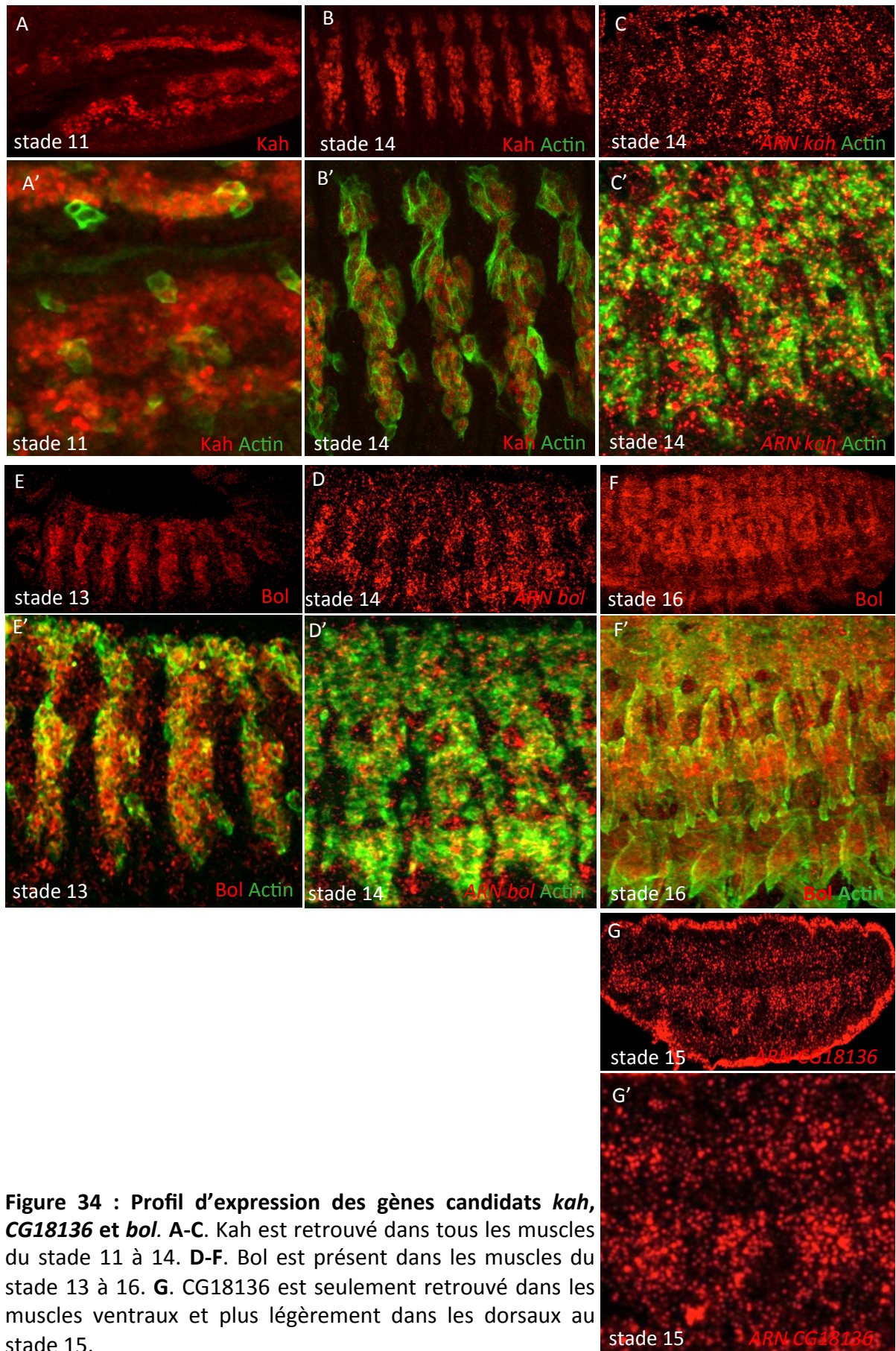
## II. Résultats complémentaires

### 2.1 Identification d'autres acteurs de la diversification musculaire

L'analyse des données issues des puces ADNC pour les 3 populations musculaires aux trois stades de développement a permis de déterminer une liste de gènes enrichis et appauvris pour chacune de ces conditions. A travers ces différentes listes nous avons identifié les gènes communs et spécifiques aux différents types de muscles. Une importante quantité de gènes provenant des populations Slou et Lms est partagée avec la population globale (Duf). Afin de mettre en évidence les gènes « différentiellement exprimés » dans les populations Lms et Slou par rapport à l'ensemble des muscles, les données issues des lignées Slou et Lms aux trois stades de développement ont été comparées avec les données Duf en appliquant un seuil  $>2$  ( $p\text{-value} < 0,05$ ). Pour la population Slou on retrouve 738, 1599 et 838 gènes respectivement enrichis aux stades 11-12 (T1), 13-15 (T2) et 16 (T3) et pour Lms 573, 1044 et 984 gènes. Cette quantité de gènes reste très importante et ne peut donc pas être testée dans sa globalité.

Afin de sélectionner les candidats les plus spécifiques de ces deux ensembles de muscles seul les gènes enrichis pour chaque population à la fois par rapport à l'embryon entier et à la musculature globale (Duf) ont été gardés. De plus les mêmes sondes associées à chaque gènes devaient être valides ( $p\text{-value} < 0,05$ ) dans les deux populations et les trois stades de développement pour être conservées. Au final cette restriction entraîne un haut degré de spécificité et donc un nombre réduit de gènes sélectionnés à 197.

Pour obtenir plus d'informations sur l'ensemble de ces gènes une recherche de GO a été réalisée avec GO Term Finder. Il s'avère que les fonctions biologiques générales associées aux muscles incluant la fusion des myoblastes ou l'organisation des sarcomères sont enrichies. Cependant une catégorie en particulier a retenu notre attention, celle liée à la régulation de l'expression des gènes. En effet à l'intérieur on retrouve 24 gènes codant pour des facteurs de transcription, des protéines de liaison à l'ARN ou encore des modificateurs de la chromatine. Dans l'objectif de découvrir de nouveaux acteurs spécifiant le destin cellulaire de la population Slou et Lms nous nous sommes focalisés sur l'analyse de ces gènes (Tableau 2, Figure 33). Certains sont déjà connus pour être impliqués dans la formation du muscle. Les facteurs musculaires généraux tels que Mef2 et Cf2 (chorion factor 2) sont retrouvés. Mef2 joue un rôle clé dans la spécification et la différenciation des muscles somatiques alors que Cf2 est impliqué dans la régulation de la taille et du nombre



**Figure 34 : Profil d'expression des gènes candidats *kah*, *CG18136* et *bol*.** A-C. *Kah* est retrouvé dans tous les muscles du stade 11 à 14. D-F. *Bol* est présent dans les muscles du stade 13 à 16. G. *CG18136* est seulement retrouvé dans les muscles ventraux et plus légèrement dans les dorsaux au stade 15.

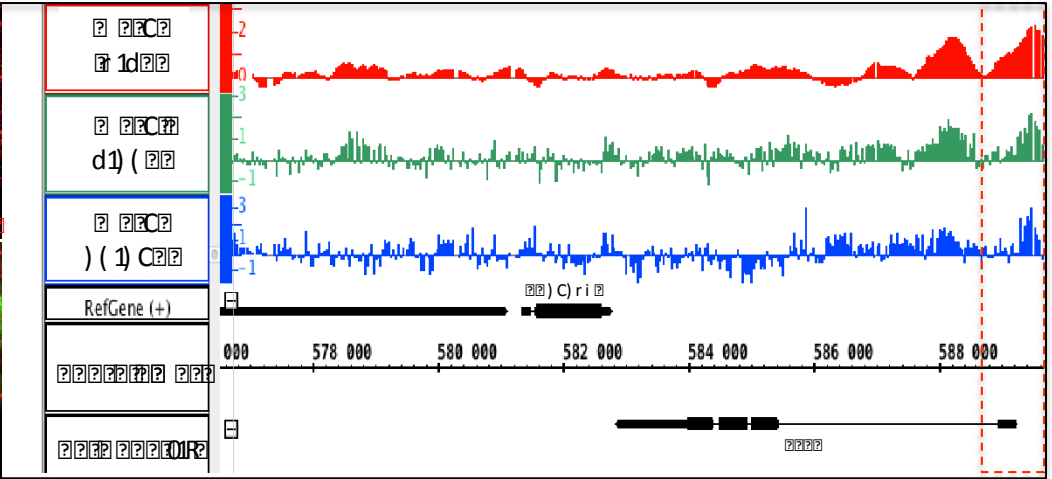
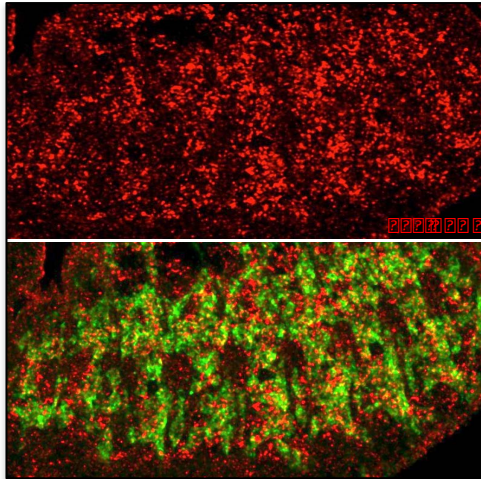
de noyaux présents dans chaque muscles (Arredondo et al., 2017). On retrouve aussi les gènes d'identité *lms* et *slou* précédemment décrits comme étant nécessaires à la spécification et la différenciation de sous populations de muscles. Pour finir des facteurs impliqués dans l'étape d'attachement aux cellules tendineuses comme *How*, *delilah* (*tx*) (Baehrecke, 1997; Nachman et al., 2015) ou dans la mise en place de la jonction neuromusculaire, *Lola*, sont retrouvés (Fukui et al., 2012). Parmi cette liste, seuls 6 gènes n'ont jamais été étudiés pour leur fonction musculaire potentielle. En regardant le profil d'enrichissement de ces derniers dans nos deux populations ainsi que leur conservation chez l'homme nous avons porté notre attention sur trois gènes en particulier qui sont *kahuli* (*kah*), *CG18136* et *boule* (*bol*).

### 2.2 Profil d'expression et régulation des gènes *Kah*, *CG18136* et *Boule*.

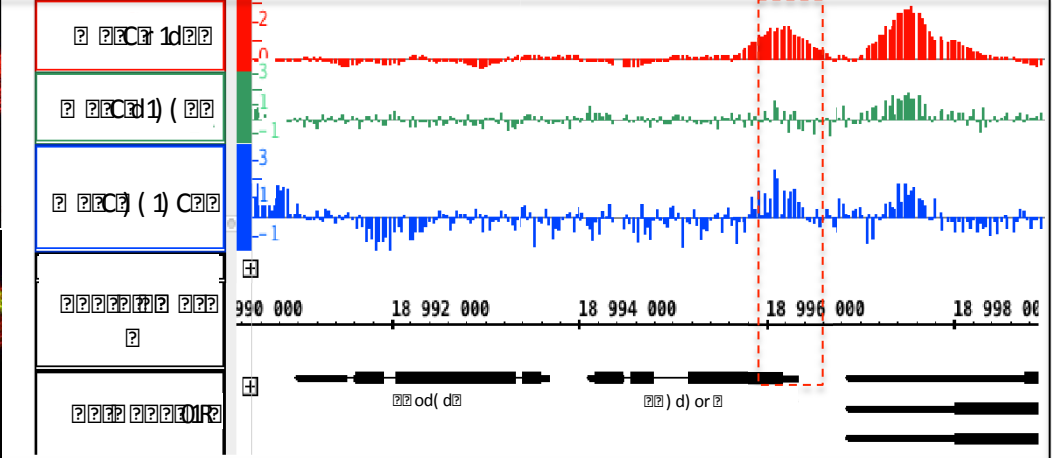
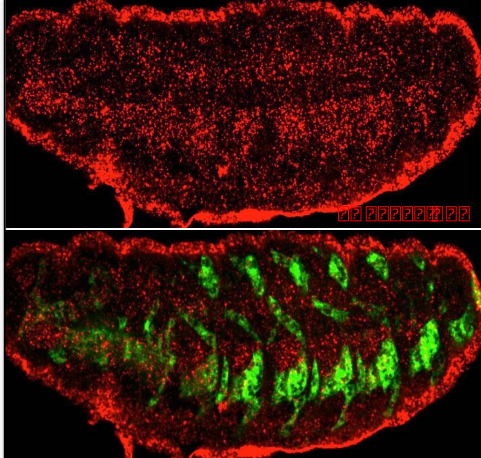
En ce qui concerne ces trois gènes ils codent tous pour des protéines aux rôles différents : *Kah* pour un facteur de transcription à doigt de zinc de la famille Snail/Scratch qui semble légèrement plus enrichi dans la population *Lms* au stade 13-15 (T2). *CG18136* code pour une histone déacétylase seulement enrichie dans les muscles *Slou* au stade 13-15 (T2) et *bol* pour une protéine de liaison à l'ARN retrouvée dans les deux populations (Tableau 2, Figure 33). Afin de déterminer leur implication potentielle dans la formation des muscles embryonnaires j'ai tout d'abord vérifié leur profil d'expression. Les ARN messagers de *Kah*, *CG18136* et *bol* présentent un patron musculaire. En effet *Kah* est exprimé dans le muscle somatique de manière générale du stade 11 au stade 14, *CG18136* est retrouvé principalement dans les muscles ventraux (*Slou* positifs) et dorsaux au stade 15. *Bol* lui est présent dans tous les muscles du stade 13 au stade 16 ainsi que dans le système nerveux périphérique dès le stade 15. Par la suite l'utilisation de la lignée *Bac-Kah-GFP* (Bloomington #64829) et de l'anticorps *Bol* (fournit par l'équipe de Wasserman) a permis de confirmer que leur expression protéique reflétait celle des transcrits. Pour *Kah* on retrouve une expression nucléaire dans les muscles du stade 11 au stade 14 et pour *Bol* une expression cytoplasmique du stade 13 à 16 (Figure 34). Ces résultats confirment donc la présence de ces gènes au sein des différentes populations *Lms* et *Slou*. Par la suite j'ai cherché à savoir comment ces gènes pouvaient être régulés. A l'aide de données issues d'expériences de ChIP (données publiées du laboratoire E. Furlong, EMBL) réalisées sur le facteur de transcription *Mef2* au stade 6-8h, 8-10h et 10-12h (AEL) j'ai pu détecté la présence de sites de liaison de



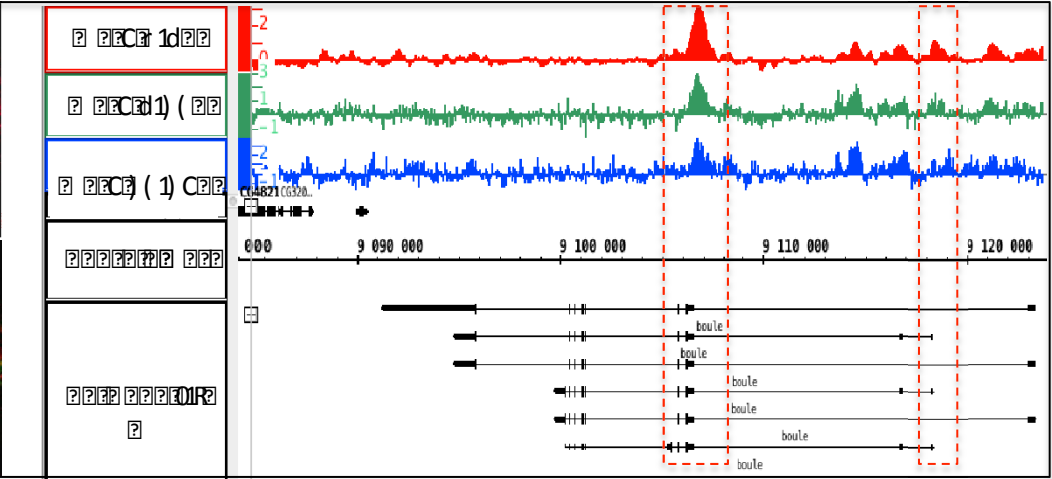
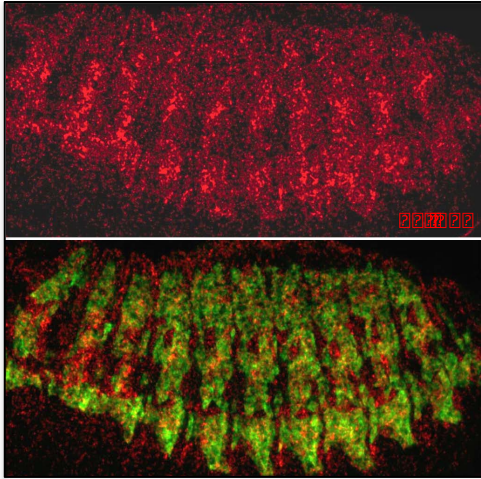
??



??



??



re 1d 1) ( 1) C RefGene (+) 000 578 000 580 000 582 000 584 000 586 000 588 000  
 1d 1) ( 1) C 990 000 18 992 000 18 994 000 18 996 000 18 998 000  
 1d 1) ( 1) C 000 9 090 000 9 100 000 9 110 000 9 120 000  
 boule boule boule boule boule boule

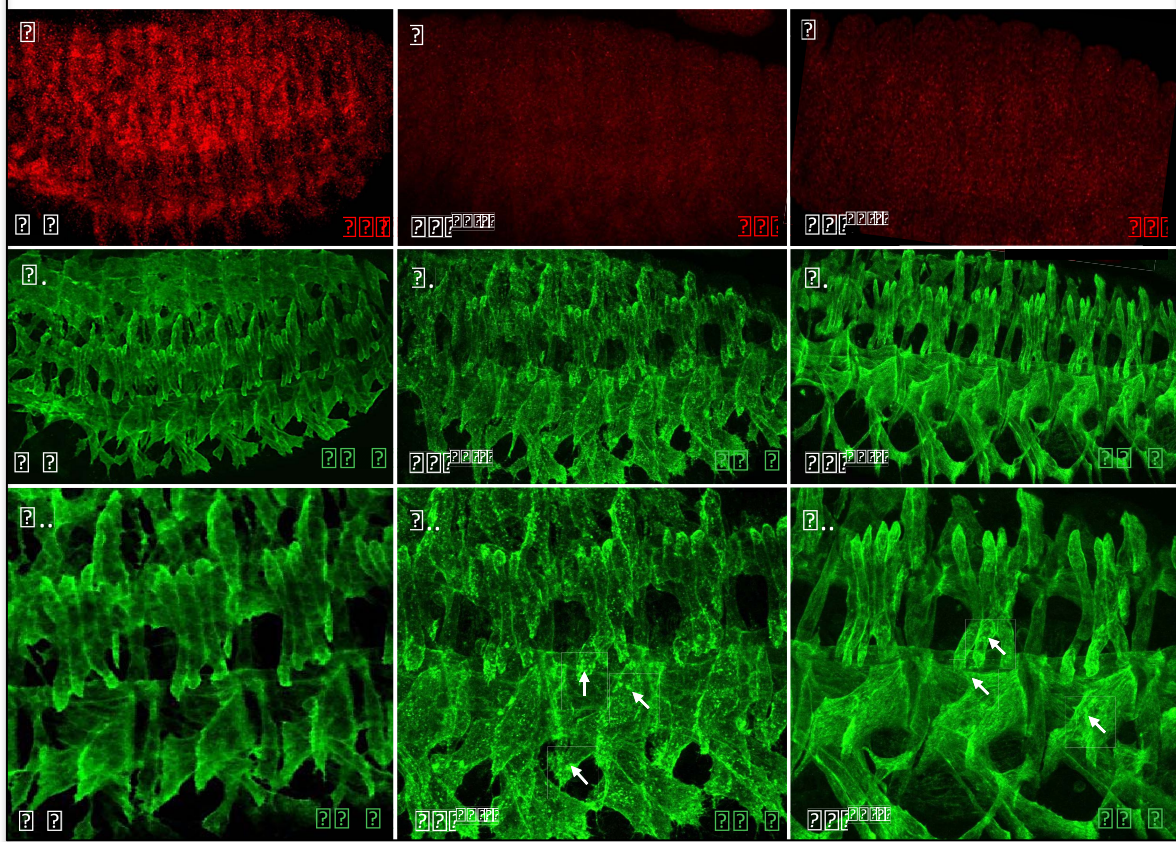
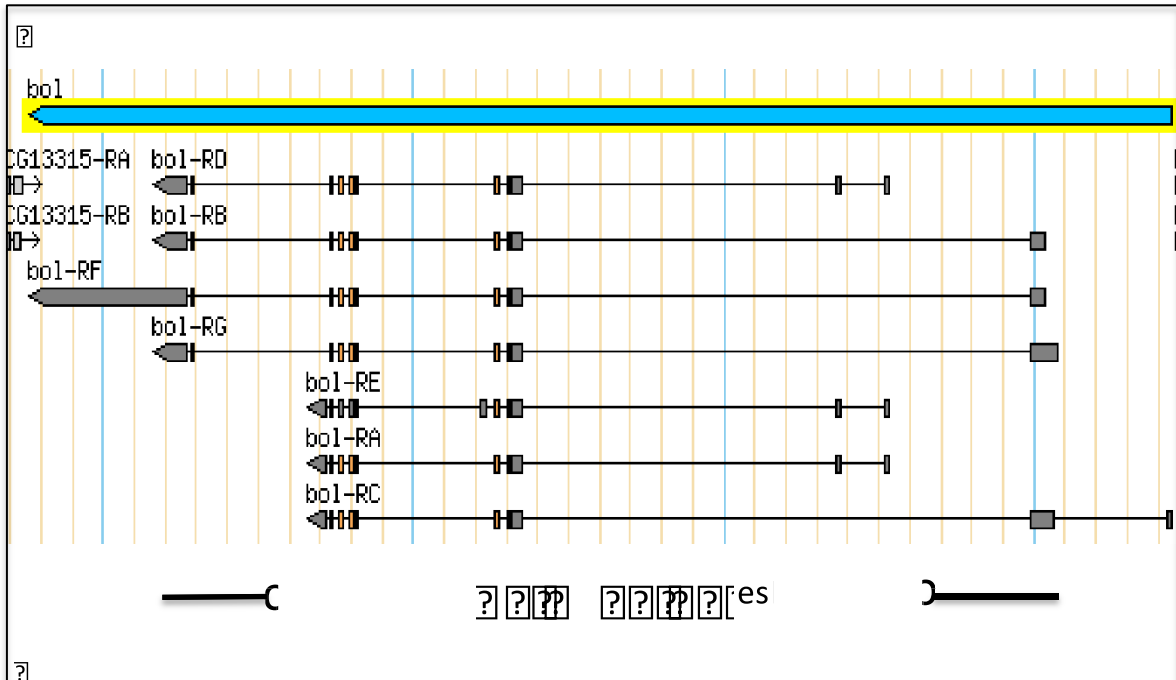
ce facteur à proximité de ces gènes. Pour *Kah* un pic important pour Mef2 (trois stades) est observé au niveau de la région promotrice ainsi que dans le premier intron. Pour le *CG18136* on distingue la fixation de Mef2 aux stades 6-8h, 10-12h AEL au niveau de son promoteur. *Bol* lui présente plusieurs pics, un au niveau du TSS (transcription start site) des transcrits les plus courts et un plus important au niveau des exons 2 ou 3 de tous les transcrits ce qui suggère un potentiel promoteur alternatif (Figure 35). Ces résultats mettent en évidence la régulation potentielle de *Kah*, *CG18136* et *Bol* par le facteur myogénique Mef2 à différents temps de développement ce qui suggère leur implication en aval de Mef2 dans la myogenèse embryonnaire.

### 2.3 Analyses fonctionnelles des gènes candidats par ARN interférent

Dans un premier temps nous avons réalisé des analyses de perte de fonction de ces gènes grâce à différentes lignées d'ARN interférent (RNAi) provenant de VDRC (Vienna Drosophila Resource Centre). L'expression ciblée de ces RNAi dans nos populations musculaires d'intérêt et aussi dans tout le mésoderme somatique a été réalisée grâce aux lignées pilotes d'expression *Lms-Gal4*, *Slou-Gal4* et *Duf-Gal4*. Grâce à des immunomarquages dirigés contre la protéine  $\beta$ 3-tubuline, un constituant majeur des microtubules, nous avons observé la paroi musculaire des embryons dont l'expression est normalement diminuée pour ces différents gènes. Pour chaque candidat deux lignées RNAi ont été testées, malheureusement aucune d'entre elles n'a permis de mettre en évidence des défauts musculaires. Bien qu'aucun phénotype n'a été observé cela ne signifie pas que ces gènes sont dépourvus de fonction au sein du muscle. En effet, le manque de pénétrance des RNAi dans le muscle ainsi que la quantité importante de transcrits à cibler, du fait des propriétés multinuclées des cellules musculaires, font de ce système un outil souvent peu efficace pour abroger correctement l'expression de ces gènes en embryon. Pour contourner cela j'ai recherché dans un second temps les mutants disponibles pour ces gènes. Seul un mutant pour *bol*, *bol<sup>40</sup>* a été publié par l'équipe de Luo en 2008. J'ai donc focalisé mon analyse sur ce gène en particulier.

### 2.4 Caractérisation fonctionnelle du gène *bol*

*Boule* appartient à une famille de protéines de liaison à l'ARN avec une fonction conservée dans la régulation du cycle cellulaire méiotique (Eberhart et al., 1996). *Bol*



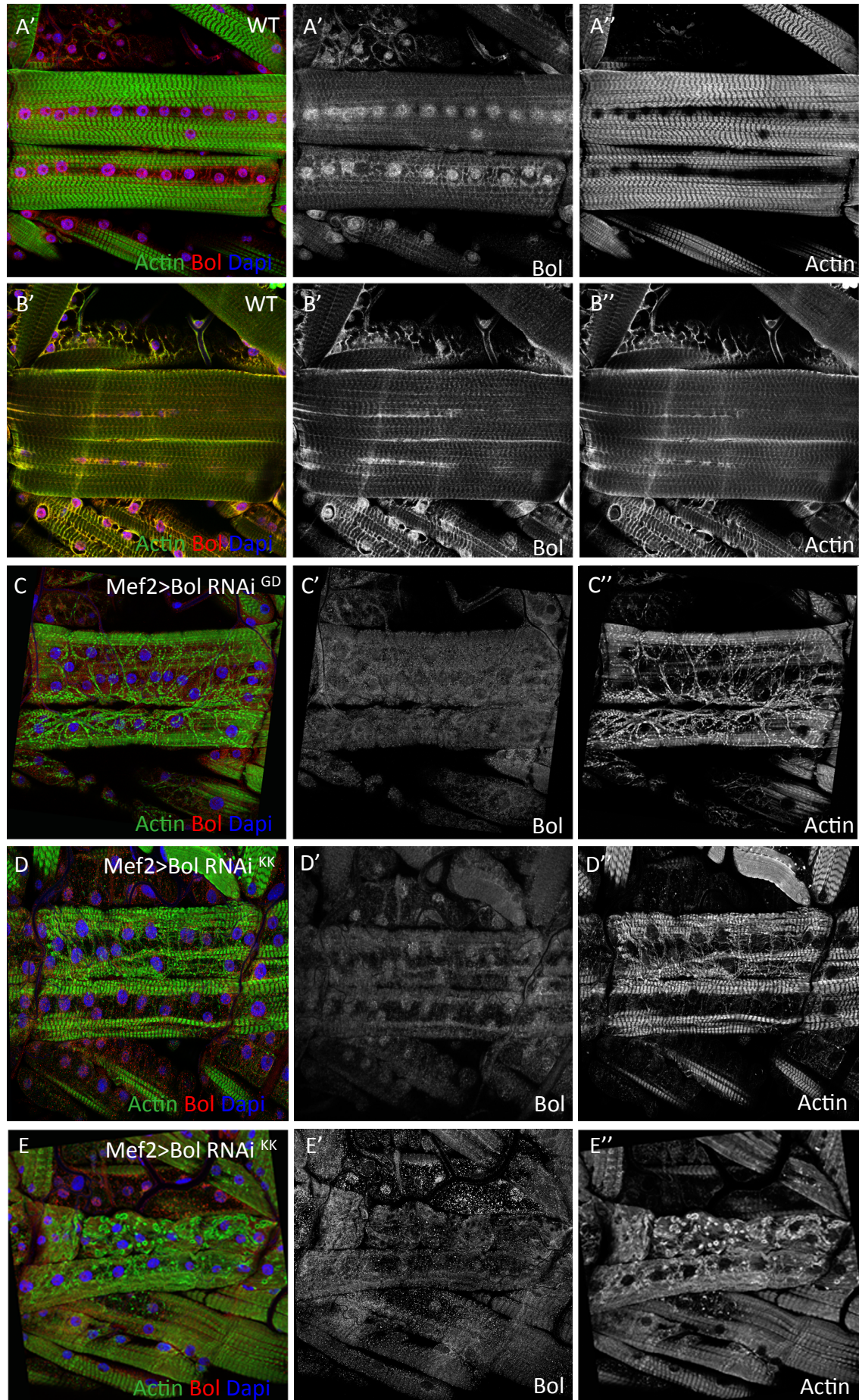
The figure displays a genetic map of the *bol* locus on chromosome CG13315, showing the arrangement of genes *bol-RO*, *bol-RB*, *bol-RF*, *bol-RG*, *bol-RE*, *bol-RA*, and *bol-RC*. Below the map, a centromere and telomere are indicated, along with a scale bar labeled "es".

The fluorescence microscopy images show a dense network of cells or fibers. The top row shows red fluorescence, the middle row shows green fluorescence, and the bottom row shows a merged image of red and green. In the bottom-middle and bottom-right images, white arrows point to specific features within the green network.

contient un domaine de liaison à l'ARN avec deux motifs de reconnaissance de l'ARN (RRM) et un domaine DAZ qui lui permet d'interagir avec d'autres protéines (Collier et al., 2005; Maegawa et al., 2002; Urano et al., 2005). Le mutant nul *bol*<sup>40</sup> a été généré par le système de recombinaison Flippase-FRT et a permis la délétion de la majorité de la séquence codante (Figure 36A). De plus, les mouches *bol*<sup>40</sup> ne sont pas viables à l'état homozygote. Après avoir reçu le mutant du laboratoire de Luo, j'ai pu générer une lignée *bol*<sup>40</sup> avec un chromosome balancier marqué afin de suivre les embryons homozygotes mutants. Des immunomarquages avec les anticorps dirigés contre l'actine et la protéine Bol ont ensuite été réalisés. J'ai tout d'abord vérifié l'absence de la protéine Bol dans les embryons homozygotes mutants. Aucune expression de la protéine n'est détectée dans ces derniers confirmant ainsi que *bol*<sup>40</sup> est bien un mutant nul. Concernant la musculature un phénotype particulier a été observé. En effet des foyers denses d'actine sont retrouvés en grand nombre à l'intérieur et aux extrémités des muscles au stade 15 ce qui n'est pas le cas dans les embryons sauvages. Plus tardivement la structure des muscles est affectée. Par exemple, au niveau des muscles latéraux et ventraux cette accumulation d'actine semble perturber la croissance directionnelle et l'extension des myotubes (figure 36B). Par conséquent, les résultats obtenus avec les mutants *bol* suggèrent que ce dernier ait une fonction en relation avec le cytosquelette d'actine et qu'il soit impliqué dans les dernières étapes de différenciation du muscle. Pour savoir si les phénotypes observés en fin d'embryogenèse sont à l'origine de défauts musculaires plus tardivement nous avons suivi le développement des larves *bol*<sup>40/-</sup>. Celles ci ont des difficultés à se déplacer et meurent donc très précocement au cours du stade larvaire L1, ce qui signifie qu'entre le stade 16 embryonnaire et le premier stade larvaire les défauts musculaires s'accroissent. Cette létalité précoce au stade larvaire rend impossible la dissection et donc la visualisation des muscles.

### 2.5 Analyses fonctionnelles du gène *bol* en larve

Afin de mieux caractériser la fonction de Bol au cours du développement musculaire larvaire nous sommes revenus à l'utilisation du système RNAi qui est plus efficace en larve. Tout d'abord nous avons confirmé l'expression du gène *bol* en larve L3. Celui-ci est retrouvé exprimé de manière granuleuse principalement autour des noyaux et bien plus faiblement au niveau de la bande Z des sarcomères. De plus une expression nucléaire est détectée cependant celle-ci semble complètement exclue du nucléole (Figure 37). Par la suite nous

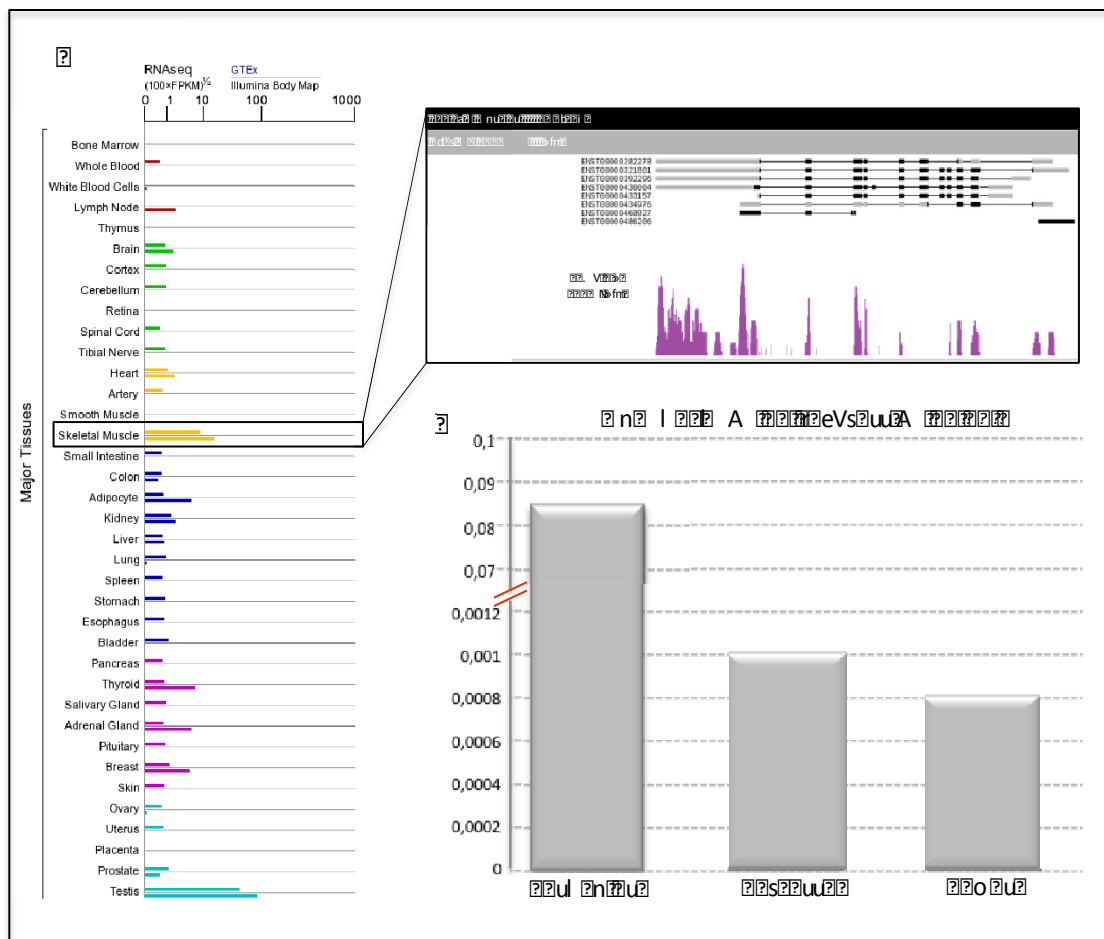


**Figure 37: La perte de Bol induit une désorganisation des fibres musculaires chez la larve.** Immunomarquage de larves L3 sauvages (A-B) et mutantes pour Bol (C-F) avec l'anticorps Bol (A'-E') et Actine (A''-E''). Une perte d'expression de la protéine Bol est observée avec les deux lignées RNAi (C'-E') entraînant un désassemblage des filaments d'actine, une atrophie des fibres (C'',D'') et parfois des agrégats d'actine à la surface des muscles (E'').

avons utilisé la lignée pilote Mef2-Gal4 pour exprimer les RNAi contre *bol* dans les muscles somatiques durant tout le stade larvaire. Après dissection, des immunomarquages ont été réalisés avec les anticorps Bol et Actine pour visualiser respectivement la perte d'expression de *bol* et la structure des muscles. Bien que l'ensemble de la musculature larvaire soit ciblé par le pilote Mef2-Gal4 les analyses phénotypiques ont été réalisées sur les muscles VL, lesquels sont représentatifs de la population musculaire globale et en même temps les plus faciles à imager. Dans les larves RNAi pour *Bol* (Mef2>*Bol*<sup>GD</sup> et Mef2>*Bol*<sup>KK</sup>) on note une diminution de l'expression de ce dernier ce qui suggère l'efficacité des lignées pour cibler les transcrits *bol*. En ce qui concerne la musculature, des défauts majeurs de la structure des muscles ont été observés. Les larves affectées présentent une perturbation sévère des filaments d'actine, une atrophie et un déchirement sporadique des fibres musculaires ainsi qu'une délocalisation des noyaux. Les mêmes phénotypes ont été retrouvés dans plusieurs muscles et dans les deux RNAi. Cependant, avec la lignée KK différents degrés de phénotypes ont été identifiés. En effet dans la majorité des cas les défauts sont similaires à ceux de la lignée GD mais dans quelques larves on retrouve une atrophie des muscles plus importante ainsi que de nombreux agrégats d'actine (Figure 37).

Tous ces phénotypes sont principalement retrouvés à la surface des muscles au niveau des filaments d'actine qui entourent normalement les noyaux et qui constituent les myofibrilles externes. Cette localisation des défauts corrèle parfaitement à l'expression endogène de *bol*. Cela suggère que ce dernier pourrait jouer un rôle dans la régulation de certaines protéines impliquées dans l'architecture et la stabilisation des filaments d'actine à la surface des fibres musculaires.

A ce jour *bol*, a surtout été décrit pour son rôle au sein des cellules germinales chez les vertébrés. Étant donné son expression musculaire chez la drosophile il était important de vérifier la présence de ce dernier au niveau du muscle squelettique chez les vertébrés. Grâce à la base de données « the human protein atlas » qui regroupe un grand nombre d'informations concernant l'expression des gènes humains j'ai pu noter une faible expression de *bol* dans le muscle squelettique. Cette donnée quantitative d'expression provient d'un séquençage ARN réalisé sur un mélange de muscles adultes (Figure 38A). Dans l'objectif d'étudier le rôle de *bol* dans le muscle chez les vertébrés j'ai recherché son expression dans l'embryon de souris. Pour cela des RT-qPCR sur les testicules, la carcasse et les muscles de la



1. The figure displays gene expression data across various tissues. The main horizontal bar chart shows RNA-seq expression (100x FPKM) for 32 tissues. The x-axis is logarithmic, ranging from 0 to 1000. The y-axis lists tissues: Bone Marrow, White Blood Cells, Lymph Node, Thymus, Brain, Cortex, Cerebellum, Retina, Spinal Cord, Tibial Nerve, Heart, Artery, Smooth Muscle, Skeletal Muscle, Small Intestine, Colon, Adipocyte, Kidney, Liver, Lung, Spleen, Stomach, Esophagus, Bladder, Pancreas, Thyroid, Salivary Gland, Adrenal Gland, Pituitary, Breast, Skin, Ovary, Uterus, Placenta, Prostate, and Testis. An inset shows a detailed view of a gene's expression profile across tissues, with a y-axis labeled 'V' and 'N'. A bar chart on the right shows p-values for three comparisons: 0,0012, 0,001, and 0,0008.

patte arrière d'embryons de souris à E11.5 ont été réalisées. La méthode de  $\Delta\Delta\text{Ct}$  est utilisée pour l'analyse des données avec *Rps17* comme gène de référence. Une forte expression de *bol* est retrouvée comme attendu dans les testicules et une très légère expression est observée dans la carcasse et les pattes (Figure 38B). La détection de l'expression de *bol* dans le muscle suggère que ce dernier pourrait être impliqué dans la myogenèse embryonnaires chez la souris.





? stress r n?



## A- La méthode TRAP : une technique efficace pour l'identification du transcriptome muscle spécifique

Pour ce projet nous avons utilisé la méthode TRAP pour identifier le transcriptome de sous population musculaire dans l'embryon de drosophile. Après optimisation, la mise en place de cette dernière nous a permis de récupérer des ARN messagers engagés en traduction spécifiques des populations ciblées avec une haute sensibilité. Par conséquent, le choix de la méthode TRAP pour réaliser ce projet fut concluant. Nous avons ainsi pu identifier pour la première fois de façon globale les gènes participant au contrôle de l'acquisition des caractéristiques individuelles de sous populations musculaires au cours de différentes étapes de l'embryogenèse.

Bien que la technique ait été optimisée et s'avère particulièrement efficace et sensible pour la récolte des ARNs messagers engagés en traduction, certains points peuvent encore être améliorés. En effet un des facteurs limitant est la quantité de matériel biologique de départ. A ce jour pour isoler assez d'ARNms en vue de les analyser, une récolte conséquente d'embryons stadés est nécessaire. Cela implique une logistique d'amplification des lignées et de récolte des embryons assez lourde qui demande beaucoup de temps. Les méthodes d'amplification d'ARN à partir d'une faible quantité (actuellement 15 ng, Kit Ovation de chez Nugen) ainsi que l'amélioration des techniques de séquençage laissent entrevoir une possibilité de réduction du matériel de départ dans un avenir proche. Ces avancées nous permettraient aussi d'utiliser le TRAP sur des tissus encore moins représentés et de réduire les fenêtres de temps de collection des embryons (ici 3h), apportant ainsi un degré de précision supplémentaire.

Un autre point d'amélioration serait l'étape d'immunopurification. En effet dans un souci de spécificité maximale des ARNs isolés il serait possible de développer une stratégie plus avancée. Pour cela une double purification (Tandem affinity purification) pourrait être réalisée grâce au système SBP-TEV (Streptavidin binding peptide- tobacco etch virus) (Bürckstümmer et al., 2006). Cette technique repose sur l'expression d'une protéine d'intérêt, dans notre cas Rpl10a, fusionnée à deux tags (SBP-TEV + GFP) permettant ainsi deux purifications successives. Le site de la protéase TEV permet de libérer la protéine



fusionnée après la première isolation et de procéder à une seconde isolation avec anticorps anti-GFP. L'ensemble des améliorations apportées permettra de démocratiser cette approche à de nombreux systèmes.

L'utilisation de puces ADNc pour analyser les échantillons a permis de mettre en évidence de nombreux transcrits différentiellement traduits au cours du temps dans ces deux populations. Cependant ce type d'analyse présente quelques limites. En effet la couverture du génome n'étant pas parfaite du au nombre et à la localisation des sondes cela engendre parfois quelques biais d'interprétation. De plus, celle-ci ne permet pas de distinguer les différents types d'isoformes exprimés. Sachant qu'un grand nombre de transcrits de gènes de la structure musculaire (protéines sarcomériques) sont connus pour être différentiellement épissés, il serait intéressant de connaître les isoformes des gènes enrichis dans les populations ciblées. A ce jour les performances accrues ainsi que la tendance à la réduction des coûts des techniques de séquençage ARN nous permettrait d'obtenir ce type d'informations (isoformes spécifiques). Dans cette optique un premier séquençage a été réalisé sur les ARN isolés de la population musculaire globale (Duf positive) au stade 16 (T3). Il s'avère que 60% des gènes enrichis dans nos données (puces à ADNc) sont retrouvés dans le RNA-seq (à fold change similaire >2). Cela semble très correct du fait de l'éloignement technologique de ces deux méthodes d'analyses. De plus, l'ensemble des gènes testés dans cette étude sont présents au sein des données du RNA-seq ce qui renforce nos résultats. Au final l'utilisation de l'approche génomique cellule spécifique TRAP suivie d'analyses par puces ADNc a été un choix cohérent et efficace qui m'a permis d'identifier un grand nombre de gènes jamais caractérisés auparavant pour leurs fonctions musculaires.

Afin de compléter les informations obtenues par le TRAP, nous pourrions combiné cette dernière à la méthode alternative du « Ribosome Footprint » qui permet une mesure globale de la vitesse de traduction *in vivo*. Cette approche est basée sur le fait que les ribosomes protègent environ 30 nucléotides des ARNs sur lesquelles ils sont fixés, de la dégradation par les nucléases (Ingolia et al., 2009). La séquence de ces fragments protégés appelée « footprint » ou « empreinte des ribosomes » fournit un enregistrement précis de la position des ribosomes à un temps spécifique. En mesurant la densité des fragments protégés d'un transcrit donné, on obtient approximativement son ratio de synthèse



protéique. De plus, cela permet de déterminer où la traduction débute, où elle se termine ainsi que la présence d'autres cadres de lecture au sein d'une même séquence. Cela a mené à l'identification de nouvelles protéines ainsi qu'à de nouvelles isoformes de protéines (Andreev et al., 2015; Ingolia et al., 2011). Dans notre cas, une fois que les complexes polysomes-ARNs messagers de nos populations d'intérêts seraient isolés sur billes magnétiques via le TRAP, les fragments d'ARNs non protégés par les ribosomes seraient digérés (nucléase) puis les ARNs restant purifiés et séquencés. Les données de séquençage seraient ensuite comparées à celles du TRAP. Cela permettraient de comprendre comment les transcrits enrichis dans ces populations sont contrôlés au niveau de la traduction ainsi que de découvrir potentiellement de nouvelles protéines musculaires.

Pour finir, comprendre comment ces gènes sont régulés au niveau transcriptionnel et post-transcriptionnel reste un enjeu majeur. Dans le but de répondre à cette question nous avons décidé de mettre en place au laboratoire la méthode INTACT qui a été développée tout d'abord chez *Arabidopsis* (Deal and Henikoff, 2010) puis adaptée à *C.elegans* et à la drosophile (Henry et al., 2012; Steiner et al., 2012). Cette approche permet d'isoler les noyaux de manière cellule spécifique. Elle est basée sur l'expression d'un transgène qui va permettre la synthèse d'une protéine étiquetée de la famille SUN qui va s'insérer uniquement dans les noyaux d'une population cellulaire d'intérêt. Son utilisation permet d'isoler les noyaux de façon cellule spécifique afin d'obtenir l'ensemble du transcriptome nucléaire et d'avoir un accès à la structure de la chromatine. L'avantage de la méthode INTACT par rapport au FANS (Fluorescent activated nuclei sorting) c'est qu'elle est aussi rapide, que la quantité de noyaux récoltés à partir du même matériel biologique est plus importante et qu'elle est accessible sans aucune machine et sans stress supplémentaire inhérent au passage par le cytomètre en flux. Une autre différence majeure est que le protocole de FANS requiert l'utilisation de formaldéhyde pour la fixation des tissus. Au contraire l'INTACT peut être réalisée en condition native ou de « cross link » ce qui permet d'adapter le protocole en fonction du type d'échantillon à analyser (ARN ou protéine de liaison à la chromatine).

Dans un premier temps nous avons donc généré grâce au système GAL4/UAS la lignée de drosophile exprimant la protéine de la membrane nucléaire UNC-84 fusionnée à la GFP (Henry et al., 2012) dans les muscles Lms positifs (Figure 39A). Une fois la lignée établie



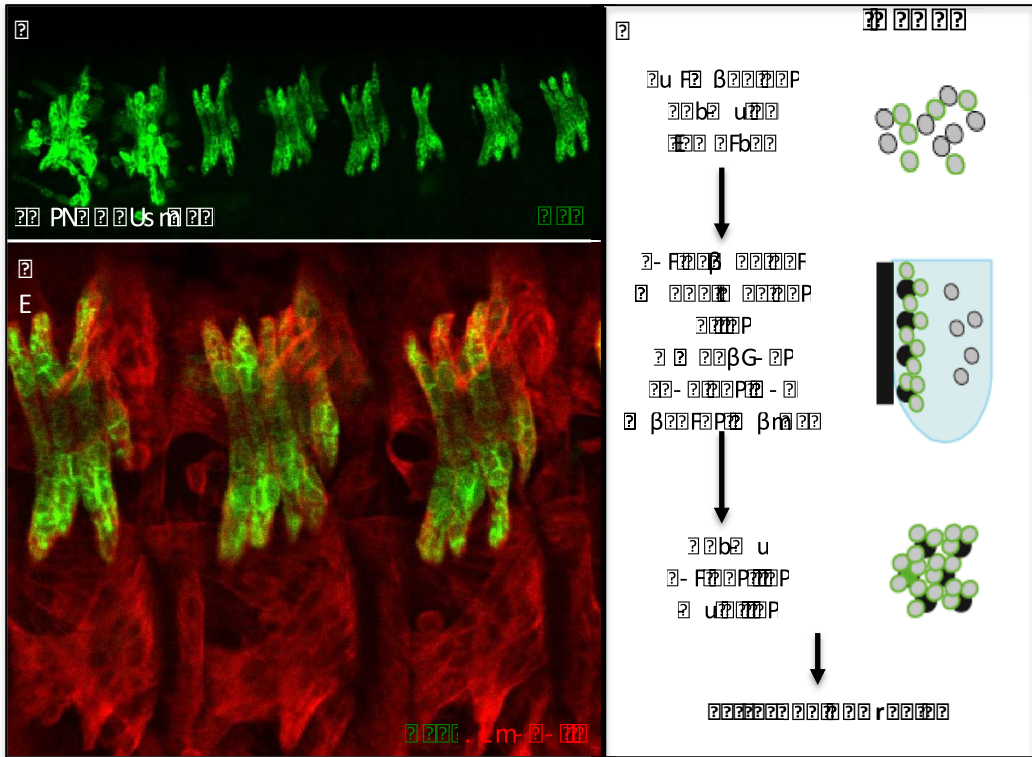


Figure 1. Fluorescence microscopy and schematic diagram of the cell cycle. Panel D shows metaphase chromosomes (green) with labels PN and Us. Panel E shows telophase and cytokinesis (green chromosomes, red cytoplasm) with label Lm. The schematic diagram illustrates the sequence of events: metaphase (chromosomes aligned at the metaphase plate), anaphase (sister chromatids separating), telophase (formation of two new nuclei), and cytokinesis (formation of the cell plate). Labels in the schematic include  $\alpha$ u, F $\beta$ ,  $\beta$ ,  $\alpha$ b $\beta$ , u,  $\alpha$ F $\beta$ ,  $\beta$ -F $\beta$ ,  $\beta$ ,  $\alpha$ b $\beta$ ,  $\beta$ G,  $\beta$ ,  $\alpha$ F $\beta$ ,  $\beta$ ra,  $\alpha$ b $\beta$ , u,  $\beta$ -F $\beta$ , u, and r.

et les embryons récoltés, la mise au point du protocole a ensuite été réalisée par un post doctorant du laboratoire (Thomas Moreau) (Figure 39B). Grâce à lui nous avons isolé dans un premier temps les ARNs des noyaux de la population Lms aux stades 13-15 et 16 (T2, T3). Le séquençage ARN de ces échantillons est actuellement en cours et nous permettra de faire le lien entre le transcriptome nucléaire et les ARN en cours de traduction dans cette population. On pourra alors identifier par exemple si un type de transcrit est préférentiellement traduit où si certains sont soumis à une régulation post-transcriptionnelle. Dans un second temps nous pourrions aussi réaliser des expériences de CHIP sur la chromatine afin d'identifier les cibles directes des facteurs de transcription présents dans cette population tels que Lms, Ap ou de nouveaux facteurs de transcription révélés par nos approches TRAP.

L'ensemble de ces méthodes appliquées à l'embryon de drosophile et plus particulièrement au cours du développement musculaire mettrait en perspective toute la dynamique de l'expression des gènes de manière cellule spécifique et donnerait une meilleur visibilité sur les acteurs impliqués dans les différents processus de l'acquisition de l'identité musculaire.

## **B- Identification de nouveaux acteurs potentiels du code d'identité et de l'acquisition des caractéristiques finales des sous populations musculaires Lms et Slou.**

### **I. De nouveaux acteurs potentiels du code d'identité des muscles Slou positifs.**

L'application de la méthode TRAP suivie d'analyses translatomiques à l'aide de puces ADNc ont permis de mettre en évidence un grand nombre de transcrits différenciellement enrichis entre la population Lms et Slou au cours du temps. Dans le but d'identifier de nouveaux acteurs du contrôle de l'acquisition de l'identité musculaire nous avons dans un premier temps réalisé une comparaison spatiale des données. Nous nous sommes focalisés plus particulièrement sur des candidats ayant une fonction dans la régulation de l'expression des gènes. Cette comparaison a révélé un ensemble de gènes codant pour des facteurs de



transcription, des protéines de liaison à l'ARN ou encore des modificateurs de la structure chromatinienne comme étant préférentiellement enrichis dans la population Slou ou Lms. Etant donné l'hétérogénéité des muscles Slou et donc la complexité du code d'identité nécessaire à la diversification de cette population nous avons préférentiellement étudiés les facteurs enrichis dans ce sous ensemble musculaire.

La recherche du profil d'expression des gènes spécifiquement enrichis dans cette population a permis de mettre en évidence 5 gènes avec un patron musculaire restreint en plus des gènes déjà connus pour être exprimé dans ces muscles tels que Slou lui même et Org-1 (Knirr et al., 1999; Schaub et al., 2012). Parmi les gènes candidats on retrouve *EndoGI*, *CG10565*, *CG10654*, *dSix4* et *Ptx1*. Ces derniers présentent une expression spécifique ou plus importante dans l'ensemble ou une partie des muscles ventraux Slou positifs. Concernant ces gènes, seul *EndoGI*, *Ptx1* et *dSix4* ont été décrit dans la littérature. *EndoGI*, un inhibiteur de l'endonucléase G, est un régulateur de la voie de signalisation Notch. En effet l'expression ectopique de ce dernier induit une activation de la voie Notch dans plusieurs tissus (O'Keefe et al., 2011). Cependant la perte de *EndoGI* ne semble pas affecter tous les processus développementaux dépendant de Notch. Malgré tout dans ces mutants certains neurones échouent à trouver leurs cibles synaptiques au niveau du muscle comme cela est retrouvé dans les mutants conditionnels pour Notch (Crown et al., 2003). *EndoGI* semble donc essentiel à la guidance axonale des motoneurones, probablement en modulant l'activité transcriptionnelle de cibles de Notch. En ce qui concerne *Ptx1*, qui code pour un facteur de transcription à homéodomaine, son expression a été caractérisée dans l'embryon de drosophile (Vorbrüggen et al., 1997). Celui ci est retrouvé dans une portion postérieure de l'intestin, dans les tubules de Malpighi, dans les muscles ventraux ainsi que dans les cellules du système nerveux. Dans cette étude, son rôle au sein du muscle n'a pas été caractérisé de manière détaillée. Pour *dSix4*, l'analyse de sa fonction, focalisée sur son expression très précoce au sein du mésoderme somatique au stade 10, n'a pas permis d'identifier son patron d'expression plus tardif ainsi que son rôle dans le code d'identité musculaire (Clark et al., 2006).

De manière intéressante, tous nos gènes candidats sont conservés chez l'homme. Les orthologues de *EndoGI*, *CG10654*, *CG10565*, *Ptx1* et *dSix4* sont respectivement *XRCC5*, *ZNF324B*, *DNAJC2*, *PITX2* et *SIX4*. A ce jour seul les orthologues de *Ptx1*, *PITX2* et *dSix4*, *SIX4* sont bien connus pour avoir des rôles clés dans la myogenèse chez les vertèbres. En effet



comme présenté en introduction *Pitx2* est exprimé dans le mésoderme cranial où il est nécessaire pour la spécification des muscles extraoculaires (Gage et al., 1999; Zacharias et al., 2011). Il est aussi requis pour la différenciation des cellules du premier arc branchial et régule MyoD dans les cellules progénitrices des membres (L'Honore et al., 2010; Shih et al., 2007) (Voir paragraphe *Pitx2*). Concernant *SIX4*, son expression ainsi que sa fonction dans la spécification des muscles hypaxiaux a bien été caractérisée. En effet ce dernier est impliqué directement dans la régulation des facteurs myogéniques Pax3, Myf5 et MyoD au cours de la myogenèse du tronc et des membres. (Giordani et al., 2007; Relaix et al., 2013; Tapscott, 2005) (Voir paragraphe *Six*).

La présence de tous ces facteurs au sein de la population Slou suggère leur implication dans le code d'identité nécessaire à l'acquisition des propriétés individuelles de ces muscles. Il serait donc intéressant de connaître la hiérarchie génétique qui se met en place dans cette population, comme cela est retrouvé chez les vertébrés au cours de la spécification et la différenciation des différents territoires musculaires (extraoculaires, branchiomériques, épaxiaux, hypaxiaux). Pour cela les perspectives à court terme sont de réaliser des analyses de perte et gain de fonction pour déterminer la contribution de chacun de ces facteurs dans la cascade myogénique. De plus une recherche *in silico* de motifs de liaison de certains facteurs de transcription tels que Mef2, Slou et Org-1 permettrait d'obtenir une indication sur la hiérarchie génique potentielle de ces facteurs au sein de la population Slou.

## II. De nouveaux gènes réalisateurs du code d'identité des muscles Lms positifs, *gel* et *dCryAB*

Dans un second temps pour identifier de nouveaux gènes qui participent aux processus clés de l'acquisition de l'identité musculaire une comparaison temporelle a été réalisée. Celle ci a permis de mettre en évidence des ensembles de gènes associés à une même fonction biologique mais présentant des signatures de traduction différentes au cours du temps entre la population Slou et Lms. A travers cette étude nous avons sélectionné un « cluster » qui contient des gènes ayant une fonction liée à la régulation du cytosquelette d'actine. Après avoir analysé leurs profils d'expression, deux gènes en particulier ont retenu notre attention, *gelsolin* et *dCryAB*. Ces deux protéines présentent un profil restreint et



notamment dans les muscles LTs correspondant à la population Lms. La perte de fonction respective de *gelsolin* et *dCryAB* dans le muscle (mutant conditionnel CRISPR/CAS9) montre des phénotypes semblables. En effet dans les deux cas nous avons observé un nombre de noyaux plus important souvent associé à une séparation des membranes à l'extrémité des muscles LTs et l'attachement à une cellule tendon supplémentaire. Cela suggère que ces gènes, en régulant le cytosquelette d'actine, soient impliqués dans le programme de fusion et d'attachement nécessaire à l'acquisition des propriétés finales des muscles LTs.

L'actine est le composant principale des filaments fins et la protéine la plus abondante dans le muscle strié squelettique. De plus elle est très bien conservée à travers le règne animal (Chhabra and Higgs, 2007). Le cytosquelette d'actine est un réseau de microfilaments dynamiques qui est essentiel pour un nombre important de fonctions cellulaires impliquant la motilité, le réarrangement structurelle des cellules, la migration, la morphogenèse (Jiang et al., 2009; Ren et al., 2009), l'adhésion cellule-cellule ou cellule matrice ainsi que l'apoptose (Papakonstanti and Stournaras, 2008). Pour remplir ces divers rôles, le cytosquelette est régulé par différents types de protéines de liaison à l'actine telles que les protéines de coiffage, de fragmentation (dépolymérisation), de nucléation (polymérisation) ou en encore de stabilisation (Pollard, 2016).

Les régulateurs de la polymérisation de l'actine interviennent dans de nombreux processus au cours du développement embryonnaire chez la drosophile et leurs rôles sont assez bien décrits. Il existe peu d'information concernant les protéines de dépolymérisation. Néanmoins, deux familles de protéines sont connues pour participer à ce processus chez la drosophile et les mammifères : la famille *gelsolin/villin* et *ADF/cofilin*. Cependant la fonction de ces protéines dans les processus musculaires n'est pas encore pleinement comprise.

Gelsoline fait partie de la famille de protéines du même nom. Cette famille Gelsoline est composée de sept membres : Gelsoline, CapG, Adseverin, Advillin, Villin, Supervillin et flightless I. Gelsoline, le membre éponyme de cette famille contient six domaines Gelsoline (G1-6) dont G1 et G4 qui sont les domaines primaires de liaison à l'actine. Les autres membres de cette famille contiennent différents nombres de domaine Gelsoline (Silacci et al., 2004).





Gelsoline est présente dans le cytoplasme et est maintenue dans un état inactif jusqu'à ce qu'elle soit requise pour la fragmentation de l'actine (Dos Remedios et al., 2003; Allen, 2003; Janmey et Stossel, 1987). La capacité de Gelsoline à cliver les filaments d'actine dépend de la présence de calcium ( $\text{Ca}^{2+}$ ). En l'absence de calcium, les domaines G1 et G4 sont incapables de se lier à l'actine. Du fait de leur localisation à l'intérieur de la protéine ils sont inaccessibles. La liaison du calcium affecte l'affinité et le ratio de liaison de Gelsoline à l'actine (Nag et al., 2013). Dès que la liaison du calcium se fait sur Gelsoline, la protéine change de conformation, permettant ainsi aux domaines G1 et G4 de s'associer avec les filaments d'actine. La protéine s'enveloppe alors autour du filament. Le clivage se fait alors par un mouvement de pince des domaines Gelsoline sur le filament. Après le clivage la protéine reste attachée à l'extrémité du filament pour le coiffer empêchant ainsi une repolymérisation (Burtnick et al., 1997).

Bien que la perte de *Gelsoline* chez la souris ne soit pas létale du fait de la redondance des protéines de dépolymérisation, plusieurs types de défauts ont pu quand même être observés. En effet les fibroblastes mutants pour *Gelsoline* ont un nombre excessif de fibres de stress résistantes à la dépolymérisation et migrent moins vite. Les plaquettes n'arrivent pas à changer de forme durant la coagulation ce qui entraîne l'augmentation de la durée de saignement (Witke et al., 1995). La rétraction des filopodes des neurites (dendrites ou axones) est aussi défectueuse (Lu et al., 1997). Pour finir la formation des podosomes dans les ostéoclastes est bloquée menant à une réduction de la résorption des os et donc à une augmentation de la masse osseuse (Chellaiah et al., 2000).

L'activité de fragmentation et de coiffage de Gelsoline sur les filaments d'actine a donc un impact sur un nombre important de fonctions biologiques. Par conséquent, cette dernière pourrait être requise au cours de différents processus de diversification des cellules musculaires embryonnaires. Par exemple pendant la fusion, le cytosquelette d'actine est réarrangé par des facteurs de nucléation suite au contact du FCM avec le myotube/FC. Un foyer dense d'actine est retrouvé coté FCM alors qu'un simple feuillet est présent dans le myotube. Le foyer d'actine au sein du FCM est à l'origine d'une structure protrusive (podosome) qui va exercer une force sur la membrane du myotube menant à la formation du pore de fusion. Les données récentes suggèrent que la tension régulée par la Myosine II du côté du myotube est nécessaire pour la formation de ce dernier (Kim et al., 2015b) ([voir paragraphe remodelage du cytosquelette d'actine](#)). Le feuillet d'actine dans le myotube



pourrait jouer un rôle dans l'augmentation de la tension au site de fusion et favoriserait donc la formation du pore. De ce fait la dépolymérisation de l'actine par Gelsoline ou la limitation de la polymérisation (coiffage) par dCryAB dans le myotube aurait une influence négative sur la tension et une conséquence directe sur la fusion. Il est donc envisageable que *via* cette dépolymérisation, Gel et dCryAB contrôlent de manière fine et spécifique le programme de fusion des muscles LTs en apportant un signal de fin de ce processus.

Cette hypothèse est en accord avec l'expression tardive de *gel* et *dCryAB* dans les muscles LTs, qui corrèle avec les derniers événements de fusion. De l'autre côté, la dépolymérisation de l'actine est supposée favoriser l'agrandissement du pore et l'endocytose du noyau (Sens et al., 2010). Ce processus, commun pour l'ensemble des événements de fusion, ne pourrait pas être contrôlé par *gel* et *dCryAB* dont l'expression est trop tardive. Il est néanmoins possible que dans le contexte mutant, une perte de dépolymérisation de l'actine induise un blocage de l'endocytose des noyaux menant à une fusion incomplète. Cette possibilité expliquerait la séparation des membranes à l'extrémité des muscles LTs. Il est difficile à ce stade de l'analyse de reconstruire les événements cellulaires qui mènent à la formation de cette séparation. Une possibilité que nous privilégions est que le premier FCM partiellement fusionnée (blocage endocytose) sert comme matrice préférentielle pour les événements de fusion surnuméraires. Par conséquent cela créerait une élongation parallèle d'un syncytium secondaire, partiellement relié au myotube primaire. Dans la majorité des cas l'élargissement d'une myotube mutant séparé favorise l'extension en forme de fourche à l'extrémité du muscle et la spécification d'une cellule tendon supplémentaire pour l'attachement.

Concernant dCryAB celle-ci fait partie de la famille de protéines « small heat shock protein » (sHSP). Chez les mammifères il existe 9 membres qui sont retrouvés exprimés dans un nombre important de tissus et notamment le muscle squelettique (de Jong et al., 1998; Sugiyama et al., 2000). Cette famille de protéines est impliquée dans plusieurs processus cellulaires tels que la réponse au stress, la modulation du cytosquelette d'actine et des filaments intermédiaires, la croissance cellulaire, la différenciation, l'apoptose ou encore la tumorigenèse (Ali and Banu, 1991; Arrigo, 1998). Elles exercent leurs rôles à travers leur activité de chaperonne. Il a cependant été démontré que les sHSP participent à l'inhibition de la polymérisation de l'actine. En effet, sHSP27 est capable de coiffer les filaments d'actine



prévenant ainsi la polymérisation et la croissance de ces derniers (Mounier and Arrigo, 2002). Une étude précédente réalisée au laboratoire a mis en évidence le rôle de dCryAB dans l'intégrité des muscles chez la larve. En effet dCryAB est présente dans le muscle larvaire autour des noyaux ainsi que dans la bande Z des sarcomères. L'atténuation de son expression mène à un patron sarcomérique altéré et à la délocalisation des noyaux (Wojtowicz et al., 2015). Etant donné la similarité des défauts observés entre la perte de fonction de *gel* et *dCryAB* dans les muscles LTs embryonnaires, dCryAB pourrait avoir une fonction proche de celle de Gelsoline dans l'inhibition de la polymérisation. En effet, en condition mutante pour dCryAB une polymérisation excessive d'actine au niveau du site de fusion côté myotube, due à l'absence de coiffage des filaments, entraînerait des défauts de fusion.

Les perspectives sont à présent de suivre la dynamique de l'actine aux cours des étapes tardives de fusion et d'élongation dans les mutants *Gel* et *dCryAB* pour mieux comprendre les évènements cellulaires à l'origine des phénotypes observés. Pour cela des expériences de « live imaging » seront réalisées. De plus, il serait nécessaire de vérifier l'importance de ces défauts en larve afin de savoir si la locomotion est affectée. L'expression restreinte de Gelsoline et dCryAB suggère que le processus de fusion et/ou d'élongation soit régulé de manière muscle spécifique validant ce qui a déjà été montré pour la fusion dans d'autres types de muscles et pour l'attachement des myotubes aux tendons (Bataillé et al., 2010; Lahaye et al., 2012) ([Voir paragraphes, fusion muscle spécifique, guidance et attachement](#)). En effet on sait déjà que la signalisation de guidance des muscles latéraux est différente de celle mise en place pour d'autres types de muscles. Celle ci ne passe pas par Slit/Robo mais par l'expression de Drl à l'extrémité des LTs (Lahaye et al., 2012). Un mécanisme muscles LT spécifique pourrait donc être également mis en jeu pour les processus biologiques tels que la fusion et la croissance des muscles. L'effet de l'expression ectopique de Gelsoline et dCryAB permettrait de déterminer si ces facteurs ont un effet sur la fusion et/ou l'élongation lorsqu'ils sont présents dans d'autres muscles. De plus, Twinstar (Tsr), un orthologue de la protéine de fragmentation de l'actine, Cofilin, est présent chez la drosophile et enrichi dans nos données au stade 16 (T3). Aucune information concernant son rôle dans le muscle n'a été démontré à ce jour. Une redondance fonctionnelle pourrait exister entre ces 2 protéines. La surexpression de Tsr dans les mutants *Gel* permettrait de



vérifier le sauvetage ou non des phénotypes par ce facteur. Pour finir comprendre comment ces gènes sont régulés dans cette sous population reste un enjeu majeur. Il serait donc important de vérifier le niveau d'expression de *gel* et *dCryAB* en conditions mutantes pour les facteurs de transcription clés de la diversification de ces muscles, *Lms* et *Apterous*. D'une autre manière la surexpression de ces facteurs dans d'autres muscles permettrait d'identifier une régulation de l'expression de *Gelsolin* et *dCryAB* par *Lms* et/ou *Ap*. Ces deux protéines semblent essentielles au contrôle des propriétés individuelles des muscles LTs. Les muscles latéraux sont des muscles bien particuliers du fait de leur localisation, leur nombre de noyaux identiques, leurs sites attachement à l'intérieur du segment et leur implication importante dans les mouvements de la larve. Cela suggère la mise en place d'une régulation spécifique et notamment pour les phases tardives de développement. *Gel* et *dCryAB* pourraient être des gènes réalisateurs du code d'identité permettant de réguler finement les étapes de fusion et d'élongation de cette population restreinte de muscles. La poursuite de ces analyses permettrait d'obtenir une meilleure compréhension de leurs fonctions musculaires.

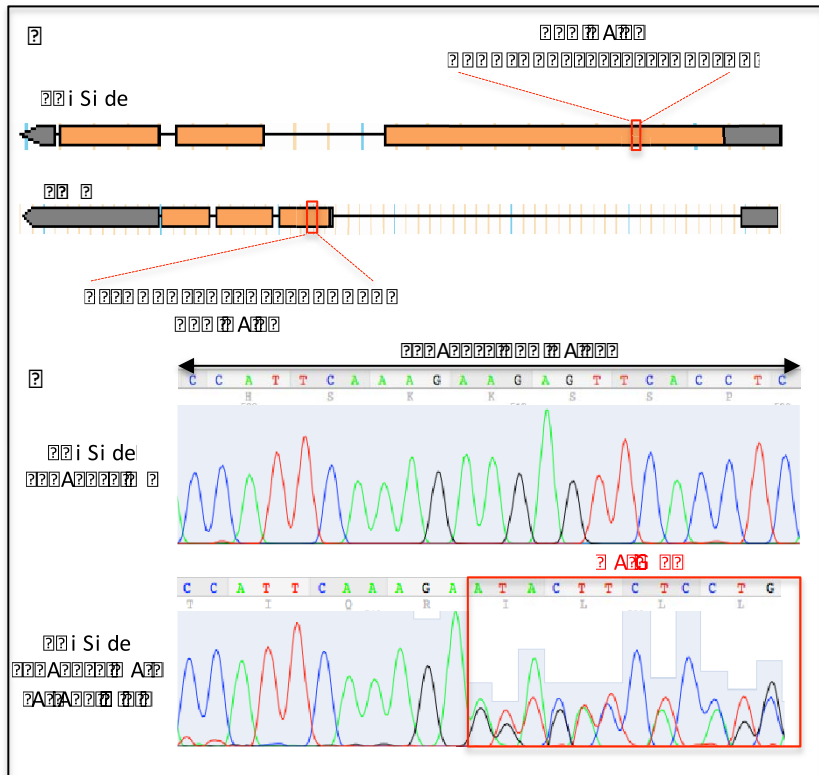
## C- D'autres facteurs potentiels de la diversification musculaire

A travers une analyse différente des gènes différentiellement exprimés entre les populations nous avons sélectionné un ensemble de gènes partageant la même fonction biologique de contrôle de l'expression des gènes. Dans le but d'identifier de nouveaux acteurs du contrôle de l'acquisition de l'identité musculaire notre attention s'est portée sur trois gènes conservés qui sont *bol*, *CG18136* et *kah*. Nous avons tout d'abord validé leur expression musculaire embryonnaire puis des analyses de perte de fonction via ARN interférent ont été réalisées. Ces dernières n'ont pas donné de résultats convaincants au stade embryonnaire. Malgré cela ces gènes restent particulièrement intéressants car ils codent pour des facteurs bien conservés.

### I. Le facteur de transcription *Kah* et l'histone déacétylase *CG18136*

Les orthologues vertébrés de *kah* et *CG18136* sont respectivement *SCRT2* et *SMYD3*. Parmi ces deux gènes seul l'orthologue de *CG18136*, *SMYD3* a une fonction connue dans le muscle squelettique. Chez le poisson zèbre la perte de *Smyd3* résulte en l'expression





The figure illustrates a genomic region and a corresponding Sanger sequencing chromatogram. The top part shows a genomic map with a red box highlighting a specific region. Below it, the Sanger sequencing chromatogram displays the DNA sequence: C C A T T C A A A G A G A G T T C A C C T C. A red box highlights a region where the sequence is A T A C T T C T C C T G.

anormale des marqueurs myogéniques tels que MyoD suggérant que *Smyd3* peut jouer un rôle dans le développement musculaire (Fujii et al., 2011). Une autre étude a montré que *Smyd3* est impliqué dans l'atrophie des muscles. En effet, en condition atrophique induite par le dexaméthasone les muscles du *tibialis anterior* surexpriment *Smyd3*. *Smyd3* régule l'atrophie du muscle squelettique en recrutant des cofacteurs transcriptionnels tels que p-TEFb (positive transcription elongation factor) et PolII<sup>Ser2P</sup> (Ser2-phosphorylated RNA polymerase II) au niveau du gène *myostatin*, qui est un régulateur négatif de la croissance musculaire (Proserpio et al., 2013). Concernant SCRT2, ce dernier est impliqué dans la différenciation de population distinctes de neurones au cours de l'embryogenèse chez les vertébrés (Dam et al., 2011; Vieceli et al., 2013).

A ce jour aucun outil efficace ne permet l'analyse et la caractérisation de la fonction de ces deux gènes en embryon chez la drosophile, il serait donc nécessaire d'en générer pour remédier à cela. Dans cette optique nous avons décidé d'utiliser la méthode d'édition du génome **CRISPR** (Clustered regularly interspaced short palindromic repeats) - **Cas9** (*CRISPR* associated protein 9) pour créer des mutants pour ces deux facteurs. Cette méthode est basée sur une coupure double brin de l'ADN provoqué par l'endonucléase Cas9 qui vient se fixer grâce à un ARN guide à la séquence de l'ADN ciblé (Jinek et al., 2012). Suite à la cassure un mécanisme de réparation de l'ADN par jonction des extrémités non homologues (NHEJ) se met alors en jeu. Ce mécanisme non conservatif contrairement à la recombinaison homologue ne restaure pas la séquence initiale de l'ADN mais seulement la partie endommagée. Ce processus n'étant pas parfait il lui arrive de commettre des erreurs qui entraînent des délétions ou des insertions (ex : codon stop prématuré) à l'origine de protéines non produites ou non fonctionnelles.

Un ARN guide ciblant le premier exon a donc été généré pour chaque gène grâce à l'outil en ligne CRISPR Optimal Target Finder (Gratz et al., 2014) puis cloné dans le plasmide pCFD3:U6:3 (plasmide #49410Addgene)(Port et al., 2014) (Figure 40A). Par la suite ces plasmides ont été injectés dans des embryons *Vasa* (bestgene souche 56552) qui expriment la Cas9 dans les cellules germinales ce qui est nécessaire à l'obtention de lignée mutante stable. J'ai criblé l'ensemble des descendants de ces embryons par PCR (polymerase chain reaction) suivi de séquençage pour identifier de potentielles mutations au niveau des séquences cibles. Pour le gène *kah* seulement une dizaine de larves ont survécu à l'injection et aucune d'entre elles ne présentaient de mutations. En ce qui concerne le *CG18136*, un



positif a été trouvé comme arborant des mutations sur un des deux allèles cependant ce dernier s'est avéré stérile et donc aucune lignée stable n'a pu être établie (Figure 40B). Les perspectives à court terme sont de réinjecter ces plasmides pour obtenir des mutants, nécessaires à la détermination du rôle de ces gènes au cours du développement musculaire. De plus une surexpression du facteur Mef2 permettrait de voir si ce dernier régule directement l'expression de *CG18136* et *kah*. Pour finir des analyses perte de fonction de ces gènes pourraient être également réalisées en larve.

## II. Bol, une nouvelle protéine de liaison à l'ARN nécessaire au cours de la myogenèse

Concernant la protéine de liaison à l'ARN Bol, l'analyse de ce dernier grâce au mutant *bol<sup>40</sup>* nous a permis d'observer des phénotypes musculaires embryonnaires. En effet la présence d'agrégats d'actine semble perturber l'élongation et la structure des muscles. De plus la mort des larves *Bol<sup>40/-</sup>* au stade L1 due à des défauts de locomotion suggère son implication dans la différenciation terminale des muscles. Pour confirmer cette hypothèse nous avons réalisé des analyses de perte de fonction en larve qui ont révélé des phénotypes de désorganisation des sarcomères validant son implication dans la régulation de gènes de structure du muscle.

Boule est un orthologue de la protéine humaine BOL/DAZL (deleted in Azoospermia like). Bol est détectée de manière prédominante dans les testicules d'un grand nombre d'espèces (Shah et al., 2010). Dans les spermatocytes pré-méiotiques, Bol occupe un sous compartiment nucléaire puis à la méiose il est transloqué dans le cytoplasme où il persiste durant la différenciation précoce des spermatides (Cheng et al., 1998; VanGompel and Xu, 2010). La perte de Boule est associée à un arrêt de la méiose chez les hommes stériles (Luetjens et al., 2004). L'absence de Bol chez la drosophile et la souris révèle des défauts de spermatogenèse identiques à ceux observés chez l'homme suggérant que sa fonction dans la régulation de la méiose soit conservée (Eberhart et al., 1996 ; VanGompel and Xu, 2010).

Au niveau régulation il a été montré que Boule se lie à l'ARN messenger et régule la synthèse protéique de ces cibles. En effet l'équipe de Wassermann a dans un premier temps étudié les mutants *Twine* (phosphatase Cdc25) et ils ont démontré que ces derniers présentaient les mêmes phénotypes que *Bol* chez la drosophile (blocage de l'entrée en



méiose)(Maines and Wasserman, 1999). De plus une réduction de l'expression de Twine seulement au niveau protéique et non ARN a été observée dans les mutants *bol*. Pour finir, les défauts d'entrée en méiose présents dans les mutants *bol* sont sauvés par la réexpression de Twine. Cela indique que la fonction essentielle de *bol* dans la transition G2/M au cours de la méiose est de contrôler la traduction de Twine. Cette régulation traductionnelle a ensuite été validée chez le poisson zèbre par le laboratoire de Kunio (Maegawa et al., 2002). Ils ont démontré que DAZL active la traduction des ARNm cibles *via* un mécanisme qui implique la liaison de son domaine RRM à la séquence 'GUUC' ainsi que son association avec des polysomes *via* son domaine DAZ. Par la suite des approches d'immunopurification des ARNs liés à DAZL dans les testicules de souris ont permis de déterminer le motif de liaison de ce dernier aux ARNs. Ce motif consensus,  $_{U2-10}[G/C]_{U2-10}$ , est reporté comme étant plus fréquemment retrouvé dans la partie 3'UTR que 5'UTR des ARNs messagers (Reynolds, 2005). Par la suite il fallait comprendre comment la fixation de DAZL sur la partie 3'UTR des gènes pouvait favoriser la traduction. Pour cela *Collier et al* ont montré que DAZL peut interagir avec des protéines de liaison à la queue de polyA (PolyA Binding Proteins, PABPs), une famille de protéines connue pour se fixer sur la partie 3'UTR des ARNm cibles et stimuler leur traduction *via* l'assemblage de ribosome. Dans cette étude ils ont identifié un site de liaison de PABP dans la séquence de DAZL et la délétion de cette dernière abolit la fonction stimulatrice de traduction de ce dernier (Collier et al., 2005). Pour finir une étude réalisée chez le poisson zèbre a démontré l'activité de répresseur traductionnel de BOL/DAZL (Zeng et al., 2009). En effet cette activité passe par l'interaction de ce dernier avec les protéines PUM2 (Pumilio-2) et DAZAP1 (Deleted in azoospermia-associated protein 1) qui sont deux protéines connues pour être impliquées dans les mécanismes de répression de la traduction (Moore et al., 2003; Morton et al., 2006). A ce jour seule une poignée d'ARN cibles de Bol/Dazl ont été confirmés par des approches biochimiques et seulement dans les cellules germinales (Maines and Wasserman, 1999; Reynolds et al., 2007; Xu et al., 2013; Zeng et al., 2009). Bien que la majorité des études montrent que Bol est restreint à ces cellules il a été aussi retrouvé comme exprimé au niveau du cerveau chez la larve et la drosophile adulte (Hoopfer et al., 2008; Joiner and Wu, 2004). La surexpression de *bol* dans le cerveau induit des défauts synaptiques dans la rétine. L'autre étude a démontré que *bol* est sous l'influence de la voie ecdysone et qu'il arbore une fonction de régulateur négatif de l'élagage synaptique des neurones MB (mushroom body)(Hoopfer et al., 2008)



Grâce à nos analyses nous avons validé que Bol n'est pas seulement une protéine de liaison à l'ARN exprimée dans le cerveau ou les cellules germinales mais bien présent dans le muscle embryonnaire et larvaire de *Drosophile*. De plus, sa localisation granuleuse péri nucléaire chez la larve L3 suggère que ce dernier pourrait être impliqué dans le transport des ARN messagers et le contrôle de la traduction de ses cibles. En effet le mécanisme de control du transport des ARNs le long du cytosquelette a été décrit dans plusieurs modèles (St Johnston, 2005). Concernant plus particulièrement DAZL des expériences de co-immunoprécipitation ont montré que DAZL se lié directement à la protéine Dlc1 (Dynein light chain 1). Cela lui permet ensuite de s'associer au complexe moteur des microtubules dyneine-dynactine (Lee et al., 2006). En suivant les ARNs associés à DAZL ils ont démontré que ces complexe protéine/ARNm s'agrégeaient autour des noyaux dans les cellules et que ces agrégats étaient positifs pour le marqueur de granules ARN, TIA1. Leur hypothèse est donc que DAZL reconnaît un ensemble de transcrits spécifiques, les lie au complexe dyneine-dynactine et les transportent au niveau des granules ARN. Les ARN messagers associés à DAZL dans les granules sont alors maintenus dans un état stable non traduits. Par la suite en réponse à un besoin spécifique de la cellule, le complexe ARNm/DAZL est relargué des granules pour que les ARNm soient traduits.

Dans un second temps nous avons montré que la perte de fonction de *Bol* entraîne des défauts majeurs dans l'organisation du cytosquelette d'actine en embryon ce qui se traduit en larve par une architecture complètement désorganisée des fibres musculaires. Ces phénotypes pourraient être liés à sa fonction dans le transport mais aussi à celle de régulateur traductionnel. En effet, en fonction de son interaction il peut tout aussi bien réprimer que favoriser la traduction d'un type spécifique d'ARNm. Par conséquent, il pourrait intervenir directement dans la régulation d'un ensemble de gènes musculaires requis pour l'organisation sarcomérique des fibres.

Dans cette optique des expériences de RIP (RNA immunopurification) suivi de séquençage ARN pourraient être réalisées afin d'identifier les cibles directes de ce facteur au cours du développement musculaire chez la drosophile. En parallèle, des approches de co-immunoprécipitation permettraient de déterminer les cofacteurs potentiels de Bol au sein du muscle. De plus des mutations des différents domaines de Bol (DAZ, RRM, PABP)





pourraient être réalisés dans le but d'observer les phénotypes associés à la perte de chacun des domaines de cette protéine. Des expériences de TRAP en condition de surexpression pour *bol* comparées à celles déjà obtenues révéleraient les gènes directement affectés au niveau traductionnel. Pour finir étant donné sa régulation potentielle par le facteur myogénique Mef2, la surexpression de Mef2 dans un tissu non musculaire permettrait de savoir si ce facteur suffit pour induire l'expression de *bol* (cible transcriptionnelle directe). Au final Bol semble être un facteur clé impliqué dans plusieurs processus de régulation de l'ARNm, notamment pour la différenciation de plusieurs types cellulaires. La poursuite de ces analyses est donc nécessaire et primordiale à la compréhension de sa fonction musculaire.



Ernest



L'objectif principal de ce travail de thèse était d'étudier les mécanismes génétiques et moléculaires impliqués dans l'acquisition des caractéristiques individuelles des muscles au cours du développement embryonnaire chez la drosophile. Pour répondre à cette question nous avons choisi d'avoir une vision cellule spécifique et dynamique de l'expression génique par l'identification d'un type particulier d'ARN, les ARN messagers en cours de traduction qui sont un reflet de la production de protéines. Pour cela, j'ai optimisé (publication 1) puis utilisé la méthode cellule spécifique TRAP au cours de la myogenèse embryonnaire chez la drosophile (publication 2).

Le projet s'est concentré sur deux populations restreintes de muscles (Lms et Slou) à trois stades de développement différents afin d'identifier de nouveaux gènes participant aux processus majeurs de spécification et de différenciation musculaire. Des analyses translomiques ont permis d'identifier 5 facteurs de liaison à l'ADN plus spécifiquement exprimés dans la population Slou suggérant leur implication dans le code d'identité nécessaire à la diversification de cette population musculaire plus hétérogène. De plus un gène en particulier, *boule*, qui code pour une protéine de liaison à l'ARN conservé chez l'homme a été caractérisé en détail. Celui-ci est exprimé dans le muscle au cours de la vie embryonnaire et larvaire et son absence entraîne des défauts structurels des fibres musculaires avec une désorganisation de cytosquelette actine et une délocalisation des noyaux. Par ces capacités d'interaction avec les ARNs, Boule pourrait directement être impliqué dans le transport et le stockage de certains ARNm musculaires ainsi que dans la régulation de la traduction de ces derniers.

Dans un second temps une comparaison temporelle de l'enrichissement des transcrits dans les deux populations musculaires a permis de mettre en évidence un ensemble de gènes dont l'expression augmente au cours du temps spécifiquement dans les muscles Lms-positifs. Ces gènes codent pour des protéines de liaison à l'actine et parmi eux *gelsolin* et *dCryAB*, présentant une expression spécifique aux muscles Lms-positifs. L'utilisation de mutants conditionnels pour ces deux gènes a révélé des phénotypes musculaires similaires. En effet dans les deux cas nous avons observé des défauts de fusion et d'élongation des myotubes suggérant que ces gènes sont des réalisateurs du code d'identité mis en place pour réguler finement l'acquisition des caractéristiques finales des muscles latéraux.



A présent un enjeu majeur sera de découvrir comment ces gènes régulateurs et réalisateurs de l'identité musculaire sont connectés entre eux, à travers un réseau de gènes spécifiant le destin cellulaire de chaque type de muscle. Avec l'analyse de l'ensemble des données générées par le TRAP et plus récemment par l'INTACT nous espérons acquérir des connaissances supplémentaires sur les gènes qui définissent l'identité de ces différentes sous populations musculaires et que cela aura un impact sur la compréhension de la diversification cellulaire en générale mais également de certaines myopathies. En effet la majorité de ces pathologies affectent différents types de cellules musculaires avec des sensibilités variables. Ces connaissances sont donc nécessaires à l'identification des mécanismes moléculaires affectés dans certains processus pathologiques et peuvent représenter dans le futur une solution pour la mise en place de nouvelles approches thérapeutiques.





??t?n??s? ?Gr?t?| A i e?s?

- Ali, N., and Banu, N. (1991). Heat shock proteins: molecular chaperones. *Biochem. Educ.* *19*, 166–172.
- Amerongen, R., and Nusse, R. (2009). Towards an integrated view of Wnt signaling in development. *Development* *136*, 3205–3214.
- Anant, S., Roy, S., and Raghavan, K.V. (1998). Twist and Notch negatively regulate adult muscle differentiation in *Drosophila*. *Development* *125*, 1361–1369.
- Andreev, D.E., O'Connor, P.B., Fahey, C., Kenny, E.M., Terenin, I.M., Dmitriev, S.E., Cormican, P., Morris, D.W., Shatsky, I.N., and Baranov, P.V. (2015). Translation of 5' leaders is pervasive in genes resistant to eIF2 repression. *Elife* *4*, e03971.
- Arredondo, J.J., Vivar, J., Laine-Menéndez, S., Martínez-Morentin, L., and Cervera, M. (2017). CF2 transcription factor is involved in the regulation of Mef2 RNA levels, nuclei number and muscle fiber size. *PLOS ONE* *12*, e0179194.
- Artero, R. (2003). Notch and Ras signaling pathway effector genes expressed in fusion competent and founder cells during *Drosophila* myogenesis. *Development* *130*, 6257–6272.
- Artero, R.D., Castanon, I., and Baylies, M.K. (2001). The immunoglobulin-like protein Hibris functions as a dose-dependent regulator of myoblast fusion and is differentially controlled by Ras and Notch signaling. *Development* *128*, 4251–4264.
- Aulehla, A., and Pourquié, O. (2006). On periodicity and directionality of somitogenesis. *Brain Struct. Funct.* *211*, 3–8.
- Aulehla, A., Wehrle, C., Brand-Saberi, B., Kemler, R., Gossler, A., Kanzler, B., and Herrmann, B.G. (2003). Wnt3a plays a major role in the segmentation clock controlling somitogenesis. *Dev. Cell* *4*, 395–406.
- Azpiazu, N., Lawrence, P.A., Vincent, J.-P., and Frasch, M. (1996). Segmentation and specification of the *Drosophila* mesoderm. *Genes Dev.* *10*, 3183–3194.
- Baehrecke, E.H. (1997). who encodes a KH RNA binding protein that functions in muscle development. *Development* *124*, 1323–1332.
- Bajard, L. (2006). A novel genetic hierarchy functions during hypaxial myogenesis: Pax3 directly activates Myf5 in muscle progenitor cells in the limb. *Genes Dev.* *20*, 2450–2464.
- Bassett, D.I. (2003). Dystrophin is required for the formation of stable muscle attachments in the zebrafish embryo. *Development* *130*, 5851–5860.
- Bataillé, L., Boukhatmi, H., Frendo, J.-L., and Vincent, A. (2017). Dynamics of transcriptional (re)-programming of syncytial nuclei in developing muscles. *BMC Biol.* *15*.
- Bataillé, L., Delon, I., Da Ponte, J.P., Brown, N.H., and Jagla, K. (2010). Downstream of Identity Genes: Muscle-Type-Specific Regulation of the Fusion Process. *Dev. Cell* *19*, 317–328.
- Bate, M. (1990). The embryonic development of larval muscles in *Drosophila*. *Development* *110*, 791–804.
- Baylies, M.K., Bate, M., and Gomez, M.R. (1998). Myogenesis: a view from *Drosophila*. *Cell* *93*, 921–927.
- Becker, S., Pasca, G., Strumpf, D., Min, L., and Volk, T. (1997). Reciprocal signaling between *Drosophila* epidermal muscle attachment cells and their corresponding muscles. *Development* *124*, 2615–2622.
- Beckett, K., and Baylies, M.K. (2007). 3D analysis of founder cell and fusion competent myoblast arrangements outlines a new model of myoblast fusion. *Dev. Biol.* *309*, 113–125.
- Ben-Yair, R., and Kalcheim, C. (2008). Notch and bone morphogenetic protein differentially act on dermomyotome cells to generate endothelium, smooth, and striated muscle. *J. Cell Biol.* *180*, 607–618.
- Berger, M.F., Philippakis, A.A., Qureshi, A.M., He, F.S., Estep, P.W., and Bulyk, M.L. (2006). Compact, universal DNA microarrays to comprehensively determine transcription-factor binding site specificities. *Nat. Biotechnol.* *24*, 1429–1435.
- Berger, S., Schafer, G., Kesper, D.A., Holz, A., Eriksson, T., Palmer, R.H., Beck, L., Klambt, C., Renkawitz-Pohl, R., and Onel, S.-F. (2008). WASP and SCAR have distinct roles in activating the Arp2/3 complex during myoblast fusion. *J. Cell Sci.* *121*, 1303–1313.

- Bertin, B., Renaud, Y., Aradhya, R., Jagla, K., and Junion, G. (2015). TRAP-rc, Translating Ribosome Affinity Purification from Rare Cell Populations of *Drosophila* Embryos. *J. Vis. Exp.*
- Bi, P., Ramirez-Martinez, A., Li, H., Cannavino, J., McAnally, J.R., Shelton, J.M., Sánchez-Ortiz, E., Bassel-Duby, R., and Olson, E.N. (2017). Control of muscle formation by the fusogenic micropeptide myomixer. *Science* *356*, 323–327.
- Birchmeier, C., and Brohmann, H. (2000). Genes that control the development of migrating muscle precursor cells. *Curr. Opin. Cell Biol.* *12*, 725–730.
- Biressi, S., Tagliafico, E., Lamorte, G., Monteverde, S., Tenedini, E., Roncaglia, E., Ferrari, S., Ferrari, S., Cusella-De Angelis, M.G., Tajbakhsh, S., et al. (2007). Intrinsic phenotypic diversity of embryonic and fetal myoblasts is revealed by genome-wide gene expression analysis on purified cells. *Dev. Biol.* *304*, 633–651.
- 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.
- Bodmer, R. (1993). The gene tinman is required for specification of the heart and visceral muscles in *Drosophila*. *Development* *118*, 719–729.
- Bonanomi, D., and Pfaff, S.L. (2010). Motor Axon Pathfinding. *Cold Spring Harb. Perspect. Biol.* *2*, a001735–a001735.
- Bonn, S., Zinzen, R.P., Perez-Gonzalez, A., Riddell, A., Gavin, A.-C., and Furlong, E.E.M. (2012). Cell type-specific chromatin immunoprecipitation from multicellular complex samples using BiTS-ChIP. *Nat. Protoc.* *7*, 978–994.
- Borello, U. (2006). The Wnt/ -catenin pathway regulates Gli-mediated Myf5 expression during somitogenesis. *Development* *133*, 3723–3732.
- Borello, U., Buffa, V., Sonnino, C., Melchionna, R., Vivarelli, E., and Cossu, G. (1999a). Differential expression of the Wnt putative receptors Frizzled during mouse somitogenesis. *Mech. Dev.* *89*, 173–177.
- Borello, U., Coletta, M., Tajbakhsh, S., Leyns, L., De Robertis, E.M., Buckingham, M., and Cossu, G. (1999b). Transplacental delivery of the Wnt antagonist Frzb1 inhibits development of caudal paraxial mesoderm and skeletal myogenesis in mouse embryos. *Development* *126*, 4247–4255.
- Borycki, A.-G., Brunk, B., Tajbakhsh, S., Buckingham, M., Chiang, C., and Emerson, C.P. (1999). Sonic hedgehog controls epaxial muscle determination through Myf5 activation. *Development* *126*, 4053–4063.
- Bothe, I., and Dietrich, S. (2006). The molecular setup of the avian head mesoderm and its implication for craniofacial myogenesis. *Dev. Dyn.* *235*, 2845–2860.
- Bothe, I., Tenin, G., Oseni, A., and Dietrich, S. (2011). Dynamic control of head mesoderm patterning. *Development* *138*, 2807–2821.
- Bothe, I., Deng, S., and Baylies, M. (2014). PI(4,5)P2 regulates myoblast fusion through Arp2/3 regulator localization at the fusion site. *Development* *141*, 2289–2301.
- Boukhatmi, H., Frendo, J.L., Enriquez, J., Crozatier, M., Dubois, L., and Vincent, A. (2012). Tup/Islet1 integrates time and position to specify muscle identity in *Drosophila*. *Development* *139*, 3572–3582.
- Bour, B.A., O'Brien, M.A., Lockwood, W.L., Goldstein, E.S., Bodmer, R., Taghert, P.H., Abmayr, S.M., and Nguyen, H.T. (1995). *Drosophila* MEF2, a transcription factor that is essential for myogenesis. *Genes Dev.* *9*, 730–741.
- Bour, B.A., Chakravarti, M., West, J.M., and Abmayr, S.M. (2000). *Drosophila* SNS, a member of the immunoglobulin superfamily that is essential for myoblast fusion. *Genes Dev.* *14*, 1498–1511.
- Boyle, E.I., Weng, S., Gollub, J., Jin, H., Botstein, D., Cherry, J.M., and Sherlock, G. (2004). GO::TermFinder--open source software for accessing Gene Ontology information and finding significantly enriched Gene Ontology terms associated with a list of genes. *Bioinformatics* *20*, 3710–3715.
- Boyle, M., Bonini, N., and DiNardo, S. (1997). Expression and function of clift in the development of somatic gonadal precursors within the *Drosophila* mesoderm. *Development* *124*, 971–982.
- 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.

- Braun, T., Bober, E., Winter, B., Rosenthal, N., and Arnold, H.H. (1990). Myf-6, a new member of the human gene family of myogenic determination factors: evidence for a gene cluster on chromosome 12. *EMBO J.* *9*, 821.
- Brent, A.E., Schweitzer, R., and Tabin, C.J. (2003). A somitic compartment of tendon progenitors. *Cell* *113*, 235–248.
- Broadie, K.S., and Bate, M. (1993). Development of the embryonic neuromuscular synapse of *Drosophila melanogaster*. *J. Neurosci.* *13*, 144–166.
- Brunelli, S., Relaix, F., Baesso, S., Buckingham, M., and Cossu, G. (2007). Beta catenin-independent activation of MyoD in presomitic mesoderm requires PKC and depends on Pax3 transcriptional activity. *Dev. Biol.* *304*, 604–614.
- Buckingham, M. (2001). Skeletal muscle formation in vertebrates. *Curr. Opin. Genet. Dev.* *11*, 440–448.
- Buckingham, M. (2006). Myogenic progenitor cells and skeletal myogenesis in vertebrates. *Curr. Opin. Genet. Dev.* *16*, 525–532.
- Buckingham, M., and 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.* *23*, 645–673.
- Buckingham, M., and Vincent, S.D. (2009). Distinct and dynamic myogenic populations in the vertebrate embryo. *Curr. Opin. Genet. Dev.* *19*, 444–453.
- Buff, E., Carmena, A., Gisselbrecht, S., Jiménez, F., and Michelson, A.M. (1998). Signalling by the *Drosophila* epidermal growth factor receptor is required for the specification and diversification of embryonic muscle progenitors. *Development* *125*, 2075–2086.
- Bürckstümmer, T., Bennett, K.L., Preradovic, A., Schütze, G., Hantschel, O., Superti-Furga, G., and Bauch, A. (2006). An efficient tandem affinity purification procedure for interaction proteomics in mammalian cells. *Nat. Methods* *3*, 1013–1019.
- Burtnick, L.D., Koepf, E.K., Grimes, J., Jones, E.Y., Stuart, D.I., McLaughlin, P.J., and Robinson, R.C. (1997). The crystal structure of plasma gelsolin: implications for actin severing, capping, and nucleation. *Cell* *90*, 661–670.
- Busser, B.W., Shokri, L., Jaeger, S.A., Gisselbrecht, S.S., Singhania, A., Berger, M.F., Zhou, B., Bulyk, M.L., and Michelson, A.M. (2012). Molecular mechanism underlying the regulatory specificity of a *Drosophila* homeodomain protein that specifies myoblast identity. *Development* *139*, 1164–1174.
- Busser, B.W., Gisselbrecht, S.S., Shokri, L., Tansey, T.R., Gamble, C.E., Bulyk, M.L., and Michelson, A.M. (2013). Contribution of Distinct Homeodomain DNA Binding Specificities to *Drosophila* Embryonic Mesodermal Cell-Specific Gene Expression Programs. *PLoS ONE* *8*, e69385.
- Cabrera, C.V., and Alonso, M.C. (1991). Transcriptional activation by heterodimers of the achaete-scute and daughterless gene products of *Drosophila*. *EMBO J.* *10*, 2965.
- Callahan, C.A., Bonkovsky, J.L., Scully, A.L., and Thomas, J.B. (1996). *derailed* is required for muscle attachment site selection in *Drosophila*. *Development* *122*, 2761–2767.
- Carmena, A., Bate, M., and Jimenez, F. (1995). Lethal of scute, a proneural gene, participates in the specification of muscle progenitors during *Drosophila* embryogenesis. *Genes Dev.* *9*, 2373–2383.
- Carmena, A., Gisselbrecht, S., Harrison, J., Jiménez, F., and Michelson, A.M. (1998a). Combinatorial signaling codes for the progressive determination of cell fates in the *Drosophila* embryonic mesoderm. *Genes Dev.* *12*, 3910–3922.
- Carmena, A., Murugasu-Oei, B., Menon, D., Jiménez, F., and Chia, W. (1998b). *inscuteable* and *numb* mediate asymmetric muscle progenitor cell divisions during *Drosophila* myogenesis. *Genes Dev.* *12*, 304–315.
- Carmignac, V., and Durbeej, M. (2012). Cell-matrix interactions in muscle disease. *J. Pathol.* *226*, 200–218.
- Carrasco-Rando, M., Tutor, A.S., Prieto-Sánchez, S., González-Pérez, E., Barrios, N., Letizia, A., Martín, P., Campuzano, S., and Ruiz-Gómez, M. (2011). *Drosophila* Araucan and Caupolican Integrate Intrinsic and Signalling Inputs for the Acquisition by Muscle Progenitors of the Lateral Transverse Fate. *PLoS Genet.* *7*, e1002186.

- Castanon, I., Von Stetina, S., Kass, J., and Baylies, M.K. (2001). Dimerization partners determine the activity of the Twist bHLH protein during *Drosophila* mesoderm development. *Development* *128*, 3145–3159.
- Cavodeassi, F., Modolell, J., and Campuzano, S. (2000). The Iroquois homeobox genes function as dorsal selectors in the *Drosophila* head. *Development* *127*, 1921–1929.
- Cavodeassi, F., Modolell, J., and Gómez-Skarmeta, J.L. (2001). The Iroquois family of genes: from body building to neural patterning. *Development* *128*, 2847–2855.
- Charvet, B., Ruggiero, F., and Le Guellec, D. (2012). The development of the myotendinous junction. A review. *MLTJ Muscles Ligaments Tendons J.* *2*, 53–63.
- Chellaiah, M., Kizer, N., Silva, M., Alvarez, U., Kwiatkowski, D., and Hruska, K.A. (2000). Gelsolin deficiency blocks podosome assembly and produces increased bone mass and strength. *J. Cell Biol.* *148*, 665–678.
- Chen, E.H., and Olson, E.N. (2001). Antisocial, an intracellular adaptor protein, is required for myoblast fusion in *Drosophila*. *Dev. Cell* *1*, 705–715.
- Chen, K., and Featherstone, D.E. (2005). Discs-large (DLG) is clustered by presynaptic innervation and regulates postsynaptic glutamate receptor subunit composition in *Drosophila*. *BMC Biol.* *3*, 1.
- Chen, E.H., Pryce, B.A., Tzeng, J.A., Gonzalez, G.A., and Olson, E.N. (2003). Control of myoblast fusion by a guanine nucleotide exchange factor, loner, and its effector ARF6. *Cell* *114*, 751–762.
- Chhabra, E.S., and Higgs, H.N. (2007). The many faces of actin: matching assembly factors with cellular structures. *Nat. Cell Biol.* *9*, 1110–1121.
- Chi, N., and Epstein, J.A. (2002). Getting your Pax straight: Pax proteins in development and disease. *TRENDS Genet.* *18*, 41–47.
- Choi, J., Costa, M.L., Mermelstein, C.S., Chagas, C., Holtzer, S., and Holtzer, H. (1990). MyoD converts primary dermal fibroblasts, chondroblasts, smooth muscle, and retinal pigmented epithelial cells into striated mononucleated myoblasts and multinucleated myotubes. *Proc. Natl. Acad. Sci.* *87*, 7988–7992.
- Ciglar, L., Girardot, C., Wilczyński, B., Braun, M., and Furlong, E.E.M. (2014). Coordinated repression and activation of two transcriptional programs stabilizes cell fate during myogenesis. *Development* *141*, 2633–2643.
- Clark, I.B.N., Boyd, J., Hamilton, G., Finnegan, D.J., and Jarman, A.P. (2006). D-six4 plays a key role in patterning cell identities deriving from the *Drosophila* mesoderm. *Dev. Biol.* *294*, 220–231.
- Clark, K.A., McElhinny, A.S., Beckerle, M.C., and Gregorio, C.C. (2002). Striated Muscle Cytoarchitecture: An Intricate Web of Form and Function. *Annu. Rev. Cell Dev. Biol.* *18*, 637–706.
- Collier, B., Gorgoni, B., Loveridge, C., Cooke, H.J., and Gray, N.K. (2005). The DAZL family proteins are PABP-binding proteins that regulate translation in germ cells. *EMBO J.* *24*, 2656–2666.
- Conti, F.J., Felder, A., Monkley, S., Schwander, M., Wood, M.R., Lieber, R., Critchley, D., and Muller, U. (2008). Progressive myopathy and defects in the maintenance of myotendinous junctions in mice that lack talin 1 in skeletal muscle. *Development* *135*, 2043–2053.
- Conti, F.J., Monkley, S.J., Wood, M.R., Critchley, D.R., and Muller, U. (2009). Talin 1 and 2 are required for myoblast fusion, sarcomere assembly and the maintenance of myotendinous junctions. *Development* *136*, 3597–3606.
- Conway, K., Pin, C., Kiernan, J.A., and Merrifield, P. (2004). The E protein HEB is preferentially expressed in developing muscle. *Differentiation* *72*, 327–340.
- Cooke, J., and Zeeman, E.C. (1976). A clock and wavefront model for control of the number of repeated structures during animal morphogenesis. *J. Theor. Biol.* *58*, 455–476.
- del Corral, R.D., Olivera-Martinez, I., Goriely, A., Gale, E., Maden, M., and Storey, K. (2003). Opposing FGF and retinoid pathways control ventral neural pattern, neuronal differentiation, and segmentation during body axis extension. *Neuron* *40*, 65–79.
- Cossu, G., and Borello, U. (1999). Wnt signaling and the activation of myogenesis in mammals. *EMBO J.* *18*, 6867–6872.
- 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.
- Cripps, R.M., Lovato, T.L., and Olson, E.N. (2004). Positive autoregulation of the Myocyte enhancer factor-2 myogenic control gene during somatic muscle development in *Drosophila*. *Dev. Biol.* *267*, 536–547.
- Crisp, S., Evers, J.F., Fiala, A., and Bate, M. (2008). The development of motor coordination in *Drosophila* embryos. *Development* *135*, 3707–3717.
- Cronmiller, C., and Cummings, C.A. (1993). The daughterless gene product in *Drosophila* is a nuclear protein that is broadly expressed throughout the organism during development. *Mech. Dev.* *42*, 159–169.
- Crowner, D., Le Gall, M., Gates, M.A., and Giniger, E. (2003). Notch steers *Drosophila* ISNb motor axons by regulating the Abl signaling pathway. *Curr. Biol.* *13*, 967–972.
- Cserjesi, P., Brown, D., Ligon, K.L., Lyons, G.E., Copeland, N.G., Gilbert, D.J., Jenkins, N.A., and Olson, E.N. (1995). Scleraxis: a basic helix-loop-helix protein that prefigures skeletal formation during mouse embryogenesis. *Development* *121*, 1099–1110.
- Cunha, P.M.F., Sandmann, T., Gustafson, E.H., Ciglar, L., Eichenlaub, M.P., and Furlong, E.E.M. (2010). Combinatorial Binding Leads to Diverse Regulatory Responses: Lmd Is a Tissue-Specific Modulator of Mef2 Activity. *PLoS Genet.* *6*, e1001014.
- D’Alessio, M., and Frasch, M. (1996). Msh may play a conserved role in dorsoventral patterning of the neuroectoderm and mesoderm. *Mech. Dev.* *58*, 217220–231.
- Dam, T.-M.-T., Kim, H.-T., Moon, H.-Y., Hwang, K.-S., Jeong, Y.-M., You, K.-H., Lee, J.-S., and Kim, C.-H. (2011). Neuron-specific expression of Scratch genes during early zebrafish development. *Mol. Cells* *31*, 471–475.
- Darabid, H., Perez-Gonzalez, A.P., and Robitaille, R. (2014). Neuromuscular synaptogenesis: coordinating partners with multiple functions. *Nat. Rev. Neurosci.* *15*, 703–718.
- Davis, G.W., Schuster, C.M., and Goodman, C.S. (1997). Genetic analysis of the mechanisms controlling target selection: target-derived Fasciclin II regulates the pattern of synapse formation. *Neuron* *19*, 561–573.
- Davis, R.L., Weintraub, H., and Lassar, A.B. (1987). Expression of a single transfected cDNA converts fibroblasts to myoblasts. *Cell* *51*, 987–1000.
- Deal, R.B., and Henikoff, S. (2010). A Simple Method for Gene Expression and Chromatin Profiling of Individual Cell Types within a Tissue. *Dev. Cell* *18*, 1030–1040.
- Diehl, A.G., Zarepari, S., Qian, M., Khanna, R., Angeles, R., and Gage, P.J. (2006). Extraocular Muscle Morphogenesis and Gene Expression Are Regulated by *Pitx2* Gene Dose. *Investig. Ophthalmology Vis. Sci.* *47*, 1785.
- Dietrich, S., Abou-Rebyeh, F., Brohmann, H., Bladt, F., Sonnenberg-Riethmacher, E., Yamaai, T., Lumsden, A., Brand-Saberi, B., and Birchmeier, C. (1999). The role of SF/HGF and c-Met in the development of skeletal muscle. *Development* *126*, 1621–1629.
- Doberstein, S.K., Fetter, R.D., Mehta, A.Y., and 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.
- Dohrmann, C., Azpiazu, N., and Frasch, M. (1990). A new *Drosophila* homeo box gene is expressed in mesodermal precursor cells of distinct muscles during embryogenesis. *Genes Dev.* *4*, 2098–2111.
- Dos Remedios, C.G., Chhabra, D., Kekic, M., Dedova, I.V., Tsubakihara, M., Berry, D.A., and Nosworthy, N.J. (2003). Actin Binding Proteins: Regulation of Cytoskeletal Microfilaments. *Physiol. Rev.* *83*, 433–473.
- Du, C., Jin, Y.-Q., Qi, J.-J., Ji, Z.-X., Li, S.-Y., An, G.-S., Jia, H.-T., and Ni, J.-H. (2012). Effects of myogenin on expression of late muscle genes through MyoD-dependent chromatin remodeling ability of myogenin. *Mol. Cells* *34*, 133–142.
- Duan, H., Skeath, J.B., and Nguyen, H.T. (2001). *Drosophila* *Lame duck*, a novel member of the Gli superfamily, acts as a key regulator of myogenesis by controlling fusion-competent myoblast development. *Development* *128*, 4489–4500.
- Duan, H., Zhang, C., Chen, J., Sink, H., Frei, E., and Noll, M. (2007). A key role of *Pox meso* in somatic

- myogenesis of *Drosophila*. *Development* *134*, 3985–3997.
- Dubrulle, J., McGrew, M.J., and Pourquié, O. (2001). FGF signaling controls somite boundary position and regulates segmentation clock control of spatiotemporal Hox gene activation. *Cell* *106*, 219–232.
- Eberhard, D., Jiménez, G., Heavey, B., and Busslinger, M. (2000). Transcriptional repression by Pax5 (BSAP) through interaction with corepressors of the Groucho family. *EMBO J.* *19*, 2292–2303.
- Echelard, Y., Epstein, D.J., St-Jacques, B., Shen, L., Mohler, J., McMahon, J.A., and McMahon, A.P. (1993). Sonic hedgehog, a member of a family of putative signaling molecules, is implicated in the regulation of CNS polarity. *Cell* *75*, 1417–1430.
- Edmondson, D.G., and Olson, E.N. (1989). A gene with homology to the myc similarity region of MyoD1 is expressed during myogenesis and is sufficient to activate the muscle differentiation program. *Genes Dev.* *3*, 628–640.
- Elhanany-Tamir, H., Yu, Y.V., Shnyder, M., Jain, A., Welte, M., and Volk, T. (2012). Organelle positioning in muscles requires cooperation between two KASH proteins and microtubules. *J. Cell Biol.* *198*, 833–846.
- Eliasson, P., Fahlgren, A., and Aspenberg, P. (2008). Mechanical Load and BMP Signaling During Tendon Repair: A Role for Follistatin? *Clin. Orthop.* *466*, 1592–1597.
- Enriquez, J., de Taffin, M., Crozatier, M., Vincent, A., and Dubois, L. (2012). Combinatorial coding of *Drosophila* muscle shape by Collier and Nautilus. *Dev. Biol.* *363*, 27–39.
- Ervasti, J.M. (2003). Costameres: the Achilles' Heel of Herculean Muscle. *J. Biol. Chem.* *278*, 13591–13594.
- Estrada, B., Choe, S.E., Gisselbrecht, S.S., Michaud, S., Raj, L., Busser, B.W., Halfon, M.S., Church, G.M., and Michelson, A.M. (2006). An Integrated Strategy for Analyzing the Unique Developmental Programs of Different Myoblast Subtypes. *PLoS Genet.* *2*, e16.
- Estrada, B., Maeland, A.D., Gisselbrecht, S.S., Bloor, J.W., Brown, N.H., and Michelson, A.M. (2007a). The MARVEL domain protein, Singles Bar, is required for progression past the pre-fusion complex stage of myoblast fusion. *Dev. Biol.* *307*, 328–339.
- Estrada, B., Gisselbrecht, S.S., and Michelson, A.M. (2007b). The transmembrane protein Perdido interacts with Grip and integrins to mediate myotube projection and attachment in the *Drosophila* embryo. *Development* *134*, 4469–4478.
- Feng, G., Mellor, R.H., Bernstein, M., Keller-Peck, C., Nguyen, Q.T., Wallace, M., Nerbonne, J.M., Lichtman, J.W., and Sanes, J.R. (2000). Imaging neuronal subsets in transgenic mice expressing multiple spectral variants of GFP. *Neuron* *28*, 41–51.
- Fernandes, J., Bate, M., and Vijayraghavan, K. (1991). Development of the indirect flight muscles of *Drosophila*. *Development* *113*, 67–77.
- Figeac, N., Daczewska, M., Marcelle, C., and Jagla, K. (2007). Muscle stem cells and model systems for their investigation. *Dev. Dyn.* *236*, 3332–3342.
- Fogerty, F.J., Fessler, L.I., Bunch, T.A., Yaron, Y., Parker, C.G., Nelson, R.E., Brower, D.L., Gullberg, D., and Fessler, J.H. (1994). Tiggrin, a novel *Drosophila* extracellular matrix protein that functions as a ligand for *Drosophila* alpha PS2 beta PS integrins. *Development* *120*, 1747–1758.
- Folker, E.S., Schulman, V.K., and Baylies, M.K. (2012). Muscle length and myonuclear position are independently regulated by distinct Dynein pathways. *Development* *139*, 3827–3837.
- Folker, E.S., Schulman, V.K., and Baylies, M.K. (2014). Translocating myonuclei have distinct leading and lagging edges that require Kinesin and Dynein. *Development* *141*, 355–366.
- Frasch, M. (1999). Controls in patterning and diversification of somatic muscles during *Drosophila* embryogenesis. *Curr. Opin. Genet. Dev.* *9*, 522–529.
- Frontera, W.R., and Ochala, J. (2015). Skeletal Muscle: A Brief Review of Structure and Function. *Calcif. Tissue Int.* *96*, 183–195.
- Fujii, T., Tsunesumi, S., Yamaguchi, K., Watanabe, S., and Furukawa, Y. (2011). Smyd3 Is Required for the Development of Cardiac and Skeletal Muscle in Zebrafish. *PLoS ONE* *6*, e23491.
- Fukui, A., Inaki, M., Tono, G., Hamatani, H., Homma, M., Morimoto, T., Aburatani, H., and Nose, A. (2012). Lola regulates glutamate receptor expression at the *Drosophila* neuromuscular junction. *Biol. Open* *1*, 362–375.



- Gage, P.J., Suh, H., and Camper, S.A. (1999). Dosage requirement of Pitx2 for development of multiple organs. *Development* 126, 4643–4651.
- Gelfi, C., Vasso, M., and Cerretelli, P. (2011). Diversity of human skeletal muscle in health and disease: Contribution of proteomics. *J. Proteomics* 74, 774–795.
- Ghavi-Helm, Y., Klein, F.A., Pakozdi, T., Ciglar, L., Noordermeer, D., Huber, W., and Furlong, E.E.M. (2014). Enhancer loops appear stable during development and are associated with paused polymerase. *Nature*.
- Gilsohn, E., and Volk, T. (2010a). A screen for tendon-specific genes uncovers new and old components involved in muscle-tendon interaction. *Fly (Austin)* 4, 149–153.
- Gilsohn, E., and Volk, T. (2010b). Slowdown promotes muscle integrity by modulating integrin-mediated adhesion at the myotendinous junction. *Development* 137, 785–794.
- Giordani, J., Bajard, L., Demignon, J., Daubas, P., Buckingham, M., and Maire, P. (2007). Six proteins regulate the activation of Myf5 expression in embryonic mouse limbs. *Proc. Natl. Acad. Sci.* 104, 11310–11315.
- Gómez-Skarmeta, J.-L., del Corral, R.D., de la Calle-Mustienes, E., Ferrés-Marcó, D., and Modolell, J. (1996). Araucan and caupolican, two members of the novel iroquois complex, encode homeoproteins that control proneural and vein-forming genes. *Cell* 85, 95–105.
- Gratz, S.J., Ukken, F.P., Rubinstein, C.D., Thiede, G., Donohue, L.K., Cummings, A.M., and O'Connor-Giles, K.M. (2014). Highly Specific and Efficient CRISPR/Cas9-Catalyzed Homology-Directed Repair in *Drosophila*. *Genetics* 196, 961–971.
- Gregorio, C.C., Trombitás, K., Centner, T., Kolmerer, B., Stier, G., Kunke, K., Suzuki, K., Obermayr, F., Herrmann, B., Granzier, H., et al. (1998). The NH2 terminus of titin spans the Z-disc: its interaction with a novel 19-kD ligand (T-cap) is required for sarcomeric integrity. *J. Cell Biol.* 143, 1013–1027.
- Grifone, R. (2005). Six1 and Six4 homeoproteins are required for Pax3 and Mrf expression during myogenesis in the mouse embryo. *Development* 132, 2235–2249.
- Grifone, R., Demignon, J., Giordani, J., Niro, C., Souil, E., Bertin, F., Laclef, C., Xu, P.-X., and Maire, P. (2007). Eya1 and Eya2 proteins are required for hypaxial somitic myogenesis in the mouse embryo. *Dev. Biol.* 302, 602–616.
- Grifone, R., Jarry, T., Dandonneau, M., Grenier, J., Duprez, D., and Kelly, R.G. (2008). Properties of branchiomeric and somite-derived muscle development in Tbx1 mutant embryos. *Dev. Dyn.* 237, 3071–3078.
- Gros, J., Scaal, M., and Marcelle, C. (2004). A two-step mechanism for myotome formation in chick. *Dev. Cell* 6, 875–882.
- Großhans, J., and Wieschaus, E. (2000). A genetic link between morphogenesis and cell division during formation of the ventral furrow in *Drosophila*. *Cell* 101, 523–531.
- Gustafsson, M.K., Pan, H., Pinney, D.F., Liu, Y., Lewandowski, A., Epstein, D.J., and Emerson, C.P. (2002). Myf5 is a direct target of long-range Shh signaling and Gli regulation for muscle specification. *Genes Dev.* 16, 114–126.
- Haenni, S., Ji, Z., Hoque, M., Rust, N., Sharpe, H., Eberhard, R., Browne, C., Hengartner, M.O., Mellor, J., Tian, B., et al. (2012). Analysis of *C. elegans* intestinal gene expression and polyadenylation by fluorescence-activated nuclei sorting and 3'-end-seq. *Nucleic Acids Res.* 40, 6304–6318.
- Halfon, M.S., Carmena, A., Gisselbrecht, S., Sackerson, C.M., Jiménez, F., Baylies, M.K., and Michelson, A.M. (2000). Ras pathway specificity is determined by the integration of multiple signal-activated and tissue-restricted transcription factors. *Cell* 103, 63–74.
- Haralalka, S., Shelton, C., Cartwright, H.N., Katzfey, E., Janzen, E., and Abmayr, S.M. (2013). Asymmetric Mbc, active Rac1 and F-actin foci in the fusion-competent myoblasts during myoblast fusion in *Drosophila*. *Development* 140, 1370–1371.
- Harel, I., Nathan, E., Tirosh-Finkel, L., Zigdon, H., Guimarães-Camboa, N., Evans, S.M., and Tzahor, E. (2009). Distinct Origins and Genetic Programs of Head Muscle Satellite Cells. *Dev. Cell* 16, 822–832.
- Hatini, V., and DiNardo, S. (2001). Distinct signals generate repeating striped pattern in the embryonic parasegment. *Mol. Cell* 7, 151–160.
- Hebert, S.L., Daniel, M.L., and McLoon, L.K. (2013). The Role of Pitx2 in Maintaining the Phenotype of

- Myogenic Precursor Cells in the Extraocular Muscles. *PLoS ONE* 8, e58405.
- Heiman, M., Schaefer, A., Gong, S., Peterson, J.D., Day, M., Ramsey, K.E., Suárez-Fariñas, M., Schwarz, C., Stephan, D.A., Surmeier, D.J., et al. (2008). A Translational Profiling Approach for the Molecular Characterization of CNS Cell Types. *Cell* 135, 738–748.
- Henry, G.L., Davis, F.P., Picard, S., and Eddy, S.R. (2012). Cell type-specific genomics of *Drosophila* neurons. *Nucleic Acids Res.* 40, 9691–9704.
- Hirsinger, E., Duprez, D., Jouve, C., Malapert, P., Cooke, J., and Pourquié, O. (1997). Noggin acts downstream of Wnt and Sonic Hedgehog to antagonize BMP4 in avian somite patterning. *Development* 124, 4605–4614.
- Hollnagel, A., Grund, C., Franke, W.W., and Arnold, H.-H. (2002). The Cell Adhesion Molecule M-Cadherin Is Not Essential for Muscle Development and Regeneration. *Mol. Cell. Biol.* 22, 4760–4770.
- Hoopfer, E.D., Penton, A., Watts, R.J., and Luo, L. (2008). Genomic Analysis of *Drosophila* Neuronal Remodeling: A Role for the RNA-Binding Protein Boule as a Negative Regulator of Axon Pruning. *J. Neurosci.* 28, 6092–6103.
- Hoppeler, H., Lüthi, P., Claassen, H., Weibel, E.R., and Howald, H. (1973). The ultrastructure of the normal human skeletal muscle. *Pflüg. Arch. Eur. J. Physiol.* 344, 217–232.
- Horsley, V., and Pavlath, G.K. (2004). Forming a Multinucleated Cell: Molecules That Regulate Myoblast Fusion. *Cells Tissues Organs* 176, 67–78.
- Horton, E.R., Astudillo, P., Humphries, M.J., and Humphries, J.D. (2016). Mechanosensitivity of integrin adhesion complexes: role of the consensus adhesome. *Exp. Cell Res.* 343, 7–13.
- Huang, R., Zhi, Q., Wilting, J., and Christ, B. (1994). The fate of somitocoel cells in avian embryos. *Anat. Embryol. (Berl.)* 190, 243–250.
- Hutcheson, D.A., Zhao, J., Merrell, A., Haldar, M., and Kardon, G. (2009). Embryonic and fetal limb myogenic cells are derived from developmentally distinct progenitors and have different requirements for -catenin. *Genes Dev.* 23, 997–1013.
- Ibarra, N., Pollitt, A., and Insall, R.H. (2005). Regulation of actin assembly by SCAR/WAVE proteins (Portland Press Limited).
- Ikeya, M., and Takada, S. (1998). Wnt signaling from the dorsal neural tube is required for the formation of the medial dermomyotome. *Development* 125, 4969–4976.
- Inaki, M., Yoshikawa, S., Thomas, J.B., Aburatani, H., and Nose, A. (2007). Wnt4 Is a Local Repulsive Cue that Determines Synaptic Target Specificity. *Curr. Biol.* 17, 1574–1579.
- Inaki, M., Shinza-Kameda, M., Ismat, A., Frasch, M., and Nose, A. (2010). *Drosophila* Tey represses transcription of the repulsive cue Toll and generates neuromuscular target specificity. *Development* 137, 2139–2146.
- Ingolia, N.T., Ghaemmaghami, S., Newman, J.R.S., and Weissman, J.S. (2009). Genome-Wide Analysis in Vivo of Translation with Nucleotide Resolution Using Ribosome Profiling. *Science* 324, 218–223.
- Ingolia, N.T., Lareau, L.F., and Weissman, J.S. (2011). Ribosome Profiling of Mouse Embryonic Stem Cells Reveals the Complexity and Dynamics of Mammalian Proteomes. *Cell* 147, 789–802.
- Ip, Y.T., Park, R.E., Kosman, D., Yazdanbakhsh, K., and Levine, M. (1992). dorsal-twist interactions establish snail expression in the presumptive mesoderm of the *Drosophila* embryo. *Genes Dev.* 6, 1518–1530.
- Jagla, T., Bellard, F., Lutz, Y., Dretzen, G., Bellard, M., and Jagla, K. (1998). ladybird determines cell fate decisions during diversification of *Drosophila* somatic muscles. *Development* 125, 3699–3708.
- Jani, K., and Schöck, F. (2007). Zasp is required for the assembly of functional integrin adhesion sites. *J. Cell Biol.* 179, 1583–1597.
- Jiang, P., Enomoto, A., and Takahashi, M. (2009). Cell biology of the movement of breast cancer cells: Intracellular signalling and the actin cytoskeleton. *Cancer Lett.* 284, 122–130.
- Jiao, Y., and Meyerowitz, E.M. (2010). Cell-type specific analysis of translating RNAs in developing flowers reveals new levels of control. *Mol. Syst. Biol.* 6.
- Johansen, J., Halpern, M.E., and Keshishian, H. (1989). Axonal guidance and the development of muscle fiber-specific innervation in *Drosophila* embryos. *J. Neurosci.* 9, 4318–4332.
- Joiner, M.-L.A., and Wu, C.-F. (2004). NERVOUS SYSTEM FUNCTION FOR THE TESTIS RNA-BINDING

- PROTEIN BOULE IN *DROSOPHILA*. *J. Neurogenet.* *18*, 341–363.
- de Jong, W.W., Caspers, G.-J., and Leunissen, J.A. (1998). Genealogy of the  $\alpha$ -crystallin—small heat-shock protein superfamily. *Int. J. Biol. Macromol.* *22*, 151–162.
- Junion, G., Jagla, T., Duplant, S., Tapin, R., Da Ponte, J.-P., and Jagla, K. (2005). Mapping Dmef2-binding regulatory modules by using a ChIP-enriched in silico targets approach. *Proc. Natl. Acad. Sci. U. S. A.* *102*, 18479–18484.
- Junion, G., Bataille, L., Jagla, T., Da Ponte, J.P., Tapin, R., and Jagla, K. (2007). Genome-wide view of cell fate specification: ladybird acts at multiple levels during diversification of muscle and heart precursors. *Genes Dev.* *21*, 3163–3180.
- Kablar, B., Asakura, A., Krastel, K., Ying, C., May, L.L., Goldhamer, D.J., and Rudnicki, M.A. (1998). MyoD and Myf-5 define the specification of musculature of distinct embryonic origin. *Biochem. Cell Biol.* *76*, 1079–1091.
- Kaipa, B.R., Shao, H., Schafer, G., Trinkewitz, T., Groth, V., Liu, J., Beck, L., Bogdan, S., Abmayr, S.M., and Onel, S.-F. (2013). Dock mediates Scar- and WASp-dependent actin polymerization through interaction with cell adhesion molecules in founder cells and fusion-competent myoblasts. *J. Cell Sci.* *126*, 360–372.
- Kardon, G. (1998). Muscle and tendon morphogenesis in the avian hind limb. *Development* *125*, 4019–4032.
- Kassar-Duchossoy, L. (2005). Pax3/Pax7 mark a novel population of primitive myogenic cells during development. *Genes Dev.* *19*, 1426–1431.
- Kassar-Duchossoy, L., Gayraud-Morel, B., and Tajbakhsh, S. (2004). Mrf4 determines skeletal muscle identity in Myf5:Myod double-mutant mice. *Nature* *431*, 461–466.
- Katzemich, A., Kreiskother, N., Alexandrovich, A., Elliott, C., Schock, F., Leonard, K., Sparrow, J., and Bullard, B. (2012). The function of the M-line protein obscurin in controlling the symmetry of the sarcomere in the flight muscle of *Drosophila*. *J. Cell Sci.* *125*, 4170–4170.
- Keleman, K., and Dickson, B.J. (2001). Short- and long-range repulsion by the *Drosophila* Unc5 netrin receptor. *Neuron* *32*, 605–617.
- Kelly, R.G. (2004). The del22q11.2 candidate gene Tbx1 regulates branchiomeric myogenesis. *Hum. Mol. Genet.* *13*, 2829–2840.
- Kiefer, J.C., and Hauschka, S.D. (2001). Myf-5 Is Transiently Expressed in Nonmuscle Mesoderm and Exhibits Dynamic Regional Changes within the Presegmented Mesoderm and Somites I–IV. *Dev. Biol.* *232*, 77–90.
- Kim, C., Ye, F., Hu, X., and Ginsberg, M.H. (2012). Talin activates integrins by altering the topology of the  $\beta$  transmembrane domain. *J. Cell Biol.* *197*, 605–611.
- Kim, J.H., Jin, P., Duan, R., and Chen, E.H. (2015a). Mechanisms of myoblast fusion during muscle development. *Curr. Opin. Genet. Dev.* *32*, 162–170.
- Kim, J.H., Ren, Y., Ng, W.P., Li, S., Son, S., Kee, Y.-S., Zhang, S., Zhang, G., Fletcher, D.A., Robinson, D.N., et al. (2015b). Mechanical Tension Drives Cell Membrane Fusion. *Dev. Cell* *32*, 561–573.
- Kim, S., Shilagardi, K., Zhang, S., Hong, S.N., Sens, K.L., Bo, J., Gonzalez, G.A., and Chen, E.H. (2007). A Critical Function for the Actin Cytoskeleton in Targeted Exocytosis of Prefusion Vesicles during Myoblast Fusion. *Dev. Cell* *12*, 571–586.
- Kirby, R.J., Hamilton, G.M., Finnegan, D.J., Johnson, K.J., and Jarman, A.P. (2001). *Drosophila* homolog of the myotonic dystrophy-associated gene, SIX5, is required for muscle and gonad development. *Curr. Biol.* *11*, 1044–1049.
- Knirr, S., and Frasch, M. (2001). Molecular Integration of Inductive and Mesoderm-Intrinsic Inputs Governs even-skipped Enhancer Activity in a Subset of Pericardial and Dorsal Muscle Progenitors. *Dev. Biol.* *238*, 13–26.
- Knirr, S., Azpiazu, N., and Frasch, M. (1999). The role of the NK-homeobox gene slouch (S59) in somatic muscle patterning. *Development* *126*, 4525–4535.
- Kolodziej, P.A., Timpe, L.C., Mitchell, K.J., Fried, S.R., Goodman, C.S., Jan, L.Y., and Jan, Y.N. (1996). frazzled encodes a *Drosophila* member of the DCC immunoglobulin subfamily and is required for CNS and motor axon guidance. *Cell* *87*, 197–204.

- Kong, P., Racedo, S.E., Macchiarulo, S., Hu, Z., Carpenter, C., Guo, T., Wang, T., Zheng, D., and Morrow, B.E. (2014). Tbx1 is required autonomously for cell survival and fate in the pharyngeal core mesoderm to form the muscles of mastication. *Hum. Mol. Genet.* **23**, 4215–4231.
- Kuroda, K., Tani, S., Tamura, K., Minoguchi, S., Kurooka, H., and Honjo, T. (1999). Delta-induced Notch signaling mediated by RBP-J inhibits MyoD expression and myogenesis. *J. Biol. Chem.* **274**, 7238–7244.
- Kurusu, M., Cording, A., Taniguchi, M., Menon, K., Suzuki, E., and Zinn, K. (2008). A Screen of Cell-Surface Molecules Identifies Leucine-Rich Repeat Proteins as Key Mediators of Synaptic Target Selection. *Neuron* **59**, 972–985.
- Labrador, J.P., O’Keefe, D., Yoshikawa, S., McKinnon, R.D., Thomas, J.B., and Bashaw, G.J. (2005). The Homeobox Transcription Factor Even-skipped Regulates Netrin-Receptor Expression to Control Dorsal Motor-Axon Projections in *Drosophila*. *Curr. Biol.* **15**, 1413–1419.
- Lagha, M., Kormish, J.D., Rocancourt, D., Manceau, M., Epstein, J.A., Zaret, K.S., Relaix, F., and Buckingham, M.E. (2008). Pax3 regulation of FGF signaling affects the progression of embryonic progenitor cells into the myogenic program. *Genes Dev.* **22**, 1828–1837.
- Lagha, M., Sato, T., Regnault, B., Cumano, A., Zuniga, A., Licht, J., Relaix, F., and Buckingham, M. (2010). Transcriptome analyses based on genetic screens for Pax3 myogenic targets in the mouse embryo. *BMC Genomics* **11**, 696.
- Lahaye, L.L., Wouda, R.R., de Jong, A.W.M., Fradkin, L.G., and Noordermeer, J.N. (2012). WNT5 Interacts with the Ryk Receptors Doughnut and Derailed to Mediate Muscle Attachment Site Selection in *Drosophila melanogaster*. *PLoS ONE* **7**, e32297.
- Lampa, S.J., Potluri, S., Norton, A.S., Fusco, W., and Laskowski, M.B. (2004). Ephrin-A5 overexpression degrades topographic specificity in the mouse gluteus maximus muscle. *Dev. Brain Res.* **153**, 271–274.
- Landgraf, M., Bossing, T., Technau, G.M., and Bate, M. (1997). The origin, location, and projections of the embryonic abdominal motoneurons of *Drosophila*. *J. Neurosci.* **17**, 9642–9655.
- Landgraf, M., Baylies, M., and Bate, M. (1999). Muscle founder cells regulate defasciculation and targeting of motor axons in the *Drosophila* embryo. *Curr. Biol.* **9**, 589–592.
- Landmesser, L. (1978). The distribution of motoneurons supplying chick hind limb muscles. *J. Physiol.* **284**, 371.
- Landmesser, L.T. (2001). The acquisition of motoneuron subtype identity and motor circuit formation. *Int. J. Dev. Neurosci.* **19**, 175–182.
- Lang, D., Powell, S.K., Plummer, R.S., Young, K.P., and Ruggeri, B.A. (2007). PAX genes: Roles in development, pathophysiology, and cancer. *Biochem. Pharmacol.* **73**, 1–14.
- Lange, S., Ehler, E., and Gautel, M. (2006). From A to Z and back? Multicompartment proteins in the sarcomere. *Trends Cell Biol.* **16**, 11–18.
- Laurin, M., Fradet, N., Blangy, A., Hall, A., Vuori, K., and Côté, J.-F. (2008). The atypical Rac activator Dock180 (Dock1) regulates myoblast fusion in vivo. *Proc. Natl. Acad. Sci.* **105**, 15446–15451.
- Le Grand, F., Grifone, R., Mourikis, P., Houbron, C., Gigaud, C., Pujol, J., Maillet, M., Pagès, G., Rudnicki, M., Tajbakhsh, S., et al. (2012). Six1 regulates stem cell repair potential and self-renewal during skeletal muscle regeneration. *J. Cell Biol.* **198**, 815–832.
- Lee, K.H., Lee, S., Kim, B., Chang, S., Kim, S.W., Paick, J.-S., and Rhee, K. (2006). Dazl can bind to dynein motor complex and may play a role in transport of specific mRNAs. *EMBO J.* **25**, 4263–4270.
- Lejard, V., Blais, F., Guerquin, M.-J., Bonnet, A., Bonnin, M.-A., Havis, E., Malbouyres, M., Bidaud, C.B., Maro, G., Gilardi-Hebenstreit, P., et al. (2011). EGR1 and EGR2 Involvement in Vertebrate Tendon Differentiation. *J. Biol. Chem.* **286**, 5855–5867.
- Leptin, M., and Grunewald, B. (1990). Cell shape changes during gastrulation in *Drosophila*. *Development* **110**, 73–84.
- L’Honoré, A., Ouimette, J.-F., Lavertu-Jolin, M., and Drouin, J. (2010). Pitx2 defines alternate pathways acting through MyoD during limb and somitic myogenesis. *Development* **137**, 3847–3856.
- Lindsell, C.E., Shawber, C.J., Boulter, J., and Weinmaster, G. (1995). Jagged: a mammalian ligand that activates Notch1. *Cell* **80**, 909–917.

- Liu, H., Zhang, C., Zhu, S., Lu, P., Zhu, T., Gong, X., Zhang, Z., Hu, J., Yin, Z., Heng, B.C., et al. (2015). Mohawk Promotes the Tenogenesis of Mesenchymal Stem Cells Through Activation of the TGF $\beta$  Signaling Pathway: Mxk Promotes the Tenogenesis of MSCs. *STEM CELLS* 33, 443–455.
- Liu, Y.-H., Jakobsen, J.S., Valentin, G., Amarantos, I., Gilmour, D.T., and Furlong, E.E.M. (2009). A Systematic Analysis of Tinman Function Reveals Eya and JAK-STAT Signaling as Essential Regulators of Muscle Development. *Dev. Cell* 16, 280–291.
- Lu, M., Witke, W., Kwiatkowski, D.J., and Kosik, K.S. (1997). Delayed retraction of filopodia in gelsolin null mice. *J. Cell Biol.* 138, 1279–1287.
- Luetjens, C.M., Xu, E.Y., Rejo Pera, R.A., Kamischke, A., Nieschlag, E., and Gromoll, J. (2004). Association of Meiotic Arrest with Lack of *BOULE* Protein Expression in Infertile Men. *J. Clin. Endocrinol. Metab.* 89, 1926–1933.
- Lum, L. (2004). The Hedgehog Response Network: Sensors, Switches, and Routers. *Science* 304, 1755–1759.
- Maeda, T., Sakabe, T., Sunaga, A., Sakai, K., Rivera, A.L., Keene, D.R., Sasaki, T., Stavnezer, E., Iannotti, J., Schweitzer, R., et al. (2011). Conversion of Mechanical Force into TGF- $\beta$ -Mediated Biochemical Signals. *Curr. Biol.* 21, 933–941.
- Maegawa, S., Yamashita, M., Yasuda, K., and Inoue, K. (2002). Zebrafish DAZ-like protein controls translation via the sequence ‘GUUC.’ *Genes Cells* 7, 971–984.
- Maines, J.Z., and Wasserman, S.A. (1999). Post-transcriptional regulation of the meiotic Cdc25 protein Twine by the Dazl orthologue Boule. *Nat. Cell Biol.* 1, 171–174.
- Malone, C.J., Fixsen, W.D., Horvitz, H.R., and Han, M. (1999). UNC-84 localizes to the nuclear envelope and is required for nuclear migration and anchoring during *C. elegans* development. *Development* 126, 3171–3181.
- Mann, T., Bodmer, R., and Pandur, P. (2009). The *Drosophila* homolog of vertebrate Islet1 is a key component in early cardiogenesis. *Development* 136, 317–326.
- Marcelle, C., Stark, M.R., and Bronner-Fraser, M. (1997). Coordinate actions of BMPs, Wnts, Shh and noggin mediate patterning of the dorsal somite. *Development* 124, 3955–3963.
- Massari, M.E., and Murre, C. (2000). Helix-loop-helix proteins: regulators of transcription in eucaryotic organisms. *Mol. Cell. Biol.* 20, 429–440.
- Massarwa, R., Carmon, S., Shilo, B.-Z., and Schejter, E.D. (2007). WIP/WASp-Based Actin-Polymerization Machinery Is Essential for Myoblast Fusion in *Drosophila*. *Dev. Cell* 12, 557–569.
- Matthes, D.J., Sink, H., Kolodkin, A.L., and Goodman, C.S. (1995). Semaphorin II can function as a selective inhibitor of specific synaptic arborizations. *Cell* 81, 631–639.
- McKinsey, T.A., Zhang, C.L., and Olson, E.N. (2002). Signaling chromatin to make muscle. *Curr. Opin. Cell Biol.* 14, 763–772.
- McMahon, J.A., Takada, S., Zimmerman, L.B., Fan, C.-M., Harland, R.M., and McMahon, A.P. (1998). Noggin-mediated antagonism of BMP signaling is required for growth and patterning of the neural tube and somite. *Genes Dev.* 12, 1438–1452.
- Menon, S.D., and Chia, W. (2001). *Drosophila* rolling pebbles: a multidomain protein required for myoblast fusion that recruits D-Titin in response to the myoblast attractant Dumbfounded. *Dev. Cell* 1, 691–703.
- Messina, G., Biressi, S., Monteverde, S., Magli, A., Cassano, M., Perani, L., Roncaglia, E., Tagliafico, E., Starnes, L., Campbell, C.E., et al. (2010). Nfix Regulates Fetal-Specific Transcription in Developing Skeletal Muscle. *Cell* 140, 554–566.
- Metzger, T., Gache, V., Xu, M., Cadot, B., Folker, E.S., Richardson, B.E., Gomes, E.R., and Baylies, M.K. (2012). MAP and kinesin-dependent nuclear positioning is required for skeletal muscle function. *Nature* 484, 120–124.
- Michelson, A.M., Gisselbrecht, S., Zhou, Y., Baek, K.-H., and Buff, E.M. (1998). Dual functions of the heartless fibroblast growth factor receptor in development of the *Drosophila* embryonic mesoderm. *Genesis* 22, 212–229.
- Millay, D.P., O’Rourke, J.R., Sutherland, L.B., Bezprozvannaya, S., Shelton, J.M., Bassel-Duby, R., and Olson, E.N. (2013). Myomaker is a membrane activator of myoblast fusion and muscle formation.

- Nature 499, 301–305.
- Miller, M.R., Robinson, K.J., Cleary, M.D., and Doe, C.Q. (2009). TU-tagging: cell type-specific RNA isolation from intact complex tissues. *Nat. Methods* 6, 439–441.
- Miner, J.H., and Wold, B. (1990). Herculin, a fourth member of the MyoD family of myogenic regulatory genes. *Proc. Natl. Acad. Sci.* 87, 1089–1093.
- Mitchell, K.J., Doyle, J.L., Serafini, T., Kennedy, T.E., Tessier-Lavigne, M., Goodman, C.S., and Dickson, B.J. (1996). Genetic analysis of Netrin genes in *Drosophila*: Netrins guide CNS commissural axons and peripheral motor axons. *Neuron* 17, 203–215.
- Miyazono, K., Maeda, S., and Imamura, T. (2005). BMP receptor signaling: Transcriptional targets, regulation of signals, and signaling cross-talk. *Cytokine Growth Factor Rev.* 16, 251–263.
- Molkentin, J.D., Black, B.L., Martin, J.F., and Olson, E.N. (1995). Cooperative activation of muscle gene expression by MEF2 and myogenic bHLH proteins. *Cell* 83, 1125–1136.
- Moncaut, N., Cross, J.W., Siligan, C., Keith, A., Taylor, K., Rigby, P.W.J., and Carvajal, J.J. (2012). Musculin and TCF21 coordinate the maintenance of myogenic regulatory factor expression levels during mouse craniofacial development. *Development* 139, 958–967.
- Moore, F.L., Jaruzelska, J., Fox, M.S., Urano, J., Firpo, M.T., Turek, P.J., Dorfman, D.M., and Pera, R.A.R. (2003). Human Pumilio-2 is expressed in embryonic stem cells and germ cells and interacts with DAZ (Deleted in AZoospermia) and DAZ-like proteins. *Proc. Natl. Acad. Sci.* 100, 538–543.
- Morimoto, T., Nobechi, M., Komatsu, A., Miyakawa, H., and Nose, A. (2010). Subunit-specific and homeostatic regulation of glutamate receptor localization by CaMKII in *Drosophila* neuromuscular junctions. *Neuroscience* 165, 1284–1292.
- Morton, S., Yang, H.-T., Moleleki, N., Campbell, D., Cohen, P., and Rousseau, S. (2006). Phosphorylation of the ARE-binding protein DAZAP1 by ERK2 induces its dissociation from DAZ. *Biochem. J.* 399, 265–273.
- Mounier, N., and Arrigo, A.-P. (2002). Actin cytoskeleton and small heat shock proteins: how do they interact. *Cell Stress Chaperones* 7, 167–176.
- Müller, D., Jagla, T., Bodart, L.M., Jährling, N., Dodt, H.-U., Jagla, K., and Frasch, M. (2010). Regulation and Functions of the *lms* Homeobox Gene during Development of Embryonic Lateral Transverse Muscles and Direct Flight Muscles in *Drosophila*. *PLoS ONE* 5, e14323.
- Nabel-Rosen, H., Volohonsky, G., Reuveny, A., Zaidel-Bar, R., and Volk, T. (2002). Two isoforms of the *Drosophila* RNA binding protein, how, act in opposing directions to regulate tendon cell differentiation. *Dev. Cell* 2, 183–193.
- Nachman, A., Halachmi, N., Matia, N., Manzur, D., and Salzberg, A. (2015). Deconstructing the complexity of regulating common properties in different cell types: Lessons from the *delilah* gene. *Dev. Biol.* 403, 180–191.
- Nag, S., Larsson, M., arten, Robinson, R.C., and Burtnick, L.D. (2013). Gelsolin: the tail of a molecular gymnast. *Cytoskeleton* 70, 360–384.
- Nathan, E., Monovich, A., Tirosh-Finkel, L., Harrelson, Z., Rousso, T., Rinon, A., Harel, I., Evans, S.M., and Tzahor, E. (2008). The contribution of *Islet1*-expressing splanchnic mesoderm cells to distinct branchiomic muscles reveals significant heterogeneity in head muscle development. *Development* 135, 647–657.
- Niro, C., Demignon, J., Vincent, S., Liu, Y., Giordani, J., Sgarlato, N., Favier, M., Guillet-Deniau, I., Blais, A., and Maire, P. (2010). *Six1* and *Six4* gene expression is necessary to activate the fast-type muscle gene program in the mouse primary myotome. *Dev. Biol.* 338, 168–182.
- Noden, D.M., and Francis-West, P. (2006). The differentiation and morphogenesis of craniofacial muscles. *Dev. Dyn.* 235, 1194–1218.
- Noden, D.M., and Trainor, P.A. (2005). Relations and interactions between cranial mesoderm and neural crest populations. *J. Anat.* 207, 575–601.
- Nose, A., Umeda, T., and Takeichi, M. (1997). Neuromuscular target recognition by a homophilic interaction of connectin cell adhesion molecules in *Drosophila*. *Development* 124, 1433–1441.
- Nose, A., Isshiki, T., and Takeichi, M. (1998). Regional specification of muscle progenitors in *Drosophila*: the role of the *msh* homeobox gene. *Development* 125, 215–223.

- Nowak, S.J., Aihara, H., Gonzalez, K., Nibu, Y., and Baylies, M.K. (2012). Akirin Links Twist-Regulated Transcription with the Brahma Chromatin Remodeling Complex during Embryogenesis. *PLoS Genet.* *8*, e1002547.
- Noyes, M.B., Christensen, R.G., Wakabayashi, A., Stormo, G.D., Brodsky, M.H., and Wolfe, S.A. (2008). Analysis of Homeodomain Specificities Allows the Family-wide Prediction of Preferred Recognition Sites. *Cell* *133*, 1277–1289.
- O'Connor-Giles, K.M., and Skeath, J.B. (2003). Numb inhibits membrane localization of Sanpodo, a four-pass transmembrane protein, to promote asymmetric divisions in *Drosophila*. *Dev. Cell* *5*, 231–243.
- O'Keefe, D.D., Edgar, B.A., and Saucedo, L.J. (2011). EndoG1 modulates Notch signaling and axon guidance in *Drosophila*. *Mech. Dev.* *128*, 59–70.
- Otabe, K., Nakahara, H., Hasegawa, A., Matsukawa, T., Ayabe, F., Onizuka, N., Inui, M., Takada, S., Ito, Y., Sekiya, I., et al. (2015). Transcription factor Mohawk controls tenogenic differentiation of bone marrow mesenchymal stem cells in vitro and in vivo: THE TRANSCRIPTION FACTOR MOHAWK CONTROLS TENOGENIC. *J. Orthop. Res.* *33*, 1–8.
- Padrick, S.B., Doolittle, L.K., Brautigam, C.A., King, D.S., and Rosen, M.K. (2011). Arp2/3 complex is bound and activated by two WASP proteins. *Proc. Natl. Acad. Sci.* *108*, E472–E479.
- Papakonstanti, E.A., and Stournaras, C. (2008). Cell responses regulated by early reorganization of actin cytoskeleton. *FEBS Lett.* *582*, 2120–2127.
- Park, M., Yaich, L.E., and Bodmer, R. (1998). Mesodermal cell fate decisions in *Drosophila* are under the control of the lineage genes numb, Notch, and sanpodo. *Mech. Dev.* *75*, 117–126.
- Parker, M.H., Seale, P., and Rudnicki, M.A. (2003). Looking back to the embryo: defining transcriptional networks in adult myogenesis. *Nat. Rev. Genet.* *4*, 497–507.
- Parker, M.H., Perry, R.L.S., Fauteux, M.C., Berkes, C.A., and Rudnicki, M.A. (2006). MyoD Synergizes with the E-Protein HEB To Induce Myogenic Differentiation. *Mol. Cell. Biol.* *26*, 5771–5783.
- Parr, B.A., Shea, M.J., Vassileva, G., and McMahon, A.P. (1993). Mouse Wnt genes exhibit discrete domains of expression in the early embryonic CNS and limb buds. *Development* *119*, 247–261.
- Parsons, M.J., Pollard, S.M., Feldman, B., Coutinho, P., Hirst, E.M., Stemple, D.L., and others (2002). Zebrafish mutants identify an essential role for laminins in notochord formation. *Development* *129*, 3137–3146.
- Penn, B.H., Bergstrom, D.A., Dilworth, F.J., Bengal, E., and Tapscott, S.J. (2004). A MyoD-generated feed-forward circuit temporally patterns gene expression during skeletal muscle differentiation. *Genes Dev.* *18*, 2348–2353.
- Philippakis, A.A., Busser, B.W., Gisselbrecht, S.S., He, F.S., Estrada, B., Michelson, A.M., and Bulyk, M.L. (2006). Expression-Guided In Silico Evaluation of Candidate Cis Regulatory Codes for *Drosophila* Muscle Founder Cells. *PLoS Comput. Biol.* *2*, e53.
- Pignoni, F., Hu, B., Zavitz, K.H., Xiao, J., Garrity, P.A., and Zipursky, S.L. (1997). The eye-specification proteins So and Eya form a complex and regulate multiple steps in *Drosophila* eye development. *Cell* *91*, 881–891.
- Pollard, T.D. (2016). Actin and Actin-Binding Proteins. *Cold Spring Harb. Perspect. Biol.* *8*, a018226.
- Port, F., Chen, H.-M., Lee, T., and Bullock, S.L. (2014). Optimized CRISPR/Cas tools for efficient germline and somatic genome engineering in *Drosophila*. *Proc. Natl. Acad. Sci.* *111*, E2967–E2976.
- Potthoff, M.J., and Olson, E.N. (2007). MEF2: a central regulator of diverse developmental programs. *Development* *134*, 4131–4140.
- Potthoff, M.J., Arnold, M.A., McAnally, J., Richardson, J.A., Bassel-Duby, R., and Olson, E.N. (2007). Regulation of Skeletal Muscle Sarcomere Integrity and Postnatal Muscle Function by Mef2c. *Mol. Cell. Biol.* *27*, 8143–8151.
- Pourquié, O., and Tam, P.P. (2001). A nomenclature for prospective somites and phases of cyclic gene expression in the presomitic mesoderm. *Dev. Cell* *1*, 619–620.
- Pourquié, O., Coltey, M., Bréant, C., and Le Douarin, N.M. (1995). Control of somite patterning by signals from the lateral plate. *Proc. Natl. Acad. Sci.* *92*, 3219–3223.
- Prokop, A. (2006). Organization of the Efferent System and Structure of Neuromuscular Junctions In

- Drosophila*. In International Review of Neurobiology, (Elsevier), pp. 71–90.
- Prokop, A., Landgraf, M., Rushton, E., Brodie, K., and Bate, M. (1996). Presynaptic development at the *Drosophila* neuromuscular junction: assembly and localization of presynaptic active zones. *Neuron* *17*, 617–626.
- Proserpio, V., Fittipaldi, R., Ryall, J.G., Sartorelli, V., and Caretti, G. (2013). The methyltransferase SMYD3 mediates the recruitment of transcriptional cofactors at the myostatin and c-Met genes and regulates skeletal muscle atrophy. *Genes Dev.* *27*, 1299–1312.
- Relaix, F., Rocancourt, D., Mansouri, A., and Buckingham, M. (2004). Divergent functions of murine Pax3 and Pax7 in limb muscle development. *Genes Dev.* *18*, 1088–1105.
- Relaix, F., Rocancourt, D., Mansouri, A., and Buckingham, M. (2005). A Pax3/Pax7-dependent population of skeletal muscle progenitor cells. *Nature* *435*, 948–953.
- Relaix, F., Demignon, J., Laclef, C., Pujol, J., Santolini, M., Niro, C., Lagha, M., Rocancourt, D., Buckingham, M., and Maire, P. (2013). Six Homeoproteins Directly Activate Myod Expression in the Gene Regulatory Networks That Control Early Myogenesis. *PLoS Genet.* *9*, e1003425.
- Ren, G., Crampton, M.S., and Yap, A.S. (2009). Cortactin: Coordinating adhesion and the actin cytoskeleton at cellular protrusions. *Cell Motil. Cytoskeleton* *66*, 865–873.
- Reuter, R., and Leptin, M. (1994). Interacting functions of snail, twist and huckebein during the early development of germ layers in *Drosophila*. *Development* *120*, 1137–1150.
- Reynolds, N. (2005). Dazl binds in vivo to specific transcripts and can regulate the pre-meiotic translation of Mvh in germ cells. *Hum. Mol. Genet.* *14*, 3899–3909.
- Reynolds, N., Collier, B., Bingham, V., Gray, N.K., and Cooke, H.J. (2007). Translation of the synaptonemal complex component Sycp3 is enhanced in vivo by the germ cell specific regulator Dazl. *RNA* *13*, 974–981.
- Richardson, B.E., Beckett, K., Nowak, S.J., and Baylies, M.K. (2007). SCAR/WAVE and Arp2/3 are crucial for cytoskeletal remodeling at the site of myoblast fusion. *Development* *134*, 4357–4367.
- Richardson, B.E., Nowak, S.J., and Baylies, M.K. (2008). Myoblast Fusion in Fly and Vertebrates: New Genes, New Processes and New Perspectives. *Traffic* *9*, 1050–1059.
- Riechmann, V., Irion, U., Wilson, R., Grosskortenhaus, R., and Leptin, M. (1997). Control of cell fates and segmentation in the *Drosophila* mesoderm. *Development* *124*, 2915–2922.
- Ritzenthaler, S., and Chiba, A. (2003). Myopodia (postsynaptic filopodia) participate in synaptic target recognition. *J. Neurobiol.* *55*, 31–40.
- Ritzenthaler, S., Suzuki, E., and Chiba, A. (2000). Postsynaptic filopodia in muscle cells interact with innervating motoneuron axons. *Nat. Neurosci.* *3*, 1012–1017.
- Romero, N.B. (2010). Centronuclear myopathies: A widening concept. *Neuromuscul. Disord.* *20*, 223–228.
- Rose, D., Zhu, X., Kose, H., Hoang, B., Cho, J., and 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.
- Ruiz, M.G., and Bate, M. (1997). Segregation of myogenic lineages in *Drosophila* requires numb. *Development* *124*, 4857–4866.
- 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. In International Review of Neurobiology, (Elsevier), pp. 1–31.
- Ruiz-Gómez, M., Romani, S., Hartmann, C., Jackle, H., and Bate, M. (1997). Specific muscle identities are regulated by Kruppel during *Drosophila* embryogenesis. *Development* *124*, 3407–3414.
- Ruiz-Gómez, M., Coutts, N., Price, A., Taylor, M.V., and Bate, M. (2000). *Drosophila* dumbfounded: a myoblast attractant essential for fusion. *Cell* *102*, 189–198.
- Ruiz-Gómez, M., Coutts, N., Suster, M.L., Landgraf, M., and Bate, M. (2002). myoblasts incompetent encodes a zinc finger transcription factor required to specify fusion-competent myoblasts in *Drosophila*. *Development* *129*, 133–141.
- Sambasivan, R., Gayraud-Morel, B., Dumas, G., Cimper, C., Paisant, S., Kelly, R.G., and Tajbakhsh, S.



- (2009). Distinct Regulatory Cascades Govern Extraocular and Pharyngeal Arch Muscle Progenitor Cell Fates. *Dev. Cell* 16, 810–821.
- Sambasivan, R., Kuratani, S., and Tajbakhsh, S. (2011). An eye on the head: the development and evolution of craniofacial muscles. *Development* 138, 2401–2415.
- Sandmann, T., Jensen, L.J., Jakobsen, J.S., Karzynski, M.M., Eichenlaub, M.P., Bork, P., and Furlong, E.E.M. (2006). A Temporal Map of Transcription Factor Activity: Mef2 Directly Regulates Target Genes at All Stages of Muscle Development. *Dev. Cell* 10, 797–807.
- Sandmann, T., Girardot, C., Brehme, M., Tongprasit, W., Stolc, V., and Furlong, E.E.M. (2007). A core transcriptional network for early mesoderm development in *Drosophila melanogaster*. *Genes Dev.* 21, 436–449.
- Sanes, J.R., and Lichtman, J.W. (1999). Development of the vertebrate neuromuscular junction. *Annu. Rev. Neurosci.* 22, 389–442.
- Sanes, J.R., and Yamagata, M. (2009). Many Paths to Synaptic Specificity. *Annu. Rev. Cell Dev. Biol.* 25, 161–195.
- Santiago, C., and Bashaw, G.J. (2017). Islet Coordinately Regulates Motor Axon Guidance and Dendrite Targeting through the Frazzled/DCC Receptor. *Cell Rep.* 18, 1646–1659.
- Sasai, Y., Kageyama, R., Tagawa, Y., Shigemoto, R., and Nakanishi, S. (1992). Two mammalian helix-loop-helix factors structurally related to *Drosophila* hairy and Enhancer of split. *Genes Dev.* 6, 2620–2634.
- Sato, T., Rocancourt, D., Marques, L., Thorsteinsdóttir, S., and Buckingham, M. (2010). A Pax3/Dmrt2/Myf5 Regulatory Cascade Functions at the Onset of Myogenesis. *PLoS Genet.* 6, e1000897.
- Schäfer, K., and Braun, T. (1999). Early specification of limb muscle precursor cells by the homeobox gene *Lbx1h*. *Nat. Genet.* 23, 213–216.
- Schäfer, G., Weber, S., Holz, A., Bogdan, S., Schumacher, S., Müller, A., Renkawitz-Pohl, R., and Önel, S.-F. (2007). The Wiskott–Aldrich syndrome protein (WASP) is essential for myoblast fusion in *Drosophila*. *Dev. Biol.* 304, 664–674.
- Schaub, C., Nagaso, H., Jin, H., and Frasch, M. (2012). *Org-1*, the *Drosophila* ortholog of *Tbx1*, is a direct activator of known identity genes during muscle specification. *Development* 139, 1001–1012.
- von Scheven, G. (2006). Neural tube derived signals and *Fgf8* act antagonistically to specify eye versus mandibular arch muscles. *Development* 133, 2731–2745.
- Schienda, J., Engleka, K.A., Jun, S., Hansen, M.S., Epstein, J.A., Tabin, C.J., Kunkel, L.M., and Kardon, G. (2006). Somitic origin of limb muscle satellite and side population cells. *Proc. Natl. Acad. Sci. U. S. A.* 103, 945–950.
- Schnorrer, F., and Dickson, B.J. (2004). Muscle building: mechanisms of myotube guidance and attachment site selection. *Dev. Cell* 7, 9–20.
- Schnorrer, F., Kalchhauser, I., and Dickson, B.J. (2007). The Transmembrane Protein Kon-tiki Couples to Dgrip to Mediate Myotube Targeting in *Drosophila*. *Dev. Cell* 12, 751–766.
- Schroter, R.H. (2004). *kette* and *blown fuse* interact genetically during the second fusion step of myogenesis in *Drosophila*. *Development* 131, 4501–4509.
- Schulman, V.K., Folker, E.S., Rosen, J.N., and Baylies, M.K. (2014). *Syd/JIP3* and *JNK* Signaling Are Required for Myonuclear Positioning and Muscle Function. *PLoS Genet.* 10, e1004880.
- Schuster-Gossler, K., Cordes, R., and Gossler, A. (2007). Premature myogenic differentiation and depletion of progenitor cells cause severe muscle hypotrophy in *Delta1* mutants. *Proc. Natl. Acad. Sci.* 104, 537–542.
- Schweitzer, R., Chyung, J.H., Murtaugh, L.C., Brent, A.E., Rosen, V., Olson, E.N., Lassar, A., and Tabin, C.J. (2001). Analysis of the tendon cell fate using *Scleraxis*, a specific marker for tendons and ligaments. *Development* 128, 3855–3866.
- Schweitzer, R., Zelzer, E., and Volk, T. (2010). Connecting muscles to tendons: tendons and musculoskeletal development in flies and vertebrates. *Development* 137, 3347–3347.
- Seale, P., Bjork, B., Yang, W., Kajimura, S., Chin, S., Kuang, S., Scimè, A., Devarakonda, S., Conroe, H.M., Erdjument-Bromage, H., et al. (2008). *PRDM16* controls a brown fat/skeletal muscle switch.

Nature 454, 961–967.

Sellin, L. (2002). NEPH1 defines a novel family of podocin-interacting proteins. FASEB J.

Sens, K.L., Zhang, S., Jin, P., Duan, R., Zhang, G., Luo, F., Parachini, L., and Chen, E.H. (2010). An invasive podosome-like structure promotes fusion pore formation during myoblast fusion. J. Cell Biol. 191, 1013–1027.

Seo, H.-C., Curtiss, J., Mlodzik, M., and Fjose, A. (1999). Six class homeobox genes in Drosophila belong to three distinct families and are involved in head development. Mech. Dev. 83, 127–139.

Shah, C., VanGompel, M.J.W., Naeem, V., Chen, Y., Lee, T., Angeloni, N., Wang, Y., and Xu, E.Y. (2010). Widespread Presence of Human BOULE Homologs among Animals and Conservation of Their Ancient Reproductive Function. PLoS Genet. 6, e1001022.

Shih, H.P., Gross, M.K., and Kiousi, C. (2007). Cranial muscle defects of Pitx2 mutants result from specification defects in the first branchial arch. Proc. Natl. Acad. Sci. 104, 5907–5912.

Shih, H.P., Gross, M.K., and Kiousi, C. (2008). Muscle development: Forming the head and trunk muscles. Acta Histochem. 110, 97–108.

Shilagardi, K., Li, S., Fengbao, L., Maikar, F., Duan, R., Jin, peng, Hoon Kim, J., Murmen, K., and Chen, E.H. (2013). Actin-Propelled Invasive Membrane Protrusions Promote Fusogenic Protein Engagement During Cell-Cell Fusion. Science 340, 356–359.

Shishido, E., Ono, N., Kojima, T., and Saigo, K. (1997). Requirements of DFR1/Heartless, a mesoderm-specific Drosophila FGF-receptor, for the formation of heart, visceral and somatic muscles, and ensheathing of longitudinal axon tracts in CNS. Development 124, 2119–2128.

Shishido, E., Takeichi, M., and Nose, C.-Y. (1998). Drosophila Synapse Formation: Regulation by Transmembrane Protein with Leu-Rich Repeats, CAPRICIOUS. Science 280, 2114–2118.

Silacci, P., Mazzolai, L., Gauci, C., Stergiopoulos, N., Yin, H.L., and Hayoz, D. (2004). Gelsolin superfamily proteins: key regulators of cellular functions. Cell. Mol. Life Sci. 61, 2614–2623.

Singh, K., and Dilworth, F.J. (2013). Differential modulation of cell cycle progression distinguishes members of the myogenic regulatory factor family of transcription factors. FEBS J. 280, 3991–4003.

Sink, H., and Whittington, P.M. (1991). Pathfinding in the central nervous system and periphery by identified embryonic Drosophila motor axons. Development 112, 307–316.

Snow, C.J., Peterson, M.T., Khalil, A., and Henry, C.A. (2008). Muscle development is disrupted in zebrafish embryos deficient for fibronectin. Dev. Dyn. 237, 2542–2553.

Sohn, R.L., Huang, P., Kawahara, G., Mitchell, M., Guyon, J., Kalluri, R., Kunkel, L.M., and Gussoni, E. (2009). A role for nephrin, a renal protein, in vertebrate skeletal muscle cell fusion. Proc. Natl. Acad. Sci. 106, 9274–9279.

Sparrow, J.C., and Schöck, F. (2009). The initial steps of myofibril assembly: integrins pave the way. Nat. Rev. Mol. Cell Biol. 10, 293–298.

St Johnston, D. (2005). Developmental Cell Biology: Moving messages: the intracellular localization of mRNAs. Nat. Rev. Mol. Cell Biol. 6, 363–375.

Starr, D., and Becker, M. (2002). Role of ANC-1 in Tethering Nuclei to the Actin Cytoskeleton. Science 298, 402–406.

Stathopoulos, A., Tam, B., Ronshaugen, M., Frasch, M., and Levine, M. (2004). pyramus and thisbe: FGF genes that pattern the mesoderm of Drosophila embryos. Genes Dev. 18, 687–699.

Steiner, F.A., Talbert, P.B., Kasinathan, S., Deal, R.B., and Henikoff, S. (2012). Cell-type-specific nuclei purification from whole animals for genome-wide expression and chromatin profiling. Genome Res. 22, 766–777.

Strünkelberg, M., Bonengel, B., Moda, L.M., Hertenstein, A., de Couet, H.G., Ramos, R.G., and Fischbach, K.-F. (2001). rst and its paralogue kirre act redundantly during embryonic muscle development in Drosophila. Development 128, 4229–4239.

Subramanian, A., and Schilling, T.F. (2014). Thrombospondin-4 controls matrix assembly during development and repair of myotendinous junctions. Elife 3, e02372.

Subramanian, A., Prokop, A., Yamamoto, M., Sugimura, K., Uemura, T., Betschinger, J., Knoblich, J.A., and Volk, T. (2003). Shortstop recruits EB1/APC1 and promotes microtubule assembly at the muscle-tendon junction. Curr. Biol. 13, 1086–1095.

- Subramanian, A., Wayburn, B., Bunch, T., and Volk, T. (2007). Thrombospondin-mediated adhesion is essential for the formation of the myotendinous junction in *Drosophila*. *Development* 134, 1269–1278.
- Sugiyama, Y., Suzuki, A., Kishikawa, M., Akutsu, R., Hirose, T., Waye, M.M.Y., Tsui, S.K.W., Yoshida, S., and Ohno, S. (2000). Muscle Develops a Specific Form of Small Heat Shock Protein Complex Composed of MKBP/HSPB2 and HSPB3 during Myogenic Differentiation. *J. Biol. Chem.* 275, 1095–1104.
- Sun, J., Kamei, C.N., Layne, M.D., Jain, M.K., Liao, J.K., Lee, M.-E., and Chin, M.T. (2001). Regulation of Myogenic Terminal Differentiation by the Hairy-related Transcription Factor CHF2. *J. Biol. Chem.* 276, 18591–18596.
- Swan, L.E. (2004). A Glutamate Receptor-Interacting Protein homolog organizes muscle guidance in *Drosophila*. *Genes Dev.* 18, 223–237.
- Swan, L.E., Schmidt, M., Schwarz, T., Ponimaskin, E., Prange, U., Boeckers, T., Thomas, U., and Sigrist, S.J. (2006). Complex interaction of *Drosophila* GRIP PDZ domains and Echinoid during muscle morphogenesis. *EMBO J.* 25, 3640–3651.
- Tajbakhsh, S., Bober, E., Babinet, C., Pournin, S., Arnold, H., and Buckingham, M. (1996). Gene targeting the *myf-5* locus with *n lacZ* reveals expression of this myogenic factor in mature skeletal muscle fibres as well as early embryonic muscle. *Dev. Dyn.* 206, 291–300.
- Tajbakhsh, S., Rocancourt, D., Cossu, G., and Buckingham, M. (1997). Redefining the genetic hierarchies controlling skeletal myogenesis: Pax-3 and Myf-5 act upstream of MyoD. *Cell* 89, 127–138.
- Tapscott, S.J. (2005). The circuitry of a master switch: MyoD and the regulation of skeletal muscle gene transcription. *Development* 132, 2685–2695.
- Tessmar, K., Loosli, F., and Wittbrodt, J. (2002). A screen for co-factors of Six3. *Mech. Dev.* 117, 103–113.
- Thomas, A., Lee, P.-J., Dalton, J.E., Nomie, K.J., Stoica, L., Costa-Mattioli, M., Chang, P., Nuzhdin, S., Arbeitman, M.N., and Dierick, H.A. (2012). A Versatile Method for Cell-Specific Profiling of Translated mRNAs in *Drosophila*. *PLoS ONE* 7, e40276.
- Thor, S., and Thomas, J.B. (1997). The *Drosophila* *islet* gene governs axon pathfinding and neurotransmitter identity. *Neuron* 18, 397–409.
- Trinick, J. (1996). Interaction of titin/connectin with the thick filament. *Adv. Biophys.* 33, 81–90.
- Tripodi, M., and Arber, S. (2012). Regulation of motor circuit assembly by spatial and temporal mechanisms. *Curr. Opin. Neurobiol.* 22, 615–623.
- Tryon, R.C., Pisat, N., Johnson, S.L., and Dougherty, J.D. (2013). Development of translating ribosome affinity purification for zebrafish: Zebrafish Translating Ribosome Purification. *Genesis* 51, 187–192.
- Tsumaki, N., and Yoshikawa, H. (2005). The role of bone morphogenetic proteins in endochondral bone formation. *Cytokine Growth Factor Rev.* 16, 279–285.
- Tzahor, E. (2003). Antagonists of Wnt and BMP signaling promote the formation of vertebrate head muscle. *Genes Dev.* 17, 3087–3099.
- Tzahor, E. (2009). Heart and craniofacial muscle development: A new developmental theme of distinct myogenic fields. *Dev. Biol.* 327, 273–279.
- Ueno, H., Sakita-Ishikawa, M., Morikawa, Y., Nakano, T., Kitamura, T., and Saito, M. (2003). A stromal cell-derived membrane protein that supports hematopoietic stem cells. *Nat. Immunol.* 4, 457–463.
- Urano, J., Fox, M.S., and Reijo Pera, R.A. (2005). Interaction of the conserved meiotic regulators, BOULE (BOL) and PUMILIO-2 (PUM2). *Mol. Reprod. Dev.* 71, 290–298.
- VanGompel, M.J.W., and Xu, E.Y. (2010). A novel requirement in mammalian spermatid differentiation for the DAZ-family protein Boule. *Hum. Mol. Genet.* 19, 2360–2369.
- Vasyutina, E., and Birchmeier, C. (2006). The development of migrating muscle precursor cells. *Brain Struct. Funct.* 211, 37–41.
- Vasyutina, E., Stebler, J., Brand-Saberi, B., Schulz, S., Raz, E., and Birchmeier, C. (2005). CXCR4 and Gab1 cooperate to control the development of migrating muscle progenitor cells. *Genes Dev.* 19, 2187–2198.

- Vasyutina, E., Lenhard, D.C., Wende, H., Erdmann, B., Epstein, J.A., and Birchmeier, C. (2007). RBP-J (Rbpsi) is essential to maintain muscle progenitor cells and to generate satellite cells. *Proc. Natl. Acad. Sci.* *104*, 4443–4448.
- Vasyutina, E., Martarelli, B., Brakebusch, C., Wende, H., and Birchmeier, C. (2009). The small G-proteins Rac1 and Cdc42 are essential for myoblast fusion in the mouse. *Proc. Natl. Acad. Sci.* *106*, 8935–8940.
- Vaz, R., Martins, G.G., Thorsteinsdóttir, S., and Rodrigues, G. (2012). Fibronectin promotes migration, alignment and fusion in an in vitro myoblast cell model. *Cell Tissue Res.* *348*, 569–578.
- Venuti, J.M., Morris, J.H., Vivian, J.L., Olson, E.N., and Klein, W.H. (1995). Myogenin is required for late but not early aspects of myogenesis during mouse development. *J Cell Biol* *128*, 563–576.
- Vieceli, F.M., Simões-Costa, M., Turri, J.A., Kanno, T., Bronner, M., and Yan, C.Y.I. (2013). The transcription factor chicken Scratch2 is expressed in a subset of early postmitotic neural progenitors. *Gene Expr. Patterns* *13*, 189–196.
- Volohonsky, G., Edenfeld, G., Klambt, C., and Volk, T. (2007). Muscle-dependent maturation of tendon cells is induced by post-transcriptional regulation of stripeA. *Development* *134*, 347–356.
- Vorbrüggen, G., and Jäckle, H. (1997). Epidermal muscle attachment site-specific target gene expression and interference with myotube guidance in response to ectopic stripe expression in the developing *Drosophila* epidermis. *Proc. Natl. Acad. Sci.* *94*, 8606–8611.
- Vorbrüggen, G., Constien, R., Zilian, O., Wimmer, E.A., Dowe, G., Taubert, H., Noll, M., and Jäckle, H. (1997). Embryonic expression and characterization of a Ptx1 homolog in *Drosophila*. *Mech. Dev.* *68*, 139–147.
- Wahl, M.B., Deng, C., Lewandoski, M., and Pourquie, O. (2007). FGF signaling acts upstream of the NOTCH and WNT signaling pathways to control segmentation clock oscillations in mouse somitogenesis. *Development* *134*, 4033–4041.
- Watson, F.L., Mills, E.A., Wang, X., Guo, C., Chen, D.F., and Marsh-Armstrong, N. (2012). Cell type-specific translational profiling in the *Xenopus laevis* retina. *Dev. Dyn.* *241*, 1960–1972.
- Wayburn, B., and Volk, T. (2009). LRT, a tendon-specific leucine-rich repeat protein, promotes muscle-tendon targeting through its interaction with Robo. *Development* *136*, 3607–3615.
- Weintraub, H., Tapscott, S.J., Davis, R.L., Thayer, M.J., Adam, M.A., Lassar, A.B., and Miller, A.D. (1989). Activation of muscle-specific genes in pigment, nerve, fat, liver, and fibroblast cell lines by forced expression of MyoD. *Proc. Natl. Acad. Sci.* *86*, 5434–5438.
- Winberg, M.L., Mitchell, K.J., and 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.
- Witke, W., Sharpe, A.H., Hartwig, J.H., Azuma, T., Stossel, T.P., and Kwiatkowski, D.J. (1995). Hemostatic, inflammatory, and fibroblast responses are blunted in mice lacking gelsolin. *Cell* *81*, 41–51.
- Wojtowicz, I., Jab onska, J., Zmojdian, M., Taghli-Lamalle, O., Renaud, Y., Junion, G., Daczewska, M., Huelsmann, S., Jagla, K., and Jagla, T. (2015). *Drosophila* small heat shock protein CryAB ensures structural integrity of developing muscles, and proper muscle and heart performance. *Development* *142*, 994–1005.
- Wolfe, R.R. (2006). The underappreciated role of muscle in health and disease. *Am. J. Clin. Nutr.* *84*, 475–482.
- Wong, M.-C., Castanon, I., and Baylies, M.K. (2008). Daughterless dictates Twist activity in a context-dependent manner during somatic myogenesis. *Dev. Biol.* *317*, 417–429.
- Wu, H., Xiong, W.C., and Mei, L. (2010). To build a synapse: signaling pathways in neuromuscular junction assembly. *Development* *137*, 1017–1033.
- Wyzykowski, J.C., Winata, T.I., Mitin, N., Taparowsky, E.J., and Konieczny, S.F. (2002). Identification of Novel MyoD Gene Targets in Proliferating Myogenic Stem Cells. *Mol. Cell. Biol.* *22*, 6199–6208.
- Xu, X., Tan, X., Lin, Q., Schmidt, B., Engel, W., and Pantakani, D.V.K. (2013). Mouse Dazl and its novel splice variant functions in translational repression of target mRNAs in embryonic stem cells. *Biochim. Biophys. Acta BBA - Gene Regul. Mech.* *1829*, 425–435.

- Yang, Z. (2005). Isolation of mRNA from specific tissues of *Drosophila* by mRNA tagging. *Nucleic Acids Res.* *33*, e148–e148.
- Yarnitzky, T., Min, L., and Volk, T. (1997). The *Drosophila* neuregulin homolog *Vein* mediates inductive interactions between myotubes and their epidermal attachment cells. *Genes Dev.* *11*, 2691–2700.
- Yokoyama, S., and Asahara, H. (2011). The myogenic transcriptional network. *Cell. Mol. Life Sci.* *68*, 1843–1849.
- Yoshihara, M., Rheuben, M.B., and Kidokoro, Y. (1997). Transition from Growth Cone to Functional Motor Nerve Terminal in *Drosophila* Embryos. *J. Neurosci.* *17*, 8408–8426.
- Zacharias, A.L., Lewandoski, M., Rudnicki, M.A., and Gage, P.J. (2011). *Pitx2* is an upstream activator of extraocular myogenesis and survival. *Dev. Biol.* *349*, 395–405.
- Zeitlinger, J., Zinzen, R.P., Stark, A., Kellis, M., Zhang, H., Young, R.A., and Levine, M. (2007). Whole-genome ChIP-chip analysis of *Dorsal*, *Twist*, and *Snail* suggests integration of diverse patterning processes in the *Drosophila* embryo. *Genes Dev.* *21*, 385–390.
- Zeng, M., Lu, Y., Liao, X., Li, D., Sun, H., Liang, S., Zhang, S., Ma, Y., and Yang, Z. (2009). *DAZL* binds to 3'UTR of *Tex19.1* mRNAs and regulates *Tex19.1* expression. *Mol. Biol. Rep.* *36*, 2399–2403.
- Zhang, M., and McLennan, I.S. (1998). Primary myotubes preferentially mature into either the fastest or slowest muscle fibers. *Dev. Dyn.* *213*, 147–157.
- Zhang, X.M., Ramalho-Santos, M., and McMahon, A.P. (2001). Smoothed mutants reveal redundant roles for *Shh* and *Ihh* signaling including regulation of L/R asymmetry by the mouse node. *Cell* *105*, 781–792.
- Zhu, C.C., Dyer, M.A., Uchikawa, M., Kondoh, H., Lagutin, O.V., and Oliver, G. (2002). *Six3*-mediated auto repression and eye development requires its interaction with members of the *Groucho*-related family of co-repressors. *Development* *129*, 2835–2849.