Drosophila myogenesis as a model for studying *cis*-regulatory networks:

identifying novel players and dissecting the role of transcriptional repression

Lucia Ciglar

(Kayserová)

2010

Dissertation

submitted to the
Combined Faculties for the Natural Sciences and for Mathematics
of the Ruperto-Carola University of Heidelberg, Germany
for the degree of
Doctor of Natural Sciences

Presented by

Bachelor of Science Lucia Ciglar
born in Šala, Slovakia

Oral examination:
6th July 2010

Drosophila myogenesis as a model for studying *cis*-regulatory networks:

identifying novel players and dissecting the role of

transcriptional repression

Referees: Dr. Jürg Müller

Prof. Dr. Herbert Steinbeisser

Inaugural-Dissertation

zur

Erlangung der Doktorwürde

der

Naturwissenschaftlich-Mathematischen Gesamtfakultät

der

Ruprecht-Karls-Universität

Heidelberg

vorgelegt von

Bachelor of Science Lucia Ciglar

aus Šala, Slovakia

Tag der mündlichen Prüfung:

6. Juli 2010

ACKNOWLEDGMENTS

Túto prácu s láskou venujem svojím najbližším, ktorým patrí moja najväčšia vďaka.

V prvom rade ďakujem môjmu milovanému manželovi za všetku jeho lásku, pochopenie, starostlivosť, a trpezlivosť. Mojím úžasným, milujúcim rodičom, bratom, a starým rodičom, za ich nikdy nekončiacu lásku a podporu, aj keď to pre nich znamenalo mnoho neľahkých chvíľ. Ďakujem aj mojím druhým rodičom a švagrovcom za láskyplné prijatie do rodiny a manžela, o akom sa mi ani nesnívalo. Veľká vďaka patrí aj všetkým ďalším členom bližšej i širšej rodiny a rodinným priateľom za vytvorenie úžasného zázemia, ktorému som nesmierne šťastná.

LIST OF FIGURES AND TABLES

Figures

Tables

SUMMARY

Recent studies have identified *in vivo* binding profiles of key mesodermal regulators across the *Drosophila melanogaster* genome. Many of the occupied sites lie in the vicinity of loci encoding yet other transcription factors. The analyzed *cis*-regulatory modules drive expression in a variety of complex spatio-temporal patterns that cannot be explained by the binding of the core regulators alone. Thus there clearly are additional, unknown transcription factors in the regulatory network that governs the process of embryonic mesoderm specification and muscle differentiation.

In order to identify novel myogenic regulators in a systematic way, and thereby enrich the underlying network, I initiated a molecular screen to uncover new players. Candidate putative transcription factors were prioritized based on their expression in mesoderm and on available ChIP-on-chip and expression profiling data. Their role in myogenesis was subsequently assayed using *Drosophila* deficiency lines whose phenotypes were analyzed with a muscle-specific marker. Altogether, 67 different deficiency and loss-of-function lines were used individually or in combination to delete 46 transcription factors with mesodermal expression. In 21 of the 46 cases, the mutant embryos displayed specific defects in the development of one or more muscle types.

One pair of partially overlapping deficiencies placed in *trans* showed a failure in myoblast fusion, a process that gives rise to muscle syncytia from mononucleated myoblasts. The corresponding deleted candidate gene was *MED24*, a subunit of the Mediator complex, which is a general co-activator of transcription. Muscle-specific knockdown of *MED24* or *MED14*, another subunit of the complex whose deletion by deficiency lines phenocopies that of *MED24*, leads to lethality. To establish whether MED24 and MED14 are indeed involved in muscle development, I generated smaller deletion lines using FRT-mediated recombination. While deletion of *MED14* does not affect myogenesis, embryos deficient for *MED24* display supernumerary mononucleated myoblasts. Both small deletion lines were then combined together to detect possible redundancy that could obscure the requirement of MED14 and MED24 in muscle development.

Another candidate transcription factor within the myogenic network based on ChIP-on-chip experiments is the transcriptional repressor Tramtrack69 (Ttk69).

Ttk69 is expressed in the primordium of visceral and, more transiently, somatic muscle. In *ttk69* mutant embryos, homozygous for a loss-of-function allele, myoblast fusion is delayed, the myoblasts aggregate in clusters, and fail to migrate towards the ectodermal attachment sites. Two distinct myoblast populations, a founder cell and multiple fusion competent myoblasts, contribute to each muscle fibre. As revealed by immunohistochemistry and *in situ* hybridization, in *ttk69* mutants there are significantly more founder cells formed while the number of fusion competent myoblasts is decreased. Consistently, ectopic expression of Ttk69 in the founder cells, but not fusion competent myoblasts, gives rise to severe myoblast fusion defects. These phenotypic analyses suggest a model where Ttk69 is required for specification of fusion competent myoblasts and in its absence, their conversion to a founder cell-like fate may occur.

According to the proposed model, Ttk69 would repress founder cell genes within the fusion competent myoblasts. To determine whether this holds true on a global scale, I performed a high-resolution ChIP-on-chip experiment in 6-8 hour wild type embryos. Indeed, Ttk69 binding was significantly enriched in the vicinity of founder cell-specific genes as compared to fusion competent myoblast-specific genes. ChIP-on-chip data generated for Lame duck, a transcriptional activator essential for fusion competent myoblast determination, showed the opposite tendency. It therefore appears that proper specification of fusion competent myoblast identity requires both positive input from Lame duck and inhibition of founder cell-specific genes by Ttk69. These findings advance our limited knowledge about the role of transcriptional repression within the myogenic regulatory network.

Finally, I re-evaluated the role of Snail, a well-established transcriptional repressor involved in early mesoderm specification and gastrulation. Multiple observations suggested that Snail may also play a positive role in regulating mesodermal genes. To investigate this possibility, I performed luciferase assays with previously characterized mesodermal enhancers and showed that Snail can elevate their activation levels. In one case, this ability of Snail was suppressed upon mutagenesis of putative Snail binding motifs, both in cell culture and *in vivo*. Moreover, expression of the enhancers and their associated genes is significantly reduced in *snail* mutant embryos. Snail thus seems to play a dual role in repressing non-mesodermal genes, but also in contributing to the activation of some early mesodermal genes.

ZUSAMMENFASSUNG

Jüngste Studien haben die *in-vivo* Bindungsprofile der bedeutendsten mesodermalen Transkriptionsfaktoren genomweit in *Drosophila melanogaster* identifiziert. Viele dieser Bindungsstellen befinden sich in direkter Nähe weiterer Transkriptionsfaktoren. Die analysierten *cis*-regulativen Module (CRMs) werden in verschiedenen räumlichen und zeitlichen Mustern exprimiert. Diese komplexen Muster können mit der Aktivität der wenigen Schlüsselfaktoren nicht erklärt werden. Es muss daher weitere, bislang unbekannte Transkriptionsfaktoren geben, die Genregulation bei der embryonalen Mesodermspezifizierung und Muskeldifferenzierung beeinflussen.

Um systematisch neue Transkriptionsfaktoren zu identifizieren, und somit das genetische Netzwerk der Muskelentwicklung zu erweitern, startete ich einen molekulare Analyse. Potenzielle Transkriptionsfaktoren wurden anhand mehrerer Kriterien, wie z.B. ihrer mesodermalen Exprimation und ihrer möglichen Regulation durch besagte Schlüßelfaktoren, bevorzugt. Deren Wichtigkeit für die Myogenese wurde anhand von Defizienz- und 'Loss-of-function' Mutanten studiert. Insgesamt wurden 67 solcher Mutanten benutzt um 46 Transkriptionsfaktoren zu untersuchen. In 21 von 46 Fällen konnte ich bestimmte abnorme Phänotypen in der Embryonalmuskulatur beobachten.

In einem besonders interessanten Fall kann man mit sich teilweise überschneidenden Defizienzmutanten einen klaren Myoblastfusionsdefekt beobachten. Interessanterweise ist das hier fehlende Gen *Med24*, welches ein Bestandteil des Mediator Komplexes ist. Da der Mediator Komplex ein genereller Co-Activator der Genexprimation ist, bestand hier die Möglichkeit, dass es sich in *Med24* um eine gewebe-spezifische Komponente des Mediator Komplexes handelt. Auch *Med14* könnte eine muskelspezifische Rolle im Mediator Komplex spielen. Ein muskelspezifischer 'knock-down' von *Med24* als auch *Med14* ist früh in der Embryonalentwicklung letal. Um die Rolle beider Faktoren in der Muskelentwicklung näher zu untersuchen habe ich durch FRT-Rekombination wesentlich kleinere, zielgerichtete Segmente entfernt. Während Verlust von *Med14* nicht zu Myogenesedefekten führt, zeigen *Med24*-defiziente Embryonen eine hohe Anzahl von mononuklearen Myoblasten auf.

Ein weiterer Transkriptionsfaktorkandidat der *Drosophila* Muskelentwicklung ist der Repressor Tramtrack69 (Ttk69). Ttk69 wird in den Anlagen der Viszeral- und, weniger stark und weniger dauerhaft, auch in der Skelettmuskulatur exprimiert.

In *ttk69* Mutanten, ist die Fusion der Myoblasten verzögert – Myoblasten aggregieren und migrieren scheinbar nicht zu ihren respektiven ektodermalen Ansatzstellen (attachment sites). Zwei bestimmte Zelltypen, Gründer Zellen (founder cells, FCs) und Fusions-kompetente Myoblasten (fusion competent myoblasts, FCMs), tragen zu jeder Muskelfibrille bei. Sowohl Immunohistochemie, als auch *in-situ* Hybridisierung zeigen, dass *ttk69* Mutanten wesentlich mehr FCs aufweisen, während die Zahl der FCMs geringer ist. Embryonen die gezwungen werden Ttk69 in FCs, nicht aber in FCMs zu exprimieren, sterben meist ab. Dies könnte darauf hindeuten, dass Ttk69 für die Spezifizierung von FCMs nötig ist, und das diese Zellen in der Abwesenheit von Ttk69 dazu neigen sich als FCs heranzubilden

Es ist daher möglich, dass Ttk69 dazu dient die Exprimation von FC-Genen in den FCMs zu unterdrücken. Um diese Möglichkeit näher zu untersuchen, habe ich eine ChIP-on-chip Studie an 6-8 Std. Wildtyp-Embryonen durchgeführt. In der Tat, Ttk69 konnte insbesondere in der Nähe von FC-spezifischen Genen gefunden werden. Interessanterweise zeigen ChIP-on-chip Studien mit Lame Duck, einem Aktivator wichtig für FCM Identität, den gegengesetzten Phänotyp auf – Lame Duck ist vor allem in der Nähe von FCM Genen zu finden. Es scheint daher, dass zur Spezifizierung der FCMs sowohl Genaktivierung durch Lame Duck, als auch negative Regulation von FC Genen durch Ttk69 nötig ist. Diese Resultate erweitern unser sehr limitiertes Verständnis der Rolle der Genrepression innerhalb des *Drosophila* Muskelnetzwerks.

Schließlich habe ich auch die Rolle des Repressorproteins Snail näher untersucht. Snail ist vor allem für die Etablierung der naïven Mesodermanlage und für den Ansatz der Gastrulation wichtig. Einige Observationen legen es nahe, daß Snail einige Gene auch positiv regulieren könnte. Um dies näher zu erschließen habe ich Luciferase Studien durchgeführt. Tatsächlich ist es der Fall, das einige CRMs durch Zusatz von Snail in ihrer Aktivität steigen. Insbesondere in einem Fall konnte ich zeigen, das die gesteigerte CRM Aktivität auch tatsächlich von den Snail Bindungsstellen abhängt, sowohl in Zellkultur, als auch *in vivo*. Zusätzlich ist auch noch die *in vivo* Aktivität einiger CRMs in Snail Mutanten gemindert. Daher scheint Snail tatsächlich eine Doppelrolle zu spielen, indem es viele nicht-mesodermale Gene

in der Mesodermanlage unterdrückt, andere mesoderm-spezifische Gene aber aktiviert.

ABBREVIATIONS

ºC	degrees Celsius	kDa	kilodaltons
AEL	after egg laying	l	litre
bp	basepairs	M	molar
BDGP	Berkeley Drosophila Genome Project	MAPK	mitogen activated protein kinase
bHLH	basic helix-loop-helix	ml	mililitres
BTB	bric-a-brac, *tramtrack*, broad-complex	mRNA	messenger RNA
cDNA	complementary DNA	ng	nanograms
ChIP	chromatin immunoprecipitation	nM	nanomolar
CNS	central nervous system	NuRD	Nucleosome Remodeling and Deacetylase
CTD	carboxy-terminal domain	PCR	polymerase chain reaction
CRM(s)	*cis*-regulatory module(s)	PNS	peripheral nervous system
DAPI	4',6-diamidino-2-phenylindole	Pol II	RNA Polymerase II
DNA	deoxyribonucleic acid	PWM	position weight matrix
DSHB	Developmental Studies Hybridoma Bank	RAS	rat sarcoma 2 viral oncogene
dsRNA	double-stranded RNA	RNA	ribonucleic acid
EMS	ethyl methanesulfonate	RNAi	RNA interference
EST	expressed sequence tag	S2	Schneider 2
FC(s)	founder cell(s)	SELEX	Systematic Evolution of Ligands by Exponential Enrichment
FCM(s)	fusion competent myoblast(s)	SOP	sensory organ precursor
FDR	false discovery rate	TAF(s)	TBP-associated factor(s)
FGF	fibroblast growth factor	TBP	TATA-box-binding protein
FLP	flippase	TF(s)	transcription factor(s)
FRT	flippase recognition target	TSS(s)	transcriptional start site(s)
GFP	green fluorescent protein	UAS	upstream activating sequence
GTF(s)	general transcription factor(s)	UTR	untranslated region
Ig	immunoglobulin	μl	microliters
kb	kilobase pairs	μM	micromolar

GENE SYMBOLS

ac	*achaete*		*Hsp70*	*Heat shock protein 70*
abd-A	*abdominal A*		*htl*	*heartless*
ase	*asense*		*jumu*	*jumeau*
Atg18	*Autophagy-specific gene 18*		*Kr*	*Krüppel*
ato	*atonal*		*l(1)sc*	*lethal of scute*
bap	*bagpipe*		*lbe*	*ladybird early*
Bcl-6	*B-cell lymphoma 6*		*lmd*	*lame duck*
bin	*biniou*		*lwr*	*lesswright*
bnl	*branchless*		*lz*	*lozenge*
btl	*breathless*		*Mdr49*	*Multi drug resistance 49*
Cdk8	*Cyclin-dependent kinase 8*		*MED1-31*	*Mediator complex subunit 1-31*
ci	*cubitus interruptus*		*Mef2*	*Myocyte enhancer factor 2*
CtBP	*C-terminal Binding Protein*		*MESR4*	*Misexpression suppressor of ras 4*
CycC	*Cyclin C*		*MHC*	*Myosin heavy chain*
CycE	*Cyclin E*		*mmy*	*mummy*
dap	*dacapo*		*msi*	*musashi*
dl	*dorsal*		*noc*	*no ocelli*
Dll	*Distal-less*		*odd*	*odd skipped*
dmrt11E	*doublesex-Mab related 11E*		*phyl*	*phyllopod*
Doc1-3	*Dorsocross1-3*		*pnr*	*pannier*
dpn	*deadpan*		*pyd*	*polychaetoid*
dpp	*decapentaplegic*		*rap*	*retina aberrant in pattern*
dTCF	*Drosophila T-cell factor*		*repo*	*reversed polarity*
duf	*dumbfounded*		*rho*	*rhomboid*
en	*engrailed*		*rib*	*ribbon*
esg	*escargot*		*row*	*relative of woc*
E(spl)	*Enhancer of split Complex*		*rst*	*roughest*
eya	*eyes absent*		*run*	*runt*
eve	*even skipped*		*sc*	*scute*
fog	*folded gastrulation*		*scrt*	*scratch*
ftz	*fushi tarazu*		*sina*	*seven in absentia*
gt	*giant*		*sinaH*	*sina homologue*
h	*hairy*		*sim*	*single-minded*
hb	*hunchback*		*slp*	*sloppy paired*
hbs	*hibris*		*sna*	*snail*
hh	*hedgehog*		*sns*	*sticks and stones*
HmgZ	*HMG protein Z*		*sog*	*short gastrulation*

srp	*serpent*		*TRF3*	*TBP-related factor 3*
stg	*string*		*ttk69*	*tramtrack69*
sug	*sugarbabe*		*ttk88*	*tramtrack88*
Su(H)	*Suppressor of Hairless*		*twi*	*twist*
svp	*seven up*		*Ubx*	*Ultrabithorax*
TAF3	*TATA binding protein 3*		*vnd*	*ventral nervous system defective*
ths	*thisbe*		*wg*	*wingless*
tin	*tinman*		*wntD*	*wnt inhibitor of Dorsal*
tll	*tailless*		*zen*	*zerknüllt*
trbl	*tribbles*		*zfh1*	*Zn finger homeodomain 1*

1   INTRODUCTION

1.1 Differential gene expression: the basis for development

Neurons, hepatocytes, myoblasts, melanocytes, fibroblasts, oligodendrocytes. These are only a fraction of the broad cell categories that are found in our bodies. It is estimated that we are made of hundreds of different cell types, each having a very unique morphology and a highly specialized role. The function of a neuronal cell cannot be replaced by a myoblast, and *vice versa.* Yet all the cells originate from one single cell, and they are identical at the level of their deoxyribonucleic acid (DNA) sequence. How such a tremendous diversity of cell types possessing the same genetic information are generated from an individual cell remains a fundamental question of developmental biology.

The key to cell diversification is the way in which DNA information is interpreted and used. The protein repertoire of neurons is very different from myoblasts, even though the genome encoding the proteins is the same. It is the differential expression of genes that leads to the protein, and thereby cellular, diversity. As the expression of neuronal rather than muscle proteins within a muscle cell would have detrimental effects for a developing organism, gene expression is subject to tight control at multiple tiers. These range from the accessibility of the DNA sequence to the transcriptional apparatus, through stability and splicing of the synthesized ribonucleic acid (RNA), to the translation, folding, and activity of the protein. All the processes leading from the DNA sequence to a protein and their control are tightly coordinated in space and time (reviewed in [1]). The first step of gene expression regulation is the initiation of transcription, when genes are transcribed into primary transcripts in a highly controlled and complex manner [2].

1.1.1 The initiation of transcription is key to differential gene expression

Transcription is fundamental to all biological processes, cell types, and species. Its initiation requires the binding of the general transcriptional apparatus to the basal promoter sequences that lie in the immediate upstream vicinity of the transcriptional start site (TSS). In eukaryotes, there are three RNA polymerase enzymes, each

transcribing a particular type of RNA, but it is RNA polymerase II (Pol II) that is responsible for transcription of protein-coding genes and regulatory, non-coding RNAs (reviewed in [3]). Eukaryotic Pol II is recruited to the promoter region via the assembly of general transcription factors (GTFs).

The regulation of transcriptional initiation is essential for differential gene expression and thereby development. It is achieved through the binding of sequence-specific transcription factors (TFs) to their binding sites, which are typically clustered in regulatory elements (reviewed in [4]). As these elements can be located at any distance from the regulated gene and can influence transcriptional output positively or negatively, they are often referred to as *cis*-regulatory modules, or CRMs. Typically, the CRMs in the *Drosophila* genome are about 500 base pairs (bp) in size and contain binding sites for multiple TFs. The activity of a given CRM is therefore a readout of the combination of inductive as well as restrictive inputs supplied by several TFs. When CRMs are analysed individually in transgenic reporter assays, they can recapitulate all or a part of the gene's expression, but only when they are examined collectively is the full expression of the gene recovered. The expression of a gene is therefore driven by multiple modular CRMs which integrate inputs from several regulators (reviewed in [5]) (Figure 1.1-1).
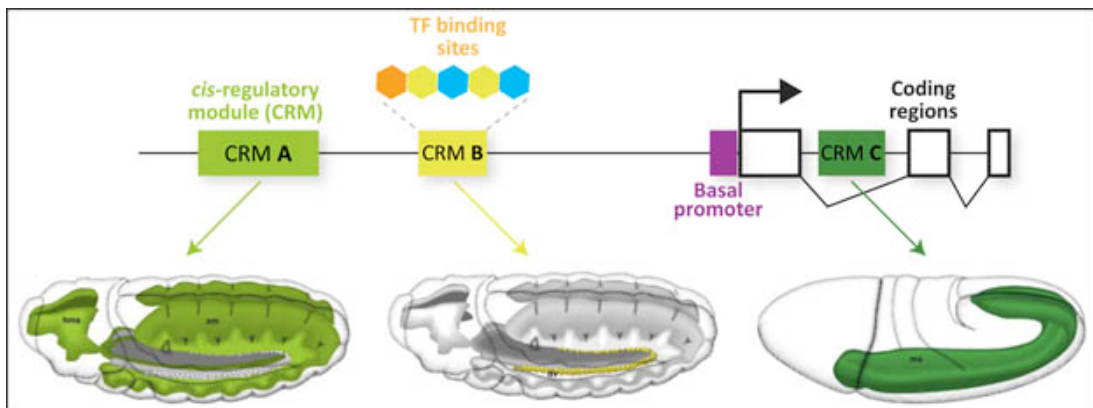
Figure 1.1-1: Tissue-specific gene expression is driven by the activity of *cis*-regulatory modules

A locus of a hypothetical gene whose expression is driven by three *cis*-regulatory modules (CRMs A, B, C) at different times and locations in the *Drosophila* embryo. Each CRM consists of clusters of binding sites (shown only for the CRM B) recognized by distinct transcription factors, TFs. The basal promoter is positioned at the start of the transcriptional unit. Note that CRMs can be located upstream, downstream, or in the intronic regions of the regulated gene. Embryo images are from [6].

1.1.2 Transcriptional repression in developmental decision making

Undoubtedly, transcriptional repression is at least as important as transcriptional activation in controlling gene expression during development. Repression is particularly important when broad morphogen gradients are translated into sharp domains of gene activity (reviewed in [7]) or when a particular cell identity is specified and maintained over time (reviewed in [8]).

There are multiple indirect mechanisms to suppress transcription which do not require the repressor to directly bind to DNA. For instance, via the inhibition of nuclear import of an activator, inaccessibility of the regulatory sequences due to chromatin composition, or titration of the activator by direct protein-protein interaction. An important category are transcriptional co-repressors which by themselves cannot bind DNA, but are recruited by sequence-specific TFs through direct or indirect protein-protein interaction. Co-repression is often mediated by localized histone deacetyletion which results in gene silencing [9].

Sequence-specific, DNA binding transcriptional repressors can block transcription by a number of mechanisms. The simplest manner is competition for overlapping binding sites between an activator and a repressor. Originally it was assumed that this mechanism was broadly used in setting up the anterior-posterior axis in *Drosophila* embryos [10]. However, mutagenesis and rearrangement of the implicated binding sites suggested that another mechanism, named quenching, might be used instead [11]. In this situation, both the activator and repressor bind to nearby sites, however the repressor prevents interaction between the activator and the basal transcription machinery, thereby quenching its transcriptional activity. An example of a TF that represses transcription via quenching is Snail, a repressor of non-mesodermal gene expression in early embryogenesis. When Snail's binding motifs are placed within a 100 bp distance from the activator sites, they can efficiently mediate short-range repression [12]. More recently, Snail was found to directly interact with two general co-repressor proteins, the C-terminal Binding Protein (CtBP) [13] and Ebi [9], that are essential for repression of its target genes.

Another mode of transcriptional repression is a direct interaction of the repressor with the basal transcriptional apparatus. While the two mechanisms described above enable repression of selective CRM activity, the outcome of direct repression is a complete abolishment of transcription, named silencing [14].

Therefore, competition and quenching contribute to the modularity of CRMs and enable their inactivation independently of other CRMs, while silencing affects the total transcriptional activity of a gene.

1.1.3 The Mediator complex and its role in the regulation of gene expression

The Mediator complex was originally isolated in yeast *Saccharomyces cerevisiae* as a component of the basal transcription apparatus necessary for the transcriptional initiation by RNA Polymerase II (Pol II) [15]. It was then successively purified from higher organisms where it was found to have an evolutionarily conserved composition and function [16]. Likely reflecting the greater organismal complexity, the metazoan Mediator complex contains additional subunits that are organized into modules with distinct functions [17]. Individual subunits have been implicated in transcription of developmentally regulated genes (see Section 1.1.3.3) and mutant alleles are often lethal at early developmental stages [18, 19]. Consequently, the Mediator complex operates as a further layer in gene expression control contributing to the developmental regulatory programs of multicellular organisms.

1.1.3.1 Mediator subunits are organized into different domains

There are currently 33 distinct genes classified by gene ontology terms [20] to encode members of the *Drosophila* Mediator complex. Similarly to the yeast complex, they are organized in four modules [21-23]. Subunits of the tail module bind sequence-specific transcription factors [24], the largest head domain interacts directly with Pol II, and the middle subunits connect the two modules together and contribute to the Pol II interaction scaffold (Figure 1.1-2). Importantly, the arrangement of the subunits into modules was also validated at the functional level as subunits within one module interact genetically in yeast [25].

The fourth and best studied Mediator domain is composed of Cyclin-dependent kinase 8 (Cdk8), Cyclin C (CycC), MED12, and MED13 (Figure 1.1-2). Unlike the three core modules, the yeast Cdk8 submodule is separable and only loosely associated with the Mediator complex [26]. In yeast, this module plays a repressive

rather than activatory function [27] and it was suggested that it does so by inhibiting stable interaction between the Mediator complex and Pol II [28]. Phenotypic analyses showed that the Cdk8-CycC and MED12-MED13 pairs play different functions during *Drosophila* development as they control expression of different genes and their loss-of-function alleles cause distinct phenotypes [18, 29]. Cdk8 in addition stimulates TFIIH-mediated phosphorylation of Pol II carboxy-terminal domain (CTD) [30]. Apart from being essential for transcription itself, hyperphosphorylation of CTD also couples transcription with messenger RNA (mRNA) processing and chromatin remodelling (reviewed in [31]).
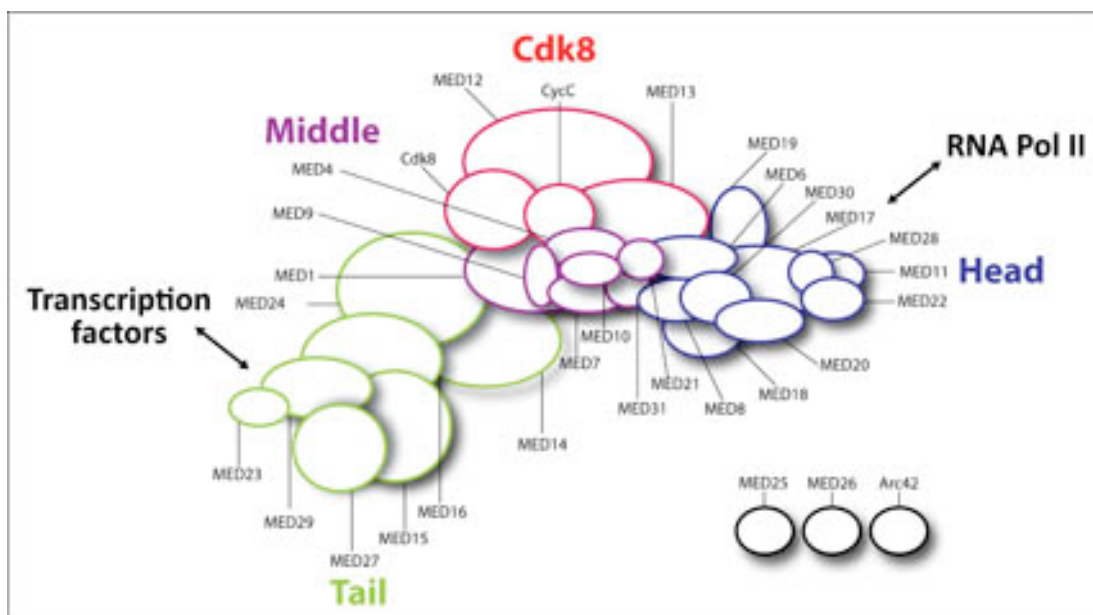
Figure 1.1-2: Organization of the *Drosophila* Mediator complex
The arrangement of the Mediator subunits was derived from pairwise interaction studies in yeast but seems to be conserved up to *Drosophila* [21, 32]. In black are *Drosophila*-specific subunits, whose physical position within the complex is unknown.

1.1.3.2 Activation of transcription by the Mediator complex

In contrast to yeast where the Mediator complex co-exists with Pol II as a holoenzyme, in higher organisms it is recruited to DNA independently [33]. Chromatin immunoprecipitation (ChIP) experiments in yeast showed that orthologues of the *Drosophila* subunits MED12, MED14, MED17, and MED22 bind to upstream activating sequences (UAS) rather than promoter regions, and they do so even when the assembly of Pol II is genetically disabled or the promoter core sequences are

deleted [34]. The yeast subunits MED7 and MED13 were found to co-occupy not only regulatory elements, but also protein-coding regions [35]. The finding that the Mediator complex is recruited in a selective manner was confirmed at a global level when ChIP coupled with microarrays was used to determine *in vivo* binding profiles of yeast MED5, MED15, and MED17 subunits [36]. Surprisingly, the Mediator complex was found to occupy DNA only under stress conditions and its binding did not fully correlate with Pol II recruitment [36].

The Mediator complex directly interacts with some, but not all, general transcription factors (GTFs), which were suggested to provide promoter specificity. However, Mediator brings also sequence-specific TFs and Pol II together [30]. The *Drosophila* complex was purified using an anti-MED31 antibody, which revealed that it physically interacts with the sequence-specific transcriptional regulators Bicoid, Krüppel, Fushi tarazu, but not with Twist or Hunchback [30]. Consequently, the Mediator complex is often pictured as a bridge integrating signals from sequence-specific regulators and the basal transcriptional machinery.

The mechanisms by which Mediator stimulates transcription remain elusive. According to a reinitiation model [37], when Pol II leaves the promoter region and enters elongation, the Mediator complex and some GTFs remain localized at the promoter to which another molecule of Pol II could then be recruited. However, the exact biochemical nature of the ability of Mediator to activate or repress transcription remains to be clarified.

1.1.3.3 Specific regulatory functions of Mediator subunits in *Drosophila* development

While yeast is an excellent system for elucidating the structure and mechanism of the Mediator complex, knowledge about its functional importance for cell viability and cell fate determination has been gathered in metazoan models, including *Drosophila* [18, 19, 38]. Differential usage of individual subunits in different developmental processes has become evident from multiple studies that are summarized in the Table 1.1-1.

Subunit(s)	Tissue/cell type	Role	Ref.
MED6	Follicle cells, imaginal discs	Cell cycle regulation, cell viability	[39]
MED12 MED13	Wing disc	Patterning along anterior-posterior axis	[40]
		Activation of Wingless targets	[41]
		Repression of *Ubx*	[29]
	Eye disc	Regulation of *atonal* and *decapentaplegic* (*dpp*); co-regulation of Atonal targets	[38, 42]
	Antennal disc	Regulation of *Distal-less* and *eyeless*	[42]
	Leg disc	Malformation of sex combs and leg shortening in mutants; expression of *bab2*	[18]
	Crystal cells	Differentiation by co-activation of Serpent and Lozenge targets	[32]
MED13 MED17	Imaginal discs	Distal sex comb and labial identity specification	[43]
Cdk8-CycC	Leg disc	Malformation of sex combs in mutants	[18]
MED15	Imaginal discs	Co-activation of Dpp targets, cell death in mutant clones	[44]
MED16	Cell line	Activation of lipopolysaccharide genes	[45]
MED17	Cell line	Activation of an antimicrobial peptide *Drosomycin* by Toll pathway	[46]
MED20 MED30	Wing	Reduced size, absence of veins, and cell death in mutant clones	[44]
MED23	Cell line	Activation of heat shock-specific gene	[45]
MED24	Salivary gland	Regulation of apoptosis	[47]
MED31	Blastoderm embryo	Establishment of anterior-posterior axis	[19]

Table 1.1-1: Mediator subunits play specific roles in *Drosophila* development
Overview of currently known roles played by the Mediator complex during *Drosophila* development with corresponding references.

1.1.3.4 Links between the general transcriptional machinery and muscle development

Individual Mediator subunits as well as GTFs have been implicated in several developmental processes, including myogenesis in *Drosophila* and vertebrates.

Although our understanding of the association between the general transcriptional apparatus and cell fate decisions is still very limited, the few known examples suggest it plays an important role that will certainly be subjected to further investigations.

It is widely accepted that Pol II is recruited to the core promoter by TFIID, a protein complex composed of TATA-box-binding protein (TBP) and several TBP-associated factors (TAFs). Therefore the observation that the expression of TBP and most TAFs upon differentiation induction in murine muscle precursor cells is significantly reduced [48], was rather surprising. Only two of the analysed subunits, TATA binding protein 3 (TAF3) and TBP-related factor 3 (TRF3) continue to be expressed and form a complex that is necessary for the myogenic regulator MyoD to activate transcription of its targets [48]. Similarly, expression of the Mediator subunits declines upon differentiation and the presence of Mediator is not required for transcription of muscle differentiation genes *in vitro* [49]. These exciting findings showing differential usage of the basal transcriptional machinery during cell differentiation provoke many questions which remain to be answered. For example, it is currently not clear whether this is a muscle-specific phenomenon or a widely used mechanism common to all differentiating cells. Requirement of other TAF proteins was previously described in Dorsal and Twist mediated activation of early mesodermal genes in *Drosophila* embryos [50]. How the levels and requirement of TAF proteins change with the progression of development has however not been addressed.

A subunit of the Mediator head module, MED28, counteracts the smooth muscle differentiation program in two different murine cell types by negatively regulating the implicated genes [51]. Although *in vivo* evidence is still lacking, the results support the notion of tissue-specific roles for distinct Mediator subunits as well as the participation of the complex in negative aspects of gene regulation.

Importantly, RNA interference (RNAi) against three subunits of the *Drosophila* Mediator complex, MED4, MED11, and MED24, were reported to lead to muscle-related defects during embryogenesis [52]. As the aim of the study was identification of novel neuronal regulators, the muscle phenotype was not presented or analyzed in more depth. Nevertheless, together with the above-mentioned examples, this observation points to a role of specific Mediator subunits in *Drosophila* muscle development.

1.1.4 General principles of developmental *cis*-regulatory networks

In higher organisms, a single sequence-specific TF can contribute to the expression of hundreds or even thousands of genes [53]. Specification of a given cell type requires many TFs, each having a specific set of target genes. As a single CRM integrates inputs from multiple TFs, the control of gene expression is highly combinatorial. Consequently, a developmental program is driven by intricate regulatory networks or circuits composed of many TFs impinging on a partially overlapping array of CRMs (reviewed in [54]).

Regulatory circuits in addition to being very elaborate and dynamic are also highly interconnected. For example, the networks governing the patterning of the *Drosophila* embryo along the anterior-posterior and dorsal-ventral axes are interconnected through the action of their key regulators [55, 56]. An emerging common theme of developmental networks is also their degree of feed-back and feed-forward regulation, where TFs activate genes encoding yet additional TFs which then either feed forward or on each other to co-regulate common sets of target genes (reviewed in [57]).

Developmental regulators are usually conserved across species, although their position within the regulatory network might have diverged. For example, members of the Eya/Six family of TFs are essential for muscle development across metazoans. However, while they are positioned at the top of the regulatory hierarchy in vertebrates, in the invertebrate model systems *Drosophila melanogaster* and *Caenorhabditis elegans* they are placed further downstream of the myogenic regulators (Figure 1.1-3). These changes in network architecture might reflect the natural differences in the physiology and complexity of the organisms. In some systems, certain developmental subcircuits may be fixed across evolution not only at the level of the TFs and regulated processes, but also in the wiring between their components [58]. As the data required to perform direct interspecies comparisons of gene regulatory networks are still very limited, many questions regarding the conservation and divergence of network topology remain to be addressed.
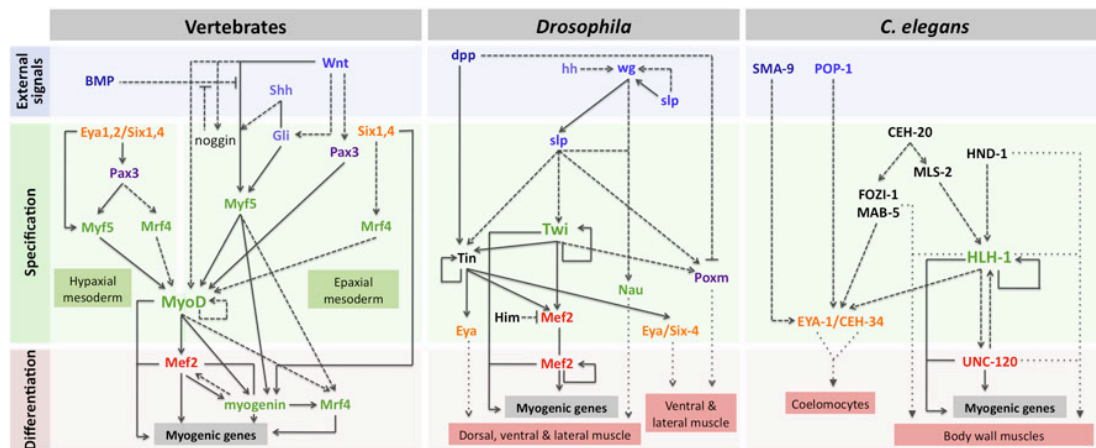
Figure 1.1-3: Conservation and divergence of regulatory networks across evolution

Comparison of the core regulatory network underlying myogenesis in vertebrates, *D. melanogaster*, and *C. elegans*. TFs that are members of the same protein family are displayed in the same colour. Full lines show direct transcriptional regulation and dashed lines represent genetic interactions currently lacking evidence for direct regulation.

1.1.5 *Drosophila* muscle development as a model for dissecting the logic of regulatory networks

The fruitfly *Drosophila melanogaster* has a long-lasting tradition in developmental genetics as an indispensable invertebrate model. Thanks to a century of research, there is an enormous toolkit available to study *Drosophila* development, behaviour, diseases, or evolution at diverse levels. Compared to vertebrate models, *Drosophila* has the great advantage of a fast generation time, relatively simple development, and a compact genome. At the same time, many of its processes are analogous to those in vertebrates, both at the morphological and molecular level.

One such process is the development of the larval musculature that occurs already during embryonic stages. Similarly to vertebrates, there are three types of muscle tissues in *Drosophila* embryos: somatic (or body wall) muscle enabling locomotion, visceral muscle lining the digestive tract, and the heart (or dorsal vessel) pumping the haemolymph. The myoblast cells that these tissues are composed of vary in their position, morphology, as well as function, yet they all come from a common pool of pluripotent mesodermal cells (Figure 1.1-4 A). Decoding the *cis*-regulatory mechanisms leading to the specification of the three muscle lineages is central to understanding muscle development and identifying general principles that are likely to be shared with other developmental regulatory networks.

Genetic, and more recently genomic, studies are beginning to reveal the molecular blueprints and the regulatory pathways leading to the specification of different muscle types. There are five core transcription factors that orchestrate specification of mesodermal precursor cells and their progressive diversification and differentiation into the three muscle tissues (Figure 1.1-4 B). The mesodermal master regulator is a basic helix-loop-helix (bHLH) factor Twist (Twi) that initiates a cascade of mesoderm-specific regulatory events. Its immediate targets are Myocyte enhancer factor 2 (Mef2) [59], a major differentiation factor required for all three muscle types, and Tinman (Tin) [60], necessary for the specification of the heart, the visceral muscle, and a subset of the somatic muscles. Tin promotes visceral muscle specification by activating Bagpipe (Bap), which then together with its target Biniou (Bin) establishes the visceral muscle identity program [61]. Mef2 is expressed until the very end of embryogenesis in all three muscle tissues where it induces differentiation by regulating expression of a variety of genes [62].
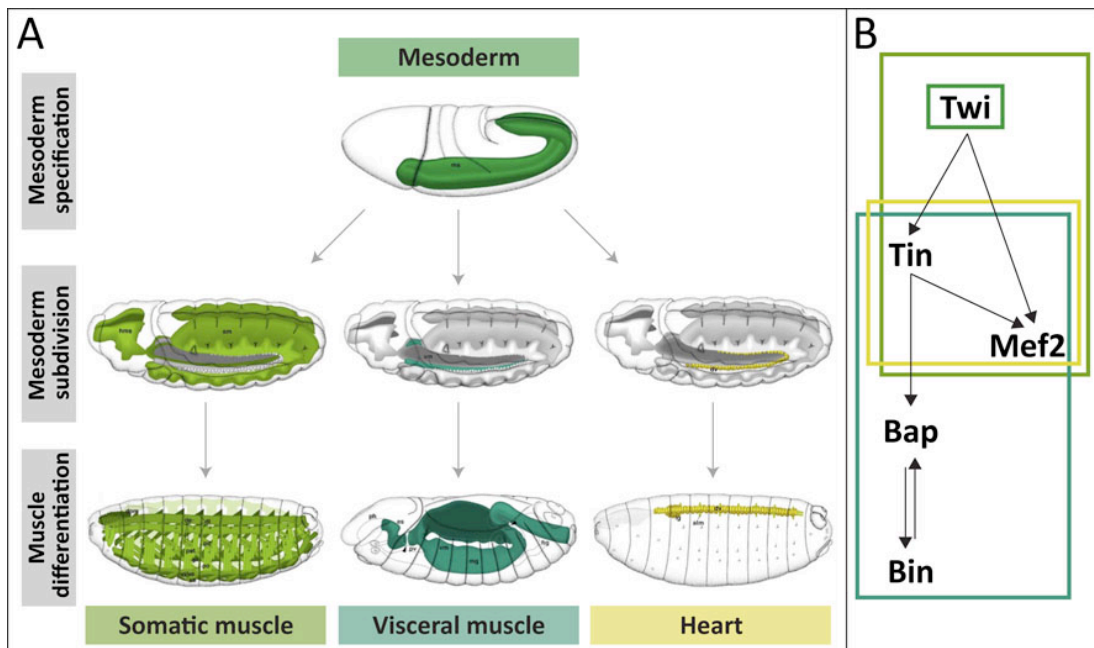
Figure 1.1-4: Muscle development in *Drosophila melanogaster*

(A) Overview of the three muscle types in the *Drosophila* embryo and an outline of the processes leading to their formation. Embryo images are from [6].

(B) The five key mesodermal TFs forming the core myogenic network. The colour of the boxes indicates the tissues from (A) that have associated phenotypes for each TF.

1.2 Early stages of mesoderm development in *Drosophila* embryos

1.2.1 A nuclear gradient of the transcription factor Dorsal subdivides the dorsal-ventral axis of the blastoderm embryo

Division of the blastoderm embryo along the dorsal-ventral axis sets the boundaries between the presumptive mesoderm and ectoderm and serves as a paradigm to study how a single factor can trigger initiation of several unique developmental programs. The TF Dorsal acts as a morphogen, where its nuclear concentration is distributed in a ventral to dorsal gradient. The nuclear concentration gradient of Dorsal is interpreted at the *cis* level and leads to different transcriptional responses along the dorsal-ventral axis, resulting in patterning of the embryo into regions of cells with different identities.

1.2.1.1 Establishment of a concentration gradient of nuclear Dorsal

The foundation of the complex events leading to the nuclear Dorsal gradient formation is laid during oogenesis, when the maternally supplied Dorsal protein is prevented from entering the oocyte nucleus via an association with the protein Cactus [63, 64]. Upon fertilization, serine proteases localized in the space between the oocyte and its vitelline shell are, through a poorly understood mechanism, activated only in the ventral region where they produce activated Spätzle protein [65]. The activated Spätzle ligand binds to the ubiquitously distributed Toll membrane receptor and induces a cascade of signalling events, which lead to the phosphorylation and proteasome-dependent degradation of Cactus (reviewed in [66]). At this stage, the embryo is still a syncytium and the cell membranes have not yet been formed, so the released Dorsal protein can readily enter the nuclei. Since the Toll signalling pathway is activated by Spätzle ligand diffusing through the perivitelline space from the ventral-most region, the concentration of nuclear Dorsal protein is highest in the ventral-most cells and gradually decreases from the ventral to dorsal side of the embryo (Figure 1.2-1).

1.2.1.2 Different Dorsal concentrations elicit different transcriptional responses

Although only two different Dorsal binding motifs have been identified [67, 68], the differences in its effective concentration combined with other regulatory factors lead to several broad classes of regulatory responses (reviewed in [69]). In the fourteen ventral-most cells, which contain the highest levels of nuclear Dorsal protein, Dorsal activates the expression of the mesodermal master regulator Twist via low affinity Dorsal binding sites [70]. Together with Twist, Dorsal then initiates the expression of a whole range of mesodermal genes, including *snail* [68]. In the lateral region of the embryo, intermediate levels of Dorsal specify the neurogenic ectoderm anlage. Depending on the input from additional signalling pathways, Dorsal targets can be expressed throughout the entire neurogenic ectoderm (such as *short gastrulation* (*sog*) [71] or *thisbe* (*ths*) [72]) or only in subparts (such as *rhomboid* (*rho*) [67] or *ventral nervous system defective* (*vnd*) [73]). Delineating the border between mesoderm and neurogenic ectoderm is a single row of mesectodermal cells marked by expression of *single-minded* (*sim*) [74].

Although Dorsal is inherently a weak activator, in certain developmental contexts it can repress transcription [75]. By recruiting the general co-repressor factor Groucho [76], Dorsal negatively regulates transcription of *decapentaplegic* (*dpp*) [77] and *zerknüllt* (*zen*) [78] in the ventral and lateral regions of the embryo. Therefore the absence or very low levels of Dorsal in the nuclei of the dorsal-most cells lead to the localized expression of *dpp* and *zen*, resulting in the specification of the dorsal ectoderm and extraembryonic tissues.
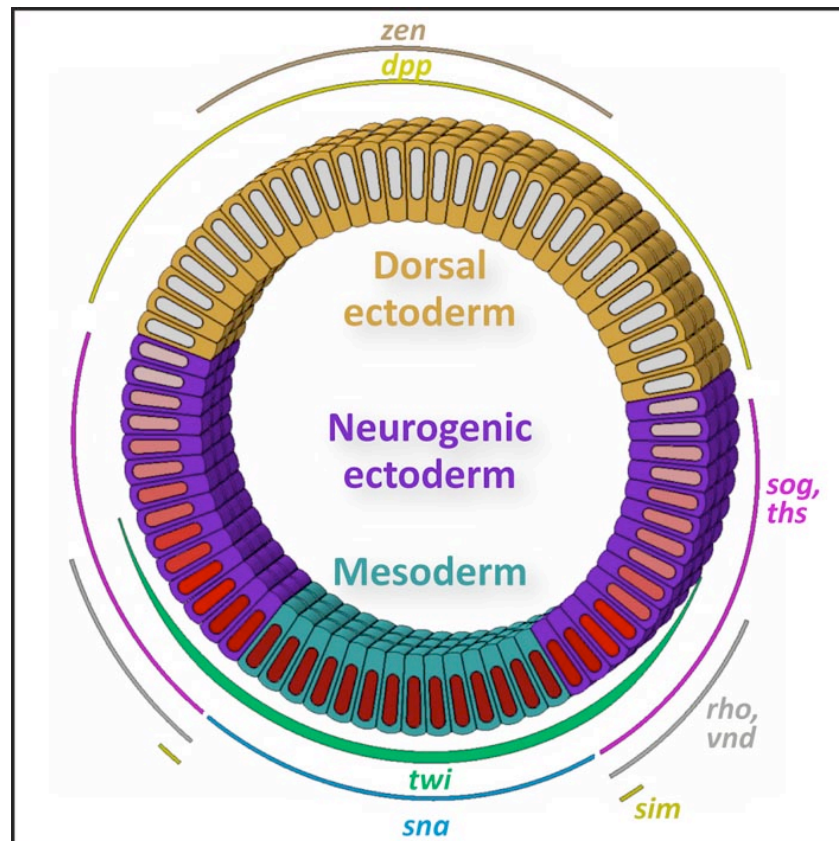
Figure 1.2-1: The establishment of the dorsal-ventral axis in the blastoderm embryo

Cross section through a blastoderm (stage 5, staging according to [79]) embryo showing the mesoderm in blue, neurogenic ectoderm in purple, and dorsal ectoderm in yellow, which are specified by a gradient of intranuclear Dorsal protein (red). Genes differentially expressed along the dorsal-ventral axis are shown next to the site of their expression.

1.2.2 Twist and Snail: two major transcriptional regulators in the presumptive mesoderm

Being a classic example of a true master regulator, the transcription factor Twist (Twi) is both essential for mesoderm specification [80] as well as being sufficient to convert non-mesodermal cells to a mesodermal fate [81]. The molecular mechanisms underlying this capacity have been recently uncovered by genome-wide studies [55, 56], which established Twi as a direct regulator of a vast collection of genes involved in transcription, cell morphogenesis, migration, and proliferation. One of the shared direct targets of Twi and Dorsal is *snail* (*sna*) [68], encoding a zinc finger transcriptional repressor that is critical for the specification of the presumptive mesoderm and its subsequent invagination during gastrulation [82].

Although the expression patterns of *twi* and *sna* are vastly overlapping, they are not fully coincident, which has important functional implications [83]. Twi is expressed in a graded fashion with the highest levels in the ventral-most mesodermal cells, but decreases in expression towards the more lateral regions [83]. The boundary of its expression is diffuse and extends up to the mesectoderm and the most ventral cells of the neurogenic ectoderm where Twi is required for the activation of respective genes (Figure 1.2-1). Sna, on the other hand, shows a more uniform expression confined to the presumptive mesodermal cells [83]. Its boundaries are rather sharp and precisely define the border of the mesoderm as it represses mesectodermal and neuroectodermal gene expression within the mesodermal domain [83].

In the mesoderm, Twist and Snail orchestrate expression of a repertoire of genes necessary for gastrulation, which is one of the first morphogenetic movements that a *Drosophila* embryo undergoes. In the course of gastrulation, mesodermal cells are internalized in a two-step process. During developmental stages 6 and 7 (staging according to [79]), the ventral-most cells undergo apical constriction where their apical surfaces become shorter. This forces the basal side of the cells to ingress into the interior of the embryo, forming a furrow. At stage 8, the new internal tube formed by the furrow dissociates and the cells spread laterally over the ectodermal cells [82]. As a result, by stage 9 the mesodermal cells are arranged in a continuous sheet beneath the outer, ectodermal cells. During stages 8 to 11, the mesodermal cells undergo three rounds of mitosis [79], hence their number increases rapidly.

Up to late stage 10, the mesoderm is a homogenous population of pluripotent cells that are recognized by the expression of general, termed panmesodermal, markers *twi* and its targets *Mef2* and *tin*. Expression of *sna* undergoes a switch where at stage 8 it ceases to be expressed in the mesoderm and instead initiates in the neuroectoderm where it remains until the end of embryogenesis [84].

1.2.2.1 A dual role for Snail as both a positive and negative regulator of transcription

Traditionally, Snail has been perceived as a dedicated repressor of transcription with a well-characterized set of direct target genes [56, 83]. Experimental results that suggested a positive regulatory input from Snail were therefore usually interpreted as

indirect effects of Snail repressing other, unknown repressor(s). However, the recent findings from genome-wide chromatin immunoprecipitation techniques have inspired re-evaluation of the possibility that Snail could indeed elicit positive regulation of transcription. These experiments showed that Snail occupies a number of enhancers that are active in mesodermal cells at the same stages as Snail binding was observed (Figure 4.3-1). In light of these results, the evidence dispersed in literature within the past two decades is consistent with the premise that Snail could indeed function as a transcriptional activator.

When some of the key mesodermal genes were described for the first time, their expression was analysed in the available mutant lines to identify their upstream regulators. Consequently, it was noticed in individual studies, that mRNA transcripts of *twi* [85], *Mef2* [86], *tin* [87], *heartless* (*htl*) [88], *folded gastrulation* (*fog*) [89], *Zn finger homeodomain 1* (*zfh1*) and *serpent* (*srp*) [90], and at least three other mesodermal genes [91] are either substantially reduced or completely absent in *sna*-deficient embryos. Absence of mesodermal gene expression in *sna* mutants was often attributed to the reduction in *twi* mRNA [85] as Twi is the major activator of these genes [59, 60]. However, Twi protein persists in *sna* mutant embryos during and even after the stages when the expression of mesodermal genes is lost [82]. Moreover, there is no known candidate repressor upstream of these genes that would be repressed by Snail, suggesting that Snail is directly needed for activation of mesodermal genes. This proposal is strongly supported by an *in vivo* experiment where low levels of Snail protein introduced genetically into *sna* mutants partially restored *tin* expression in the presumptive mesoderm [92]. In contrast, the repression of *rho* and *sim* was not rescued by low levels of Snail [92], suggesting that repression and activation might need high and low levels of the TF, respectively.

Another line of evidence for activatory input from Snail comes from reporter assays in *Drosophila* Schneider 2 (S2) cells. The reporter was driven by a Snail binding sequence element identified in an electrophoretic mobility shift assay [93]. When a Snail expression vector was co-transfected into the cells, the reporter activity levels were estimated to increase up to 32-fold [93].

In *sna* or t*wi* mutant embryos, mesodermal cells fail to invaginate but can form irregular folds or a furrow, respectively [82]. When both *sna* and *twi* are absent, the mesodermal cells do not undergo any apparent cell shape changes and form a furrow only rarely [82]. Expression of Sna can induce partial invagination even in absence of

Twi while Twi is not able to rescue the gastrulation defect in *sna* mutant [92]. In addition, mesodermal cells expressing *sim* and *rho* can still invaginate [92], indicating that derepression of neuroectodermal genes is unlikely to be the only reason for gastrulation failure in *sna* mutants. One of the proposed models explaining the gastrulation defects in *sna* mutant embryos therefore relies on Snail positively regulating genes that initiate the ventral furrow invagination [92].

Finally, a Snail-like TF can activate reporter expression in *Caenorhabditis elegans* animals through a Snail motif that is necessary for the reporter activity [94]. In a human liver carcinoma cell line, Snail co-operates with two transcriptional activators and through direct DNA binding stimulates transcription of their target gene [95]. Therefore, the potential of Snail to perform a dual function might not be confined to *Drosophila* but may be a general feature of Snail-like TFs in multiple species.

1.3 Muscle development in *Drosophila* embryos

1.3.1 Inductive signals from the ectoderm subdivide the mesoderm into four domains

After gastrulation, the mesodermal cells are pluripotent and can give rise to any of the *Drosophila* muscle types [96]. The observation that the developmental fate of the transplanted mesodermal cells is related to their position in the embryo [96] suggested an involvement of other germ layers in mesoderm specification. A molecular understanding of this process was initiated when the signals necessary for mesoderm patterning were identified. Secreted ligands from the overlying ectoderm in combination with mesoderm intrinsic factors (for example Twist) subdivide the mesoderm into roughly four domains per each of the eleven thoracic and abdominal hemisegments.

Along the dorsal-ventral axis, the *dpp* morphogen restricts the expression of Tin to the dorsal mesoderm [97] where it is required to specify the primordia of heart, visceral muscle, and dorsal somatic muscle [98]. The subdivision of the mesoderm along the anterior-posterior axis is under the control of the pair rule genes *even skipped* (*eve*) and *sloppy paired* (*slp*) along with the segment polarity genes *hedgehog* (*hh*) and *wingless* (*wg*), respectively [99]. Wg signalling via its effector *Drosophila* T-cell factor (dTCF) directly activates *slp* only in the Wg-compartment [100]. In the dorsal Tin-expressing domain of the Wg-compartment, the primordia of the cardiac and dorsal somatic mesoderm are established [101, 102] (Figure 1.3-1). The ventral mesodermal cells that are not exposed to Dpp signalling will contribute to the remaining somatic musculature. The dorsal part of the *eve* compartment, that does not receive Wg signalling, forms the visceral mesoderm primordium [101] while the ventral region gives rise to fat body (Figure 1.3-1). The key regulator of fat body lineage is Serpent, a GATA-like TF whose deletion leads to an absence of fat body progenitors [103]. Differentiated fat body cells fill in the cavities between somatic and visceral musculature and are non-contractile cells that perform many of the metabolic functions of the mammalian liver [104, 105].

As a consequence of the subdivision of the mesoderm along the dorsal-ventral and anterior-posterior axes, there are three different types of muscle primordia in *Drosophila* embryos. These give rise to somatic, visceral, and cardiac muscles. Once

their fate is committed, the myoblasts of the somatic and visceral mesoderm initiate differentiation during which they change their morphology, migrate, divide, undergo fusion, interact with other cells, and finally produce functional, contractile myofibers.
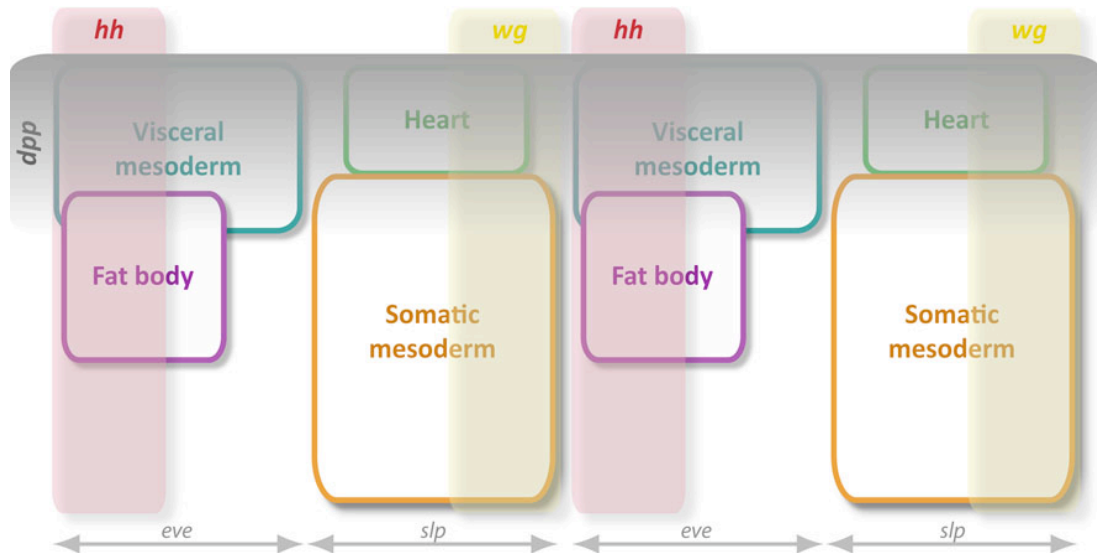
Figure 1.3-1: Subdivision of mesoderm by synergism between extrinsic and intrinsic signals

The presumptive mesoderm is specified by ectodermally-derived *dpp*, *hh*, and *wg* signalling molecules in combination with mesodermally expressed *eve* and *slp* pair rule genes into four primordia. Two of the eleven hemisegments are shown.

1.3.2 Specification and morphogenesis of the visceral musculature

The visceral musculature (VM) in the *Drosophila* embryo consists of four muscle types surrounding the gut: foregut VM, hindgut VM, trunk or midgut circular VM, and the longitudinal VM which forms a second muscle layer only around the midgut (reviewed in [106]).

The establishment of the circular VM is the best characterized and its specification is initiated by the interplay of positive signals from Dpp and Tinman and negative regulation by Sloppy paired [101]. The outcome of their combined action is the activation of Bagpipe (Bap) only in the dorsal region of the *eve* compartment in each parasegment. Bagpipe along with the effectors of Dpp signalling and Tin activate Biniou (Bin) [61]. Bin then feeds back and positively regulates Bap and deletion of either Bin or Bap results in a complete failure of circular VM formation [61, 107]. While Bin is a chief transcriptional regulator targeting a broad repertoire of

developmental genes, Bap has a more restricted regulatory potential and typically serves as a contributor to Bin-mediated regulation [108].

Once specified, the circular VM progenitors migrate along the anterior-posterior axis to form a continuous row of cells that ingress inside the embryo. These cells are however not a homogenous population as two distinct cell types, a founder cell and a fusion competent myoblast, undergo heterotypic fusion to form binucleated muscle syncytia [109]. These muscle cells then elongate vertically and cover the entire gut circumference. Finally, the visceral muscle is further subdivided along the anterior-posterior axis [110] and crosstalk between homeotic genes and signalling pathways specifies precise positions of the gastric caeca and the three constrictions around which the midgut will undergo morphogenetic looping at stages 15 and 16 [111].

1.3.3 Cardioblasts specification and the formation of a functional heart

The *Drosophila* heart is a dorsally located linear tube, the main function of which is to pump the haemolymph throughout the larval body. It is composed of two parts, the aorta and heart proper in the anterior and posterior portion of the embryo, respectively. At a cellular level, two major cell groups contribute to the heart; the contractile cardiomyocytes or cardioblasts themselves and the supporting pericardial cells. Both cell types arise from the dorsal embryonic mesoderm and are segregated through asymmetric cell division [112]. The integration of inputs from Dpp and Wg signalling and Tin, Dorsocross (Doc) and Pannier (Pnr) transcription factors leads to cardioblast specification within the cardiac mesoderm [113, 114]. In each parasegment, there are six cardiac cells that are further patterned, as four cells are *tin*-positive, two of which express another TF, Ladybird early (Lbe) [115], while the two other cardioblasts are *tin*-negative but *seven up* (*svp*)-positive [116]. Eventually, the distinct cell fates translate into different physiological functions as, for example, the cells marked by *svp* in the three most posterior segments participate in the formation of ostia, the inflow tracts for the haemolymph [117].

1.3.4 The specification of somatic muscle progenitors is orchestrated by crosstalk between multiple pathways

The somatic muscle primordium contains persistently high levels of Twist protein, which requires Wg signalling and *slp* expression [118, 119]. In contrast, the visceral mesoderm and fat body are derived from the low-Twist expressing domain, as a result of Notch-mediated Twist repression ([120], see Section 1.3.4.1). High levels of Twist are essential for somatic mesoderm specification as when they are genetically reduced, the somatic progenitors are specified as visceral mesoderm and fat body [81]. The population of somatic mesoderm cells is further patterned by the proneural gene *lethal of scute* (*l(1)sc*) into 19 myogenic competence fields per hemisegment [121]. Within each *l(1)sc* cluster, the *l(1)sc* expression is progressively restricted to a single cell by Notch-dependent lateral inhibition [122]. This cell is a muscle progenitor and in a process requiring Notch, asymmetric cell division, it forms two daughter cells [123]. In most cases, both daughter cells are different muscle founder cells (FCs) and each FC will eventually give rise to one of the thirty muscle fibres per hemisegment. The remaining, *l(1)sc*-negative population of somatic myoblasts with activated Notch signalling, forms fusion competent myoblasts (FCMs) [124]. FCMs are thought to be a naïve population of myoblasts that become "entrained" upon fusing with a FC to form a multinucleated muscle syncytium (reviewed in [125]).

1.3.4.1 Notch signalling is required at multiple stages of somatic muscle development

The Notch pathway is one of the key signalling pathways and is heavily used throughout the developmental life cycle of presumably all metazoans. It is best known for its role in amplifying the differences between two similar, locally communicating cells, to give rise to two alternative developmental fates (reviewed in [126]). In a simplified view, the Notch signalling operates via membrane-associated Notch receptor and its membrane-associated ligands. If two adjacent cells, one expressing the Notch receptor and the other one its ligand, come in contact and binding of the ligand to Notch occurs, the intracellular Notch domain is released. It then enters the nucleus where it constitutes a TF with Suppressor of Hairless (Su(H))

protein. Among its targets are genes encoding members of the Enhancer of split (E(spl)) complex of transcriptional repressors that typically block a particular cell fate.

In the somatic muscle lineage, there are three steps during which Notch signalling plays a critical function (Figure 1.3-2):

1. Setting up the myogenic competence field by restricting Twist expression: In embryos expressing a constitutively active form of the Notch receptor in the entire mesoderm at stage 10, the levels of Twist protein are reduced and the somatic mesoderm is not specified [81]. In *Notch* mutants, *twist* is expressed throughout the mesoderm and somatic muscle is formed at the expense of other muscle tissues. Differences in the levels of Twist along the anterior-posterior axis of the mesoderm are therefore essential for the correct subdivision of the muscle primordia and Notch signalling is hence another important pathway involved in this process.

2. Progenitor selection by lateral inhibition: Restriction of *l(1)sc* expression to a single muscle progenitor is likewise dependent on the Notch pathway as in embryos with ectopic Notch signalling throughout the mesoderm, no progenitors can be recognized by the expression of *l(1)sc* or later progenitor specific markers [121, 127]. In wild type embryos, active Notch signalling leads to the restriction of *l(1)sc* expression to a single cell ("signal sending cell") which will become the progenitor cell. In the neighbouring cells receiving Notch input ("signal receiving cells"), the expression of *l(1)sc* is suppressed and as a result, they acquire an FCM identity.

3. Asymmetric cell division of the muscle progenitors: After specification of the progenitor cells, Notch signalling continues to be essential as it controls their asymmetric cell division [123]. The progenitors do not express an identical repertoire of proteins as they already initiate the program responsible for differentiation of the 30 different muscle fibres (Figure 1.3-3). When a progenitor divides, one of its daughter cells continues to be under active Notch signalling unlike the other one that expresses asymmetrically localized repressor of Notch, *numb* [123]. As a result, the cells differ in the collection of expressed genes that will eventually be reflected at morphological level.
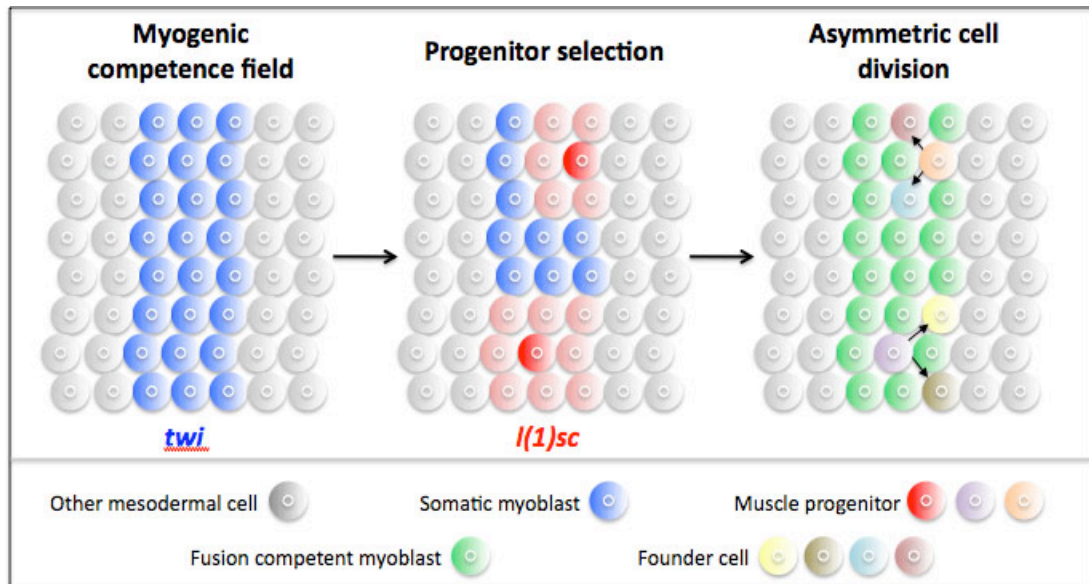
Figure 1.3-2: Notch-dependent processes in somatic muscle development

Schematic representation of key processes leading to somatic muscle specification. At all three steps, Notch signalling plays an essential role. For simplicity, only two progenitors per segment are shown.

1.3.5 Myoblast fusion: when founder cells met fusion competent myoblasts

In the suggested multistep process of myoblast fusion (reviewed in [125]), each of the 30 founder cells (FCs) in one hemisegment fuses with a defined number of neighbouring fusion competent myoblasts (FCMs) to form a multinucleated myotube. It is the molecular traits intrinsic to the FCs that dictate the unique features of each muscle fibre, such as its size, shape, or attachment. These unique traits are conferred by transcription factors known as muscle identity genes, such as Krüppel, Slouch, or Ladybird [128-130], that are expressed in overlapping combinations in the individual FCs. The currently known muscle identity factors and their combined expression patterns cannot account for the identity of all 30 muscle fibres, suggesting that more remain to be discovered.

All FCMs, on the other hand, express *lame duck* (*lmd*), coding for a zinc finger transcription factor essential for their differentiation [131, 132]. In *lmd* mutant embryos, the undifferentiated FCMs undergo apoptosis and hence there are no multinucleated muscle fibres formed [131, 132]. The only known direct target of Lmd is the differentiation regulator *Mef2* [131] and ectopic Lmd alone is sufficient to

drive *Mef2* expression in the nervous system [133]. Lmd possibly also directly activates *sticks and stones* (*sns*), a transmembrane protein mediating recognition of, and adhesion to, FCs, as its mRNA levels are reduced in *lmd* mutants [131, 132]. Interestingly, in *lmd* mutants *twist* expression remains inappropriately high in FCMs whereas in wild type embryos its expression in somatic mesoderm is normally reduced after stage 11 [132]. There is currently no evidence that Lmd acts to repress Twist directly, however, other members of the Gli family of TFs that Lmd belongs to can act as both transcriptional activators and repressors (reviewed in [134]).

As a result of their distinct identities, the two types of myoblasts also differ in the machinery executing the fusion itself. FCs express redundant Immunoglobulin (Ig)-domain transmembrane proteins *dumbfounded* (*duf*) and *roughest* (*rst*) [135] while the FCMs express, again redundantly functioning, *sns* and *hibris* (*hbs*) [136]. In the first step of the myoblast fusion, the four transmembrane proteins mediate heterotypic recognition and subsequent adhesion between the two cell types. Once the cells have tightly adhered to one another, electron dense plagues form at the points of cell contact and the plasma membrane breaks down. In the first round of myoblast fusion only limited fusion occurs and myotubes contain on average only two to three nuclei [137]. In the second phase, initiated by the migration of unfused FCMs, multiple rounds of fusion events occur, leading to the largest muscle fibres, which are estimated to contain up to 25 nuclei [138].

Not surprisingly, the dynamic events are accompanied by changes in the shape and thus in the underlying cytoskeleton of the involved cells. In recent years, many components and regulators of the actin cytoskeleton involved in myoblast fusion have been elucidated in *Drosophila* [139-141]. Nevertheless, there are still many aspects to be clarified, such as how the number of fusion events is controlled or what the mechanisms of the pore formation and plasma membrane breakdown are.

1.3.6 From myoblast fusion to muscle attachment

Already during myoblast fusion, the growing myotubes are guided towards their attachment sites to which they will eventually establish firm connections. The migration events can be separated into three consecutive steps [142]. During stage 12, the founder cells that originated from a common muscle progenitor cell migrate away from each other and the first rounds of fusion commence [142]. At stage 13, the

myotubes extend cytoplasmic protrusions, or filopodia, that scan their surrounding for guidance cues [142]. During this time, myoblast fusion continues in the central part of the myotubes. Finally, once the ectodermal tendon cells are reached, filopodia formation terminates and integrin-containing adhesion complexes are brought into the contact site of the two cells. By the end of stage 16, all muscle fibres have completed the fusion process and are attached to the ectoderm.

The epidermal tendon cells to which the somatic muscles attach are regularly arranged at the segmental borders (intersegmental) as well as within individual segments (intrasegmental) (reviewed in [143]). Their specification is driven by the transcriptional factor Stripe that is able to induce expression of several tendon cell specific genes when expressed ectopically in the ectoderm [144]. At a molecular level, the mechanisms behind the directed myotube guidance and tendon cell recognition are poorly understood. However, it seems that the signalling pathways involved in axonal pathfinding in the nervous system are utilized in myogenesis as well.

By stage 17, when embryogenesis is completed and the transition to a larva is about to initiate, a stereotypic pattern of 30 somatic muscles is formed in each of the abdominal segments 2 to 7 (Figure 1.3-3). Based on the dorsal-ventral axis, somatic muscles can be subdivided into three groups: dorsal, lateral, and ventral. Each muscle fibre has a characteristic size, shape, and orientation with respect to the other muscle fibres. This unique muscle morphology is reflected in the nomenclature system based on the muscle's position and orientation [145] or an assigned number [146].
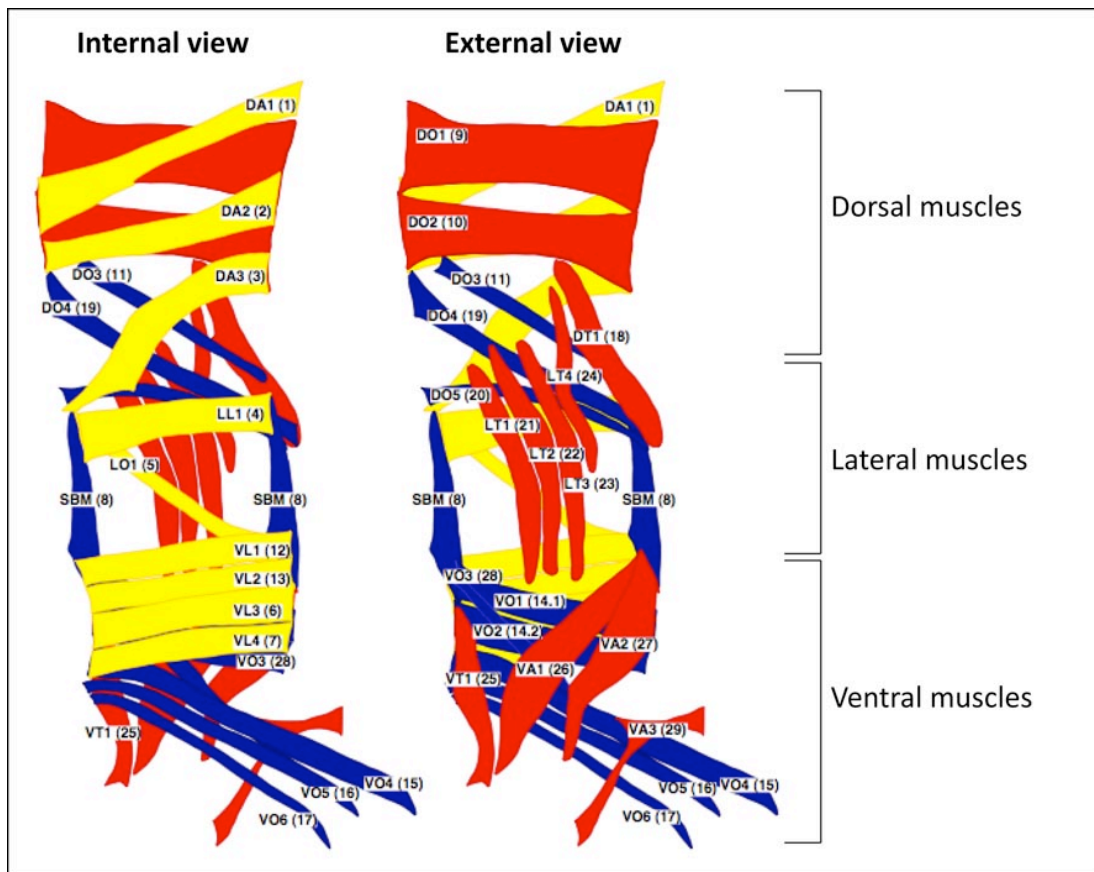
Figure 1.3-3: Larval somatic muscles have a stereotypic pattern

Internal (left) and external (right) muscle fibres in A2-A7 segments of stage 17 embryos. The colour coding indicates position of the muscle (red- external, yellow and blue- internal). Nomenclature follows M. Bate [145] and AC. Crossley [146]. From bottom to top: VO, ventral oblique; VA, ventral acute; VL, ventral longitudinal; SBM, segment border muscle; LT, lateral transverse; LO, lateral oblique; LL, lateral longitudinal; DT, dorsal transverse; DO, dorsal oblique; DA, dorsal acute. The images are from [128].

1.4 Tramtrack is a transcriptional repressor with numerous developmental roles

Tramtrack is a well-characterized transcriptional repressor and negative regulation of its target genes is essential for cell fate specification at several stages of the *Drosophila* life cycle. In order to understand the contribution of Tramtrack to muscle development and to explain its loss-of-function and gain-of-function muscle phenotypes, it is important to introduce its diverse developmental roles, target genes, and regulators in other developmental systems.

1.4.1 Molecular characteristics and embryonic expression of Ttk69 and Ttk88

1.4.1.1 Gene model of the *ttk* locus

The two Tramtrack protein isoforms, Ttk69 and Ttk88, are encoded by genomic sequence spanning over 20 kb on the right arm of the third chromosome. There are three different transcriptional start sites (TSSs), each of which contributes to two distinct mature mRNA transcripts encoding the two protein isoforms (Figure 1.4-1 A). The two mRNA transcripts are probably generated through alternative *trans*-splicing, a commonly observed phenomenon among TFs of the BTB (*broad*, *tramtrack*, *bric-a-brac*) protein family to which Tramtrack belongs [147].

As the precursor mRNA transcripts originating from the three TSSs differ only in their untranslated regions, there exist only two different protein coding sequences (Figure 1.4-1 B). The encoded proteins were named after their molecular weight Ttk69 (69 kDa) and Ttk88 (88 kDa). Both proteins have a BTB protein-protein interaction domain [148] at their amino-termini, but they differ at their carboxy-termini where one pair of C_2H_2 zinc finger domains resides [149].

1.4.1.2 DNA binding properties of Ttk proteins

Because of the different zinc fingers pairs, the two Ttk isoforms have different DNA binding preferences [150]. Ttk69 binding motifs were derived from DNAse I footprinting assays on the *fushi tarazu* (*ftz*) zebra element [151, 152], the *even skipped* (*eve*) promoter, the *eve* autoregulatory element [150], and on regulatory regions of

tailless (*tll*) [153]. Even though the identified sites slightly vary, they all share a TCCT (or AGGA) core motif that is essential for Ttk69 binding [153]. Ttk88, on the other hand, does not occupy the same sites and instead was found to protect a region upstream of the *engrailed* (*en*) locus which, in turn, is not recognized by Ttk69 [150]. The *in vitro* binding studies have thus shown that the differences in the amino acid sequence of the zinc finger domains are reflected at the DNA sequence level of bound sites (Figure 1.4-1 B).
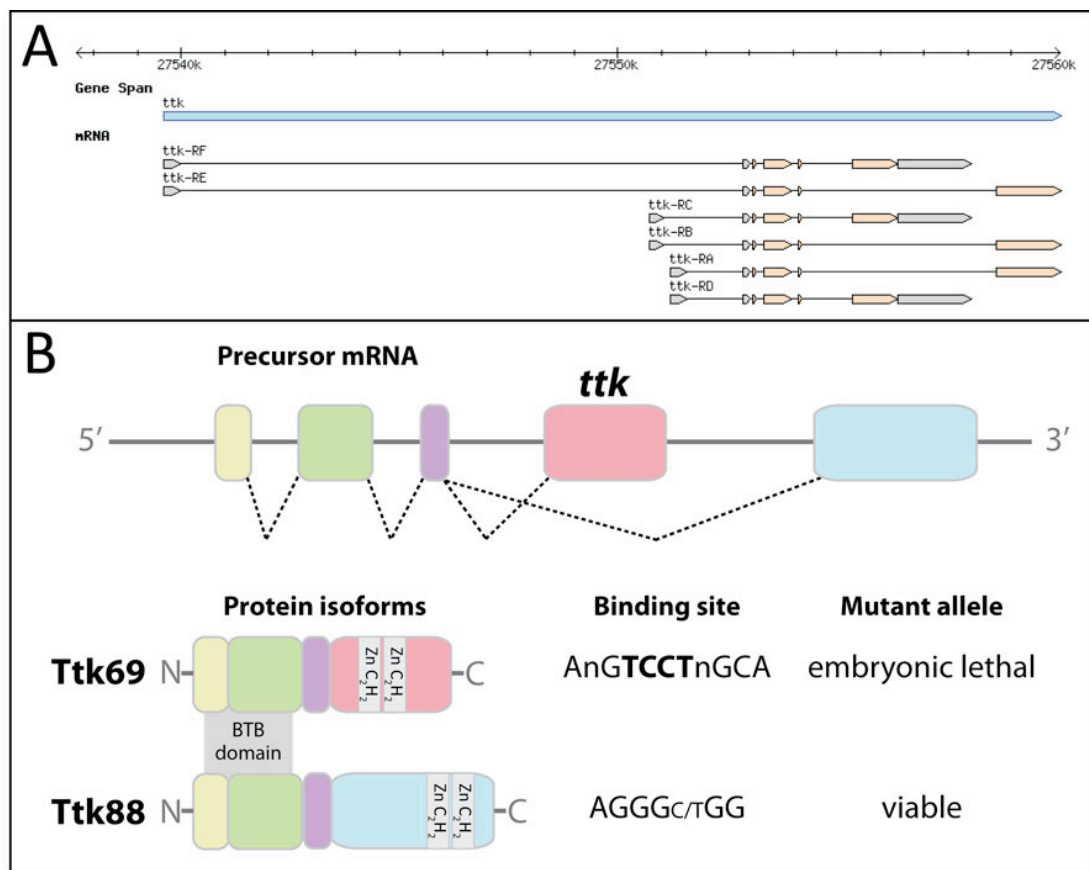
Figure 1.4-1: Gene model of the *ttk* locus and basic features of Ttk proteins

(A) Genomic plot of the *ttk* locus (in blue) with the different mRNA transcripts generated by alternative splicing below. Transcription from three TSSs gives rise to distinct transcripts that differ only in their non-coding regions (in grey) while the protein coding sequences are identical (in orange). Generated with GBrowse genome viewer [154].

(B) Schematic representation of *ttk* precursor mRNA (only exons are shown) and the two different encoded protein isoforms. Coloured boxes represent the exons and the dashed lines symbolize splicing events. The two different protein isoforms, Ttk69 and Ttk88, share the amino-terminus where a BTB domain is located but they differ in their zinc fingers at the carboxy-terminus. As a result, the two protein isoforms have different DNA binding sites and their deletion has distinct effects on viability. Modified from [155].

1.4.1.3 Binding partners of Ttk69 protein

One of the mechanisms of Ttk69 transcriptional repression has been elucidated in a yeast two-hybrid screen which uncovered a physical association between Ttk69 protein and Mi-2, a subunit of the histone deacetyletion and chromatin remodelling complex NuRD [156]. The region mediating the interaction lies in proximity of the zinc fingers and is specific to the Ttk69 isoform. Mi-2 and Ttk69 interact also genetically in the embryonic nervous system and co-localize on polytene chromosomes [156]. Recruitment of the chromatin remodelling machinery is therefore one way in which Ttk69 represses the transcription of its targets.

Another binding partner of Ttk69 is the general co-repressor C-terminal Binding Protein (CtBP). A well-characterized CtBP binding P-DLS motif [13, 157] is located in the carboxy-terminal portion of Ttk69 protein and its mutation interferes with the ability of Ttk69 to block neuronal development in eye precursors [158]. Physical interaction between Ttk69 and CtBP was confirmed in *in vitro* pulldown assays and the two genes are also linked genetically [158]. At least in the context of eye development, Ttk69 does not seem to interact with the co-repressor Sin3A or the histone deacetylase Rpd3, even though the BTB domain of the human B-cell lymphoma 6 (Bcl-6) protein interacts with their mammalian orthologues [159].

1.4.1.4 Embryonic expression of Ttk69 and Ttk88

The embryonic expression patterns of both Ttk isoforms at the mRNA and protein levels have been analysed in multiple previous studies [150, 152, 160] and are summarized in Table 1.4-1. There are only minor differences in the expression of the two isoforms and their protein expression recapitulates mRNA distribution (with a few exceptions that are discussed in the Section 1.4.3) [150]. There are two main phases of Ttk embryonic expression: a maternal phase, when Ttk69 and Ttk88 are supplied maternally and diminish before the onset of cellularization, and a zygotic phase, starting at stage 5 with expression of Ttk69 in the endoderm primordia. The zygotic expression however starts to peak at stage 11 when both Ttk isoforms are strongly expressed in trachea, derivatives of endoderm, epidermis, and nervous system. Ttk69 can also be detected in the mesoderm (Figure 4.2-1), but the strong signal from surrounding tissues obscures its mesodermal expression component, which has thus remained unidentified in most of the previous studies.

Tissue/cell type	Stage (st.)	Ttk69	Ttk88	Reference(s)
follicle cells	all stages of oogenesis	✔	✖	[161]
maternally deposited mRNA	syncytial blastoderm	✔	✔	[152]
yolk nuclei	st. 5 - 10	✔	✖	[150]
anterior endoderm	st. 5 - 6	✔	✖	[150]
midgut and its primordia	st. 7 - 17	✔	✔	[150, 152]
stomatodeum	st. 11 - 17	✔	✔	[150, 152]
trachea	st. 11 – 17	✔	✔	[150, 162]
glia in CNS	st. 12 – 17	✔	✖	[155]
support cells in PNS	from st. 11	✔	✔	[163]
epidermis	st. 11 – 17	✔	✔	[160]
ring gland	unspecified	✔	✔	[164]
visceral mesoderm	st. 11 – 17	✔	✔	[160], this study
somatic mesoderm	st. 11	✔	✔	this study

Table 1.4-1: Overview of embryonic sites of Ttk69 and Ttk88 expression
Green check marks indicate expression and red 'x' marks lack of expression. Examples of references where the given expression was reported are indicated.

1.4.2 Extensive developmental roles of Ttk69 and Ttk88

As the complex expression of Tramtrack proteins suggests, they play multiple roles during embryonic as well as later stages of *Drosophila* development (summarized in Table 1.4-2). Even though the expression patterns of the two protein isoforms are largely overlapping, their functions, as well as target genes, differ as a result of their different DNA binding preferences. While an allele deleting Ttk88 only is homozygous viable [165], alleles that delete Ttk69 are lethal at embryonic stages. Based on the isoform specific and shared phenotypes, the roles of Ttk69 and Ttk88 can be categorized into 4 different classes (Table 1.4-2) that are discussed in more detail below.

	Process	Stage	Targets	Ttk69	Ttk88
I	mitotic-to-endocycle transition of follicle cells	oocyte	*stg*, *rap*, *dap*, *ci*	[166, 167]	no
	polar cell specification			[168]	no
	chorion synthesis, dorsal appendage morphogenesis			[161]	no
II	maternal-zygotic transition; axis specification	syncytial blastoderm	*ftz*, *eve*, *tll*, *run*, *h*, *odd*	[160, 169]	no
			en	no	[170]
I	CNS glial cell development	embryo	*ase*, *dpn*, *scrt*, *repo*, *CycE*	[155, 171]	no
	trachea morphogenesis	embryo	*mmy*, *esg*, *pyd*, *bnl*	[162]	no
III	repression of neuronal fate in PNS	embryo		[163, 172]	[163]
		larva	*ac, sc, ase, CycE*	[163, 173, 174]	NA
			ac, sc, ase *	NA	[163, 174]
IV	repression of photoreceptor fate	larva	*lz*	[165, 175]	NA
	repression of R7 fate	larva	*en*	NA	[165, 176]
I	mitosis regulation in eye development	larva	*stg*	[177]	no
	maintenance of photoreceptor cell fate	late pupa		[178]	no

Table 1.4-2: Ttk69 and Ttk88 have overlapping as well as distinct developmental roles and target genes

A comprehensive summary of the described roles of Ttk69 and Ttk88 in developmental processes (column 1) during different stages (column 2). Known downstream genes are in the third column, in green colour are genes that have been asserted to be positively regulated. Those genes whose regulatory regions were experimentally confirmed to be bound by Ttk are underlined. Last two columns contain literature references as well as information on whether the role and/or targets are specific or common to the two Ttk isoforms (NA- not applicable).
Class I, Ttk69-specific roles; class II, shared role but different target genes; class III, shared role and target genes; class IV, different role in the same developmental process.
*repression of *ac*, *sc*, and *ase* by Ttk88 is weaker compared to Ttk69

1.4.2.1 Class I: Ttk69-specific roles in oogenesis, embryogenesis, and eye development

There are multiple developmental roles that have been assessed as unique to the 69 kDa isoform of Tramtrack. Some of these are linked to maternally supplied Tramtrack69 protein that is expressed in the follicle cells throughout oogenesis. There it contributes to the control of the transition between mitosis and endocycle [166, 167], the number of polar cells connecting the single egg chambers [168], and chorion synthesis and morphogenesis of dorsal appendages [161]. Differentiation and cell division of the follicle cells surrounding the oocyte are regulated by Notch signalling (reviewed in [126]). The levels of Ttk69 protein in *Notch* mutant follicle cells are reduced and the cells also undergo prolonged mitosis similar to a *ttk* mutant [166]. Ttk69 is however not simply an effector of Notch signalling as the differentiation associated defects in *Notch* mutants are not phenocopied in *ttk* mutants [166]. Ttk88 is not required in oogenesis as it is, unlike Ttk69, absent from the follicle cells and the mutant phenotypes were observed in Ttk69-specific alleles *ttk^{1e11}* and *ttk^{twk}* [161, 166, 167].

Ttk69, but not Ttk88, is also expressed in all glia cells of the embryonic central nervous system (CNS) and controls their number as well as proper differentiation by repression of the default neuronal program [155]. Its principal targets within the CNS are the pan-neural genes *asense* (*ase*), *deadpan* (*dpn*), and *scratch* (*scrt*) [171]. In addition, ectopically expressed Ttk69 downregulates the expression of the cell cycle regulator *Cyclin E* (*CycE*) and in *ttk* mutant embryos the number of glial cells is increased [171]. Therefore, similar to its role in oogenesis, Ttk69 is also involved in control of proliferation as it prevents supernumerary rounds of cell division in the CNS.

Another example of a role restricted to Ttk69 is the orchestration of multiple steps of late trachea development. Unlike in the above-mentioned cases, Ttk88 also shows tracheal expression, yet mutants deleting or misexpressing Ttk88 isoform develop normal trachea [162]. *ttk^{D2-50}* homozygous embryos (affecting *ttk69* and possibly *ttk88*) express reduced levels of *branchless* (*bnl*), the ligand of the Fibroblast growth factor (FGF) receptor *breathless* (*btl*), both of which together mediate tracheal migration and branching [179]. One of the sources of *bnl* is the visceral mesoderm, where it is co-expressed with Ttk69 and downregulated in *ttk*-deficient embryos,

suggesting a positive input of Ttk69 into *bnl* regulation [162]. Ttk69 is also involved in later steps of trachea development when subtle differences between its expression levels in presumptive fusion cells and adjacent cells are essential for their subsequent fusion [162]. In addition, Ttk69 is required for complex cell shape changes and rearrangements that are necessary for tubule and intracellular lumen formation. In *ttk* mutants, intercalation of trachea cells is impaired due to downregulation of *polychaetoid* (*pyd*), an adherens junction associated protein necessary for this process [180]. Altogether Ttk69 acts at multiple stages of trachea differentiation both positively and negatively as a regulator of complex morphogenetic events. Two targets, *bnl* and *pyd*, were suggested to be activated by Ttk69 [162]; however, it should be noted that this assertion is based only on their misexpression phenotypes in *ttk* loss- and gain-of-function scenarios, and no evidence for their direct activation by Ttk69 (e.g. binding of Ttk69 to their regulatory regions, reduced expression upon mutagenesis of Ttk69 sites) has been reported.

One more case where Ttk69 (but not Ttk88) was suggested to play a positive regulatory role (presumably indirectly) are the late differentiation events of eye development. Ttk69 is the only Tramtrack isoform expressed in all photoreceptor cells during pupal stages and it is required for their proper terminal differentiation [178]. Ttk69, however, represses the development of photoreceptor neurons earlier in development [165], hence it would be interesting to know what causes the switch in its function and what target genes Ttk69 regulates at these two stages.

1.4.2.2 Class II: Both Ttk69 and Ttk88 are involved in embryonic axis specification but they regulate different target genes

The earliest studied role of Tramtrack that also led to its identification is the repression of pair rule genes in early *Drosophila* embryos. A series of classic genetic experiments showed that ectopic expression of Ttk69 through heat shock inducible promoter represses, albeit to a variable extent, the expression of the pair rule genes *ftz*, *eve* [160, 181], *run* (*runt*), *hairy* (*h*) and *odd skipped* (*odd*) [160], but has no effect on the expression of the upstream gap genes *hunchback* (*hb*), *Krüppel* (*Kr*) [160, 181], or *giant* (*gt*) [181]. Binding of Ttk69 protein to the regulatory regions of *ftz* and *eve* has been extensively characterized in *in vitro* experiments [150-152]. Upon deletion of the Ttk69 binding sites in the enhancer of *ftz*, the expression of a linked reporter gene

initiated prematurely and was not resolved in the typical striped pattern along the anterior-posterior embryonic axis [151].

On the other hand, the Ttk88 isoform does not cause misexpression of any of the above mentioned pair rule genes [181], but instead binds the promoter region of the segment polarity gene *en* [150] and, in cooperation with the pair rule protein Runt, establishes its stereotypic expression in 14 parasegmental stripes [170]. Therefore, even though both maternally supplied Ttk isoforms play a role in axis specification, they appear to do so by regulating distinct sets of downstream genes.

The levels of maternally supplied Ttk proteins diminish before the onset of zygotic transcription and their gradual dilution is thought to be one mechanism by which the maternal-zygotic transition in *Drosophila* is orchestrated. If the levels of maternal Ttk69 protein are increased, the onset of *ftz* expression is delayed, whereas in *ttk* mutants it initiates prematurely [169]. According to a suggested model, zygotic transcription does not initiate before the 8th nuclear cleavage as it is actively repressed by maternal Ttk69 protein [169]. With subsequent nuclear division cycles the nucleocytoplasmic ratio gradually increases and so the effective concentration of Ttk69 protein decreases, permitting transcription of its targets. However, there must be more such maternally loaded repressors as, consistently with the previous results, altering the levels of Ttk69 does not influence initiation of *Kr* expression [169].

1.4.2.3 Class III: Shared role and targets of both Ttk isoforms in the peripheral nervous system

One of the most extensively characterized roles of Ttk is repression of the neuronal cell fate in the developing peripheral nervous system (PNS) during embryonic and larval stages. Ttk proteins act as cell fate regulators in at least two types of *Drosophila* sensory organs, the chordotonal and the external sensory organs [163]. In wild type embryos and adults, these consist of one neuronal cell and three or four, respectively, support cells of distinct identities. The different cell types originate via a series of asymmetric cell divisions from a single sensory organ precursor (SOP). For example, in the case of external sensory organs, the SOP divides into IIa and IIb progenitor cells that further divide to produce different types of support cells and a third progenitor, IIIb, which will give rise to the neuronal cell and yet another support cell (Figure 1.4-2 B). The bristles visible on the notum of an

adult fly are therefore formed by one shaft and one socket cell (both support cells), while the neuron and two other support cells are positioned below the cuticle (Figure 1.4-2 A).

In *ttk* mutants, there are more neurons formed at the expense of support cells as a result of two IIb precursors being formed instead of one IIa and one IIb precursor cell [163]. At a phenotypic level, this change in cell identity is reflected in a so-called "bald" phenotype where the external support cells on the notum are missing and thus no bristles are visible (Figure 1.4-2 C). In contrast, when the levels of Ttk are increased, the SOP divides into two IIa precursor cells. Twice as many shaft and socket cells are formed, giving rise to a double-bristle phenotype (Figure 1.4-2 D). These cell fate changes resemble those in *Notch* loss- or gain-of-function mutants, respectively [182], and it has been established that Ttk functions downstream of *numb* and executes the developmental program directed by *Notch* signalling only in the *numb*-negative daughter cell [163].

The repression of neuronal identity by Ttk in the PNS lineage was also confirmed by molecular studies in larval wing imaginal discs, where the specification of adult sensory organs is initiated [174]. These experiments also extended the potential of Ttk as a neuronal fate repressor as ectopic expression of Ttk69, and to a lesser extent Ttk88, prior the formation of SOPs leads to their ablation and almost complete absence of wing sensory organs [174]. The targets of Ttk69 and Ttk88 responsible for this phenotype are the proneural genes *achaete* (*ac*), *scute* (*sc*), and *asense* (*ase*), albeit their repression by ectopic Ttk88 is milder compared to Ttk69. In addition, ectopic expression of Ttk69, but not Ttk88, reduces the size of the wing due to a block in cell proliferation [174]. Given that many of the Ttk69-specific roles are linked to cell cycle regulation, it is likely that the repression of cell cycle progression is a general feature of Ttk69, but not Ttk88.
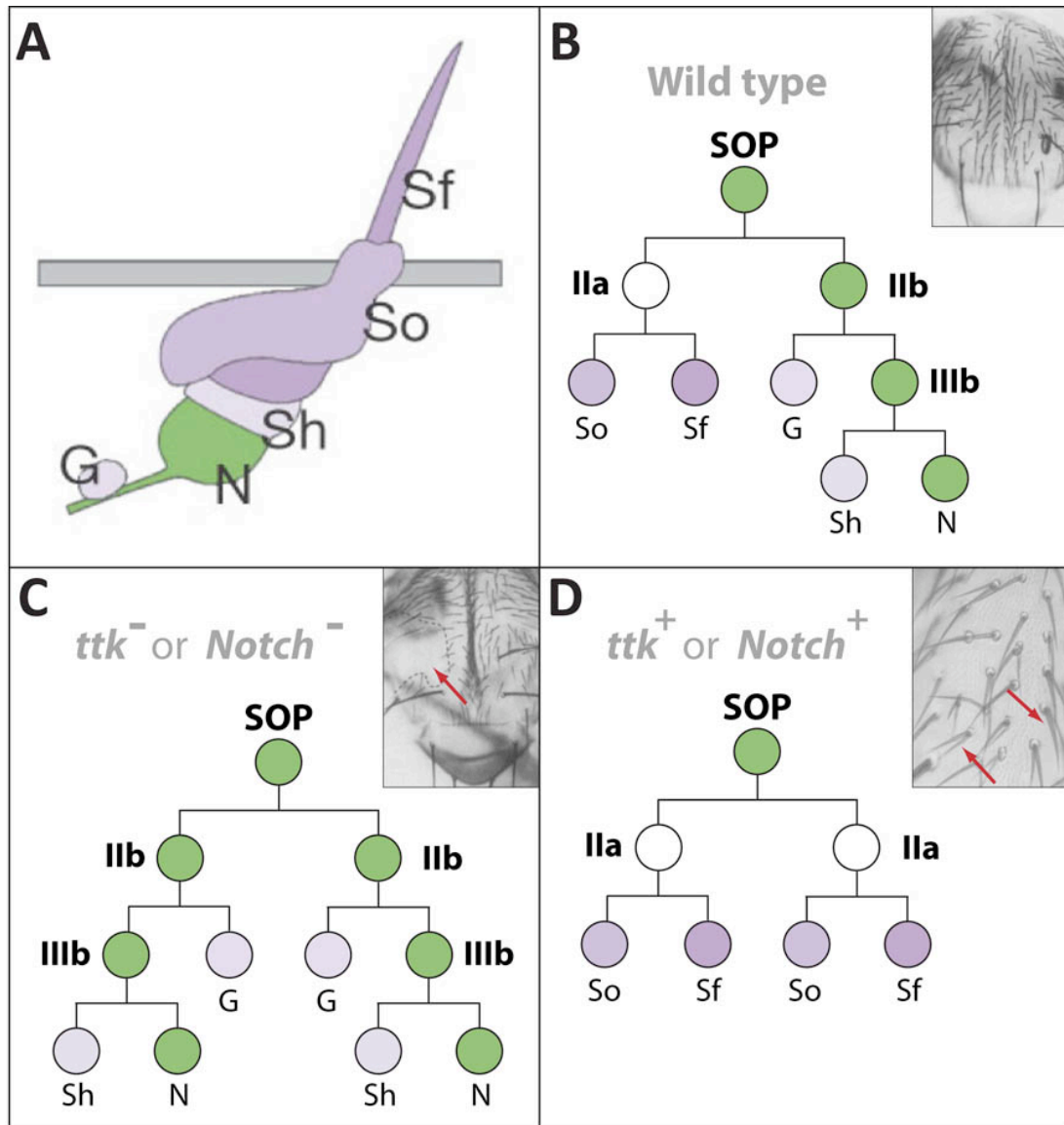
Figure 1.4-2: Cell fate changes in the PNS of *ttk* loss- and gain-of-function backgrounds

(A) Schematic representation of cell composition of one external sensory organ on the adult's notum. Sf, shaft cell; So, socket cell; Sh, sheath cell; N, neuron; G, glia. Adapted from [183].

(B) Scheme of cell divisions in the lineage of an external sensory organ in wild type adults. Inset shows the notum of a wild type fly. Modified from [183] and [163].

(C) In *ttk* or *Notch* mutants, the aberrant cell divisions produce more neurons, glia and sheath cells at the expense of socket and shaft cells. As a result, the notum of a *ttk* mosaic fly (in the inset; the *ttk*-deficient area is outlined by a dashed line and red arrow) appears "bald" as the outer shaft and socket cells are missing (red arrow). Modified from [183] and [163].

(D) Ttk69 overexpression or constitutively active Notch signalling lead to lack of neurons, but twice as many socket and shaft cells formed, giving rise to a so-called double-bristle phenotype (in the inset; the red arrows point to doubled bristles on a notum of a Ttk69-overexpressing fly). Modified from [183] and [163].

1.4.2.4 Class IV: Ttk69 and Ttk88 play divergent roles in eye development

The compound eyes of *Drosophila* adults are composed of approximately 800 ommatidia, each formed by six outer (R1-R6) and two centrally positioned (R7, R8) photoreceptor neuronal cells surrounded by non-neuronal support cells. If the *ttk88* isoform is deleted, the average number of outer photoreceptor cells per ommatidium decreases, some ommatidia have more than one R7 cell, and the number of R8 cells remains unaffected [165]. It is however unlikely that the ectopic R7 cells were formed at the expense of R1-R6 cells as many ommatidia with increased number of R7 cells had wild type number of R1-R6 cells. In *ttk88* mutant larval eye discs, the levels of En protein are increased and in some cases its spatial expression pattern is affected as well [165]. The mechanism of Ttk88 action in R7 fate suppression seems to be novel as it acts independently of *seven in absentia* (*sina*), a major positive regulator of R7 determination [176]. As in the nervous system, Ttk88 acts as a repressor of neuronal cell fate in the developing eye and elaborate regulation is required to eliminate its function and enable the initiation of differentiation (see Section 1.4.3.3).

Removal of Ttk69, on the other hand, interferes with specification of all cell types and leads to absence of ommatidial tissue as the ommatidia and sensory bristles are not formed [165]. The severe loss-of-function phenotype can be partially explained by Ttk69 regulating *lozenge* (*lz*) expression, encoding a transcriptional regulator of genes essential for the differentiation of R1/6/7 neurons and cone cells [184]. Combining *ttk69* and *lz* mutant alleles attenuates the severity of the *ttk* phenotype and the formation of ectopic R7 neurons in *ttk* mutants is *lz*-dependent [175]. Expression levels of *lz* are responsive to changes in Ttk69 levels and as there are multiple putative Ttk69 binding motifs in its upstream regions [184], it is likely to be directly repressed by Ttk69.

In the pool of precursor cells that will form the adult eye, Ttk69 directly represses transcription of *string* (*stg*) [177], a Cdc25 homologue that is essential for cell cycle progression from G2 to mitosis. Ttk69 competes for binding with the activator Pointed and it was suggested that the ratio between the levels of the two proteins and/or changes in the affinity of their binding influences the regulatory trajectory [177]. According to a proposed model, Ttk69 keeps the precursor cells in the eye disc in an arrested state, but when Pointed is activated during the second

mitotic wave, it can compete Ttk69 away from *stg* regulatory regions and initiate mitosis.

Consequently, it appears that even though both Ttk isoforms are required for proper eye development, they exert different roles, possibly by regulating distinct sets of target genes. These however remain to be identified, as the knowledge about genes downstream of Ttk69 and Ttk88 in the eye is currently very limited.

1.4.3 Dynamic expression of Ttk69 and Ttk88 is achieved by diverse regulatory mechanisms

The complex and dynamic expression pattern of Tramtrack is accomplished through extensive regulation at multiple tiers. Currently, there is only limited information on transcriptional regulation of Ttk, but there is a wealth of data reporting regulation of Ttk at post-transcriptional and post-translational levels (summarized in Figure 1.4-3).

1.4.3.1 Post-transcriptional regulation of *ttk69* mRNA

Ttk69 protein, but not *tt*k69 mRNA, is localized to only one of two daughter cells generated in Notch-mediated asymmetric cell division in the PNS (see Section 1.4.2.3). Only the *numb*-positive daughter cell expresses, through unknown means, an RNA binding protein Musashi (Msi) that directly binds a sequence within the 3' untranslated region (UTR) of *ttk69* mRNA [183]. Via this sequence, Musashi directly represses translation of *ttk69* mRNA in cell culture and also negatively regulates synthesis of Ttk69 protein *in vivo* [183]. As a result, Ttk69 protein is present only in the *numb*-negative daughter cells where it represses the neuronal cell fate.

The 3' UTR of Ttk69 was also computationally predicted and then experimentally verified to be targeted by two microRNAs, miR-92b and miR-312 [185]. Whether these microRNAs reduce the levels of Ttk69 protein also *in vivo* is unknown. Another example of microRNA-mediated downregulation of Ttk69 protein levels was described to be active during early *Drosophila* embryogenesis. *miR-184* mutant phenocopies overexpression of Ttk69 as expression of the pair rule genes, but not gap genes, is delayed [186]. In a quantitative assay, the levels of Ttk69 protein were found to be reduced in the *miR-184* mutant and both genes interact genetically as

well [186]. Repression of Ttk69 by maternally expressed miR-184 might therefore contribute to timely depletion of the Ttk69 protein to allow zygotic transcription to proceed.

1.4.3.2 Covalent modifications of Ttk proteins in the developing eye

During eye development, Ttk88 inhibits differentiation of R7 photoreceptor cells and activation of the differentiation program by the RAS/MAPK signalling cascade releases this block by decreasing the levels of Ttk88 protein. The mechanism behind Ttk88 downregulation has been elucidated in genetic as well as biochemical studies, which linked the proteins Sina and Phyllopod (Phyl) to Ttk88 degradation. Sina via its E3 ubiquitin ligase activity, ubiquitinates Ttk88 and hence targets it for proteasome-mediated degradation [187]. Consistently, Ttk proteins contain stretches of PEST sequences [152] that are indicative of a short half-life and are often associated with proteasome-dependent proteolysis [188, 189]. However, the physical interaction between Ttk and Sina alone is unstable and Phyl, an adaptor protein whose expression is transcriptionally induced by RAS/MAPK signalling, is required to stimulate the ubiquitination activity of Sina [190]. The region of Ttk that interacts with Phyl resides in the BTB domain that is common to both Ttk isoforms and the interaction is highly specific as closely related BTB domains of other *Drosophila* proteins do not associate with Phyl [190]. Genetic evidence supports the findings from *in vitro* experiments as overexpression of Phyl can counteract the block in photoreceptor differentiation caused by ectopic Ttk88 only in the presence of Sina [191]. An additional layer of Ttk control is brought by Msi, a negative regulator of *ttk69* mRNA translation in the PNS [183]. Msi genetically interacts with Sina in the eye imaginal disc in a redundant fashion as derepression of Ttk69 protein is more severe in *msi sina* double mutant compared to single mutants [192].

A recently identified novel homologue of Sina, named SinaH, specifically recognizes a sequence unique to the Ttk69 isoform [193]. Similarly to Sina, it targets Ttk69 for proteolysis and it also requires an adaptor protein whose identity remains to be discovered. In the presence of Phyl, SinaH can also interact with full-length Ttk88 and Ttk69 through the BTB domain, yet the recognition of the Ttk69-specific motif seems to be more efficient [193]. The importance of regulating the protein levels of each Ttk isoform specifically is underscored by the finding that Ttk69 protein, unlike Ttk88, is a target of the enzyme Lesswright (formerly Ubc9) that covalently

conjugates an ubiquitin-like molecule Smt3 [194]. Ttk69 co-localizes with Smt3 in certain PNS precursor cells as well as on chromosomal sites in salivary-gland polytene chromosome spreads [194]. However, the effect of the conjugated moiety on Ttk69 levels or function is unknown.

1.4.3.3 Additional mechanisms to regulate Ttk69 protein function

Another modification of Ttk69 that was detected in uncellularized *Drosophila* embryos is phosphorylation by MAP kinases [153]. The ability of MAP kinases to phosphorylate Ttk69 was confirmed only *in vitro* and so its existence and function *in vivo* remain to be investigated.

The Ttk69 isoform has been linked to cell proliferation in multiple developmental contexts, where it blocks mitosis. In the embryonic CNS it contributes to the regulation of glial cell number by repressing the S phase promoting factor *CycE* [171]. It is therefore not surprising that the expression of Ttk69 protein itself oscillates with cell cycle progression. During S phase, when DNA is replicated, Ttk69 protein is expressed in lateral glia at undetectable levels but its expression increases once the cells exit S phase [171]. Expression of Ttk69 therefore seems to be incompatible with DNA replication, possibly also in other cell types as ectopic Ttk69 was found to interfere with DNA replication in wing imaginal discs as well [174].
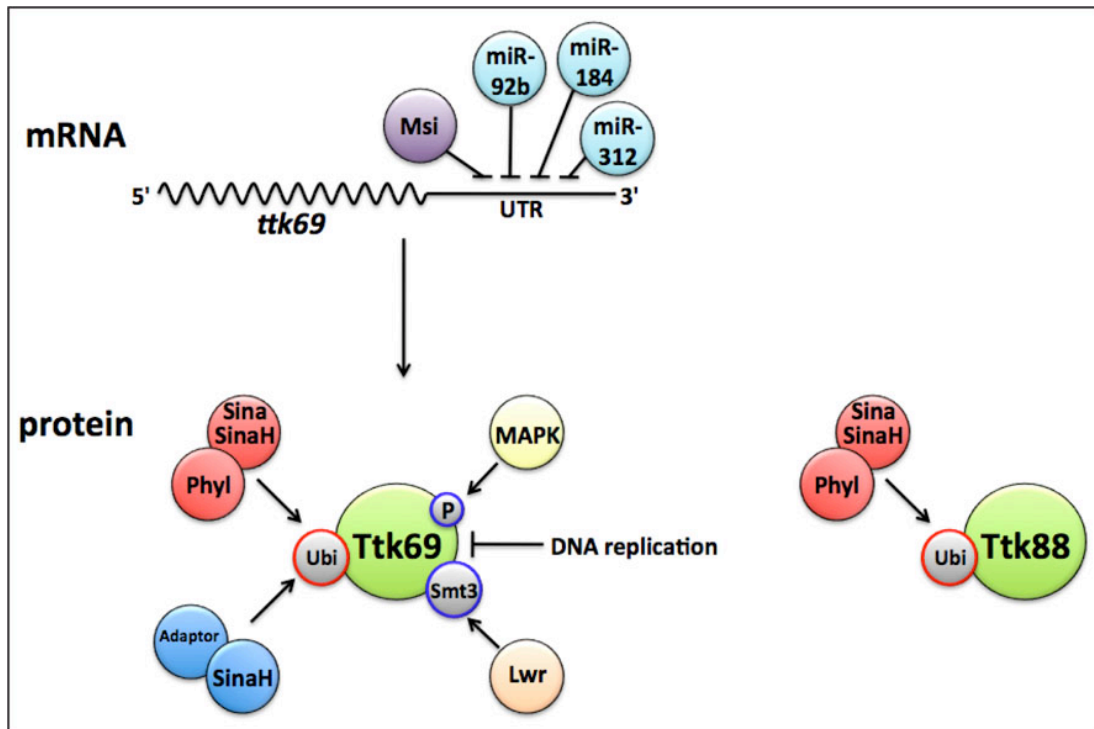
Figure 1.4-3: Mechanisms of Ttk69 and Ttk88 mRNA and protein control

Schematic summary of all known regulators of Ttk69 and Ttk88 mRNA and protein (not drawn to scale). *ttk69* mRNA is via its 3' UTR repressed by RNA-binding protein Musashi (Msi) and at least three different microRNAs. Ttk69 protein is ubiquitinated (Ubi) by two different E3 ligases, Sina and SinaH, both requiring an adaptor protein for Ttk69 binding (Phyl for Sina, an unknown protein for SinaH). Ubiquitination enhances degradation of the protein (red border). Ttk69 is in addition modified by phosphorylation by MAPK (P-phosphate group), and SUMO-like moiety Smt3 mediated by Lwr. The effect of both modifications on Ttk69 levels or function is unknown (blue border). Ttk69 protein levels are also reduced during DNA replication. Ttk88 is only known to be subjected to ubiquitination by Sina/Phyl and SinaH/Phyl complexes.

2 AIMS OF THE THESIS

To date, approximately 50 of the 700 putative *Drosophila* transcription factors have been implicated in embryonic muscle development. However, several experimental observations indicate that this number is greatly underestimated. Similarly, genome-wide studies have identified thousands of *cis*-regulatory modules (CRMs) occupied by the core myogenic transcription factors, yet our knowledge about the diverse regulatory inputs that they integrate remains limited. My thesis aimed to bridge this gap by systematically identifying novel myogenic regulators that would contribute to the underlying regulatory network.

The vast majority of known mesodermal transcription factors regulate transcription positively. However, studies in other developmental processes revealed that repression in developmental decision-making is at least as important as transcriptional activation. A key aim of my thesis was therefore to understand the contribution of negative regulatory inputs to mesodermal gene expression. One such transcription factor, Ttk69, was implicated in my screen to be a new potential negative regulator of mesoderm development. This was rather surprising, given the established role of Ttk69 as an essential regulator of neuronal development. I used a combination of genetic and genomic approaches to characterise the role of Ttk69 in muscle development, and specifically addressed the following questions: At which particular step of myogenesis is Ttk69 involved? What are the downstream targets of Ttk69? What is the regulatory relationship between Ttk69 and other myogenic factors? Where is Ttk69 positioned within the muscle regulatory network? Is Ttk69 a dedicated repressor or could it potentially activate transcription as well?

Many developmental regulators have the capacity to both activate and repress transcription. Similar to Ttk69, Snail has been traditionally perceived as a dedicated repressor, yet several lines of evidence suggested that it could contribute to transcriptional activation. Exploring the possibility that Snail is a bimodal transcription factor regulating mesoderm development was therefore another objective of my thesis.

By seeking answers to these questions, I hope that my thesis will contribute to our understanding of the regulatory logic driving mesoderm development and the general complexity of metazoan gene expression.

3 MATERIALS AND METHODS

3.1 MATERIALS

3.1.1 Instruments

Name	Vendor
10x ocular (35mm, 2.5x, 4"x5")	Zeiss
10x/0.25 A-Plan Objective	Zeiss
20x/0.50 Plan-Neofluar Objective	Zeiss
63x/1.4 Plan-Apochromat Objective	Zeiss
ABI PRISM 7500 Sequence Detection System	Applied Biosystems
Analytical scales, AE50	Mettler
Axiophot Light Microscope	Zeiss
Bioruptor UCD-200TM ultrasonic cell disruptor water bath	CosmoBio
Centrifuges 5415D, 5417C, 5810	Eppendorf
Dounce tissue grinder (homogeniser; 15ml)	Wheaton
Gel electrophoresis chamber Mini-Sub Cell GT	BioRad
Gel electrophoresis chamber Wide Mini-Sub Cell GT	BioRad
LSM 510 FCS	Zeiss
LSM 510 META	Zeiss
Nanodrop Spectrophotometer ND-1000	Nanodrop Technologies
Shaker table Gyrotory (Model G2)	New Brunswick Scientific
Shaker table Nutator (220V)	Adams
Sieves (0.125mm; 0.122mm; 0.335mm; 0.71mm)	VWR or Buddeberg/Quadrolab
Speed-Vac Concentrator 5301	Eppendorf
Stereo Microscope Stemi SV6	Zeiss
Stereo Microscope Lamps KL 1500 LCD	Schott
Thermal Cycler PTC-200 Multicycler DNA engine	MJ Research Peltier
Thermal Cycler PTC-225 DNA Engine Tetrad	MJ Research Peltier
VICTOR 1420 Luminescence Counter	Perkin Elmer
Water bath Thermomix 5BU	B.Braun Biotech International
Water bath GD100	Grant
Water bath MP	Julabo

3.1.2 Chemicals and reagents

Name	Catalogue number	Vendor
1 kB DNA ladder	N3232L	New England Biolabs
100 bp DNA ladder	N3231S	New England Biolabs
4-Nitro blue tetrazolium chloride (NBT)	11383213001	Roche
5-Bromo-4-chloro-3-indolyl-phosphate (BCIP)	1383221	Roche
Agarose (LMP)	15517-022	Gibco
Ampicillin	A9518	Sigma
AmpliTaq® DNA polymerase	N8080171	Applied Biosystems
Aprotinin	A1153	Sigma
Biorad protein assay	500-0001	Biorad
Biotin (BIO) RNA Labeling mix	11685597910	Roche
Boric acid (99.5%)	B7660	Sigma
Bovine Serum Albumine (BSA, fraction V)	A-7906	Sigma
Cellfectin	10362-010	Invitrogen
Chloroform	C2432	Sigma
DAB substrate	1718096	Roche
DEPC treated water	9920	Ambion
DFS-Taq DNA polymerase	101100	Bioron
Digoxigenin (DIG) RNA labeling mix	11277073910	Roche
DNaseI	0 776785	Roche
DMSO	8.02912.10	Merck
Deoxynucleoside-5'-triphosphate (dNTPs) for PCR	110002	Bioron
Deoxyribonucleic acid, single stranded from salmon testes	D9156	Sigma
Dual-Luciferase® Reporter Assay System	E1960	Promega
EDTA	E6758	Sigma
EGTA	E3889	Sigma
Ethanol	1.00983.2500	Merck
Ethidium bromide	E-1385	Sigma
Fluorescein (FITC) RNA labelling mix	11685619910	Roche
Formamide	F5786	Sigma
Formaldehyde (16%, methanol)	18814	Polyscience Europe
Glycogen	901393	Roche
Glycerol	4043-00	J.T.Baker
Hydrogen peroxide	H1009	Sigma

Name	Catalogue number	Vendor
Leupeptin	L2884	Sigma
Lithium chloride	L9650	Sigma
MEGAscript® T7 Kit	AM1334	Ambion
Methanol	1.06009.2500	Merck
n-heptane	H9629	Sigma
N-Laurylsarcosine	D-6750	Sigma
Normal Goat Serum (NGS)	S-1000	Vector Labs
NP 40	13021	Sigma
2-propanol	1.09634.2500	Merck
Pepstatin	P5318	Sigma
Phenol:Chloroform:Iso-amyl-alcohol	9732	Ambion
Phenyl methyl sulfonyl fluoride (PMSF)	P7626	Sigma
Protein-A Sepharose® 4B	P9424	Sigma
Proteinase K	745723	Roche
QuickChange® site-directed mutagenesis kit	200523	Stratagene
RNase A	1006693	Qiagen
Sodium acetate	9740	Ambion
Sodium deoxycholate	D6750	Sigma
Sodium chloride	1.06404.5000	Merck
Sodium dodecyl sulfate	L6026	Sigma
Sp6 RNA Polymerase	M0207s	New England Biolabs
SybrGreen® PCR Master Mix	4309155	Applied Biosystems
Pfu DNA polymerase		EMBL
ProLong® Gold antifade reagent	P36931	Invitrogen
Restriction endonucleases		New England Biolabs
Ribonuclease inhibitor	15518-012	Invitrogen
T3 RNA polymerase	1 031 163	Roche
T4 DNA ligase	799099	Roche
T7 RNA polymerase	881 767	Roche
Tris-base	T6791	Sigma
Triton-X-100	T8787	Sigma
TSA™ Plus Cyanine 3 & Cyanine 5 System	NEL752001 KT	Perkin Elmer
TSA™ Plus Cyanine 3 & Fluorescein System	NEL753001 KT	Perkin Elmer
Yeast tRNA	15401-029	Invitrogen
Vectastain Elite ABC Kit	PK-611	Vector Laboratories

Name	Catalogue number	Vendor
Western blocking reagent	11921673001	Roche

3.1.3 Miscellaneous materials

Name	Catalogue number	Vendor
ABIprism 96-well optical reaction plates	4306737	Applied Biosystems
AeroDuster 100		Servisol
Brown microcentrifuge tubes (1.5 ml)	1-6180	NeoLab
Brushes (various sizes)	9.172.050	Buddeberg
GeneChip Drosophila Tiling 1.0R Array	900588	Affymetrix
QIAprep Spin Miniprep kit	27106	Qiagen
QIAquick PCR Purification kit	28106	Qiagen
QIAquick Gel Purification kit	28706	Qiagen
Parafilm	PM-996	Pechiney
PCR tubes (0.5 ml, thin walled)	0030 124.502	Eppendorf
PCR tubes (0.2 ml, thin walled)	0030 124.332	Eppendorf
Phase-lock heavy gel tubes (2 ml)	0032-005-152	Eppendorf
RNeasy Mini kit	74104	Qiagen
Safe-Lock Tubes 1.5 ml	28004	Eppendorf
Safe-Lock Tubes 2.0 ml	0030 120.094	Eppendorf
SafeSeal Micro Tubes 1.5 ml	72.706.200	Sarstedt
Siliconized microcentrifuge tubes (1.6 ml)	710176	Biozym

3.1.4 Oligonucleotides

Name Sequence	Stock number	Restrict. site	Purpose
WntD-F GCGGTACCATAGCCTGCAAATCCCAAGCC	1141	KpnI	Cloning *WntD-lacZ* into pGL3
WntD-R GCGAGATCTGCAACAATACTGGCAGTTCCC	1148	BglII	Cloning *WntD-lacZ* into pGL3
Tin-F GCGGTACCTCAAGCGTTGAGCGTTGAGC	1149	KpnI	Cloning *Tin B-374* into pGL3

Name Sequence	Stock number	Restrict. site	Purpose
Tin-R CGAGATCTTGCGGGAAAGCAGGAAAATGG	1150	BglII	Cloning *Tin B-374* into pGL3
Mef2-F GCGGTACCCTGTAAAAATCACGCATAACCG	1157	KpnI	Cloning *Mef2 I-D[L]* into pGL3
Mef2-R CGAGATCTCCTGAAGAAACCCCTGCCAAG	1164	BglII	Cloning *Mef2 I-D[L]* into pGL3
Mdr49-F GCGGTACCGCAACAAAGTCGATCGTATAAC	1209	KpnI	Cloning *Mdr49 early mesoderm* into pGL3
Mdr49-R GCAGATCTCAGATTTCCTCTCGGTTAGTC	1216	BglII	Cloning *Mdr49 early mesoderm* into pGL3
sna-F AGCGGCCGCGGCCGCCAACTACAAAAG	1372	NotI	Cloning *sna* CDS into pAc5.1 A
sna-R CGCGTCTAGAGCACCAAAACCGAATCGACTTAG	1373	XbaI	Cloning *sna* CDS into pAc5.1 A
dl-F ATGGTACCCATGGATGTTTCCGAACCAG	1374	KpnI	Cloning *dl* CDS into pAc5.1 A
dl-R CGTCTAGATGGTTCGTTGTGAAAAGGTAT	1376	XbaI	Cloning of *dl* CDS into pAc5.1 A
Mef2-23-F CTGCATGTTGCATGCACTCATCACATGTGCAATACTCGGCATCTGCGGCAGTAGC	1888	-	Mutagenesis of Sna2 and Sna3 in *Mef2 I-D[L]*
Mef2-23-R GCTACTGCCGCAGATGCCGAGTATTGCACATGTGATGAGTGCATGCAACATGCAG	1889	-	Mutagenesis of Sna2 and Sna3 in *Mef2 I-D[L]*
Mef2-1-F GCGACGTACGGTTGATGCTGAGTATTGCATGCACTCATCACATG	1890	-	Mutagenesis of Sna1 in *Mef2 I-D[L]*
Mef2-1-R CATGTGATGAGTGCATGCAATACTCAGCATCAACCGTACGTCGC	1891	-	Mutagenesis of Sna1 in *Mef2 I-D[L]*
Tin-123-F GATCCCTTCTGGGCTGGTTGACATGTGAGATTCTCATGTGAGGACCGCCGCACAGGGGCG	1892	-	Mutagenesis of Sna1, 2 and 3 in *Tin B-374*
Tin-123-R CGCCCCTGTGCGGCGGTCCTCACATGAGAATCTCACATGTCAACCAGCCCAGAAGGGATC	1893	-	Mutagenesis of Sna1, 2 and 3 in *Tin B-374*
Tin-4-F CTTTGTATTGCCTTGTTTCTAACTTCTCGCTTTTCCATTTTCCTGCTTTCC	1902	-	Mutagenesis of Sna4 in *Tin B-374*
Tin-4-R GGAAAGCAGGAAAATGGAAAAGCGAGAAGTTAGAAACAAGGCAATACAAAG	1903	-	Mutagenesis of Sna4 in *Tin B-374*

Name Sequence	Stock number	Restrict. site	Purpose
lacZ-F CGACTGATCCACCCAGTCCC	1842	-	Mapping Δ*MED24* by single-embryo PCR
lacZ-R GCGATGTCGGTTTCCGCGAG	1843	-	Mapping Δ*MED24* by single-embryo PCR
pos-F CGATGACACTATCGCAGTTACATCC	1844	-	Mapping Δ*MED24* by single-embryo PCR
pos-R CTGGTTTTAAGTTGGAATTTAGAAAGAAC	1845	-	Mapping Δ*MED24* by single-embryo PCR
msk-F CATCCACGATCGCAAACTTTG	1846	-	Mapping Δ*MED24* by single-embryo PCR
msk-R ACCGATTGTCGCAGCATACAT	1847	-	Mapping Δ*MED24* by single-embryo PCR
Med24-F GCTGTGCACCAACAAGGTCCAG	1848	1848	Mapping Δ*MED24* by single-embryo PCR
Med24-R GGCAGGTGACCAGCATCGACT	1849	-	Mapping Δ*MED24* by single-embryo PCR
CG7375-F AGCACAGCCAAACGCGGAAG	1948	-	Mapping Δ*MED24* by single-embryo PCR
CG7375-R CCACCACGCATCGCCTTCTT	1949	-	Mapping Δ*MED24* by single-embryo PCR
Atg18-F CCAATGCAGTTGTTCCGCGTAAAT	1950	-	Mapping Δ*MED24* by single-embryo PCR
Atg18-R AACGAGGAAGAATCATCGCTGACA	1951	-	Mapping Δ*MED24* by single-embryo PCR
CG8005-F CGATTGGGCTCCCCTTGAGC	1952	-	Mapping Δ*MED24* by single-embryo PCR
CG8005-R GAGCCTCCGCAGATCGGACA	1953	-	Mapping Δ*MED24* by single-embryo PCR
Med24-1-F CCTCGCGCATCTTCCCTTG	2291	-	Genomic primer to validate Δ*MED24*
RB-1-R TCCAAGCGGCGACTGAGATG	2286	-	RB element primer to validate Δ*MED24*
Med24-2-R CGACACCGGTTGCATGTC	2291	-	Genomic primer to validate Δ*MED24*
RB-2-F CCTCGATATACAGACCGATAAAAC	2287	-	RB element primer to validate Δ*MED24*
WH-F GACGCATGATTATCTTTTACGTGAC	2288	-	WH element primer to validate Δ*MED14*

Name Sequence	Stock number	Restrict. site	Purpose
XP-R AATGATTCGCAGTGGAAGGCT	2289	-	XP element primer to validate Δ*MED14*
jumuB-F GCGCAGATCTAACCAACTAATTGCCCCAAA	2039	BglII	Cloning *jumuB* into pDuo2n
jumuB-R ACTGGTACCCTGCCCCCCCAAAAAATTC	2040	KpnI	Cloning *jumuB* into pDuo2n
jumuA-F GCGAGATCTGGTTGTGTTTATTTTGCTAGCCAC	2041	BglII	Cloning *jumuB* into pDuo2n
jumuA-R ACTGGTACCGGAATCAAGTATTTGCCCGC	2042	KpnI	Cloning *jumuA* into pDuo2n
rib-F GCGAGATCTAGACTGGGACACAGCTATGCC	2057	BglII	Cloning *rib* into pDuo2n
rib-R ATAGGTACCCTTGAGTTTTCCGCTCGCC	2058	KpnI	Cloning *rib* into pDuo2n
jumuB-1 CTCTCGGTTTAGCGGGTTTTAATTCGGTATATTCTCCGGGTTATTTCATTCAT	2203	-	Mutagenesis of Ttk2 and Ttk3 in *jumuB*
jumuB-2 GTAACATCCTTAAACACACAGGATGAGTTAGGGTCGTAAATCG	2204	-	Mutagenesis of Ttk4 in *jumuB*
jumuB-3 CACTCAGGTAAAACCAACGATTGCTGTTGTTCGTTCTGGACGAATCTTCCC	2205	-	Mutagenesis of Ttk7, 8, and 9 in *jumuB*
jumuB-4 CTTTCTCGGCATTTTTCGTCACTGGCGATGGTTGTAAG	2206	-	Mutagenesis of Ttk12 in *jumuB*
jumuB-5 CCTTTTCAAGTTGTGCAACGAACCAATTTGTGACCGACGAGGGTATTTTCGTC	2207	-	Mutagenesis of Ttk14 and Ttk15 in *jumuB*
jumuA-1 CGAAAATAACTCATGTTGCAACACTTTAACCAGTACAACAAAAAAGTATCAGC	2198	-	Mutagenesis of Ttk2 in *jumuA*
jumuA-3 GGGTCCGAATACGAATCCGATAGTTCTGTGCTGATGTTTGTGGG	2200	-	Mutagenesis of Ttk6 in *jumuA*
jumuA-5 CAAGTATTTGCCCGCTATGCTTTAATATACTATTTCATTTTATCCGCTACG	2202	-	Mutagenesis of Ttk10 in *jumuA*
rib-1 CGAATGCGAAGGGGAAACATTTAAACTGTGCTACGAATCCACCC	2208	-	Mutagenesis of Ttk1 in *rib*
rib-2 CGTGTGCAGCAAGTGTGTGAAGACCAAATTGTTCGAGCGCTGCCAGAGC	2209	-	Mutagenesis of Ttk2 and Ttk3 in *rib*

3.1.5 Antibodies

Antigen/ Antibody	Conjugated moiety	Host species	Dilution	Vendor/Source
GFP	-	chicken	1:300	Abcam
β–Galactosidase	-	chicken	1:300	Abcam
β–Galactosidase	-	mouse	1:100	Promega
β3 tubulin	-	rabbit	1:300	R. Renkawitz-Pohl
Mef2	-	rabbit	1:200	E. Furlong
Ttk69	-	rabbit	1:200	A. Travers
Ttk69	-	rabbit	1:200	A. Azorin
Bin	-	rabbit	1:50	E. Furlong
Kr	-	guinea pig	1:300	H. Jäckle
Eve	-	guinea pig	1:200	H. Jäckle
Asense	-	rabbit	1:600	Y. Jan
Repo (8D12)	-	mouse	1:5	DSHB
Achaete	-	mouse	1:15	DSHB
PhosphoH3	-	rabbit	1:200	Millipore
Lmd	-	rabbit	1:100	E. Furlong
Futch (22C10)	-	mouse	1:10	DSHB
FasIII (7G10)	-	mouse	1:10	DSHB
Digoxigenin	Alkaline phosphatase	sheep	1:1000	Roche
Digoxigenin	Peroxidase	sheep	1:1000	Roche
Fluorescein	Peroxidase	sheep	1:1000	Roche
Biotin	Peroxidase	sheep	1:1000	Roche
Rabbit IgG	Biotin	donkey	1:200	Jackson Immunoresearch
Mouse IgG	Biotin	donkey	1:200	Jackson Immunoresearch
Guinea pig IgG	Biotin	donkey	1:200	Jackson Immunoresearch
Chicken IgG	Alexa488	goat	1:300	Invitrogen
Mouse IgG	Alexa 488	donkey	1:300	Invitrogen
Rabbit	Cy5	donkey	1:200	Jackson Immunoresearch
Streptavidin	Cy2	-	1:200	Jackson Immunoresearch
Streptavidin	Cy3	-	1:300	Jackson Immunoresearch
Streptavidin	Cy5	-	1:200	Jackson Immunoresearch
Streptavidin	Horse reddish peroxidase	-	1:200	Perkin Elmer

3.1.6 Plasmids and vectors

Name	Purpose	Source
pH-pelican	Generation of transgenic *Drosophila melanogaster* *lacZ*-reporter lines	[195]
pDuo2n	Generation of transgenic *Drosophila melanogaster* *lacZ*-reporter lines	[196]
pAc5.1 A	Constitutive protein expression in *Drosophila* Kc cells	Invitrogen
pGL3-Hsp70	Firefly luciferase reporter vector	[108]
pRL-Hsp70	Renilla luciferase control vector	[108]
pCRII-TOPO	Template for *in vitro* transcription reactions	Invitrogen

3.1.7 Software

Name	Purpose	Source
Vector NTI	Visualization and design of cloning projects, DNA sequence analysis	Invitrogen
ImageJ	Export and processing of microscopy data files	W. Rasband, NIH
Integrated Genome Browser (IGB)	Visualization of ChIP-on-chip data	Affymetrix

3.1.8 Media, solutions, and buffers

LB^{+Amp}–medium:

    10 g Trypton Peptone
    5 g Bacto-yeast extract
    10 g NaCl
    ad 1 l with H$_2$O
    after autoclaving:
    ad 1 ml 1000x Ampicillin

LB^{+Amp}–plates:

    10 g Trypton Peptone
    5 g Bacto-yeast extract
    10 g NaCl
    15 g Agar-Agar
    ad 1 l with H$_2$O
    after cooling off to 50°C:
    add 1 ml 1000x Ampicillin

SOC -medium:	5 g Bacto-yeast extract
	20 g Bacto-Peptone
	20 g Dextrose
	10 mM NaCl
	2,5 mM KCl
	10 mM $MgSO_4$
	ad 1 l with H_2O
Standard fly medium:	1 l H_2O
	12 g agar
	80g corn powder
	18 g dry yeast
	22 g sirup
	10 g soy powder
	6.2 g propionic acid
	80 g malt extract
	2.4 g nipagin
1% Agarose:	1% Agarose (w/v) in 1x TAE buffer
1kb DNA ladder:	#N 3232 L (NEB)
	1000µl of 500 µg/µl 1kb-ladder diluted to 1 µg/10µl in 700 µl 6x Loading Buffer (without xylene cyanol) and 3300 µl 1x TE-buffer
	used at 1 µg/10µl working stock
100bp DNA-ladder:	# N323L (NEB)
	500 µl of 500 µg/µl 100bp DNA-ladder diluted to 1 µg/10 µl in 350 µl 6x Loading Buffer (without xylene cyanol) and 1650 µl 1x TE-buffer
	use: 0.5 µg/10 µl working stock
100x PMSF:	100 mM in 2-propanol
1000x Ampicillin:	100 mg/ml, sterile
1000x Aprotinin:	10 mg/ml in H_2O
1000x Leupeptin:	10 mg/ml in DMSO

20x PBS:	175.2 g NaCl
	44.8 g KCl
	46.6 g Na_2HPO_4 x 12 H_2O
	4.2 g KH_2PO_4
50x TAE:	2 M Tris/glacial acetic acid, pH 7.7
	5 mM EDTA in H_2O
6x Loading Dye:	30% glycerol
	0.25% bromophenol blue
	0.25% xylene cyanol
Cell lysis buffer for Drosophila embryos:	5 mM Hepes, pH 8.0
	85 mM KCl
	0.5% NP40 (IGEPAL)
	+ protease inhibitors
Cross-linking solution:	50 mM Hepes
	1 mM EDTA
	0.5 mM EGTA
	100 mM NaCl
	1.8% Formaldehyde, pH 8.0
DAB Staining solution:	3,3' Diaminobenzidine (DAB) staining solution diluted 1:20 in 3% H_2O_2
DNA extraction buffer:	10 mM Tris-HCl
	25 mM NaCl
	1 mM EDTA
	0.02 mg/ml Proteinase K
Fixing solution:	175 μl 16% formaldehyde (4% final)
	525 μl PBS
Fix/Heptane:	700 μl Fixing solution
	700 μl n-heptane
LiCl – buffer:	250 mM LiCl
	10 mM Tris-HCl, pH 8.0
	1 mM EDTA
	0.5% NP-40
	0.5% sodium deoxycholate
Methanol/PBT:	50% methanol
	50% PBT

Nuclear lysis buffer:

50 mM Hepes, pH 8.0
10 mM EDTA•Na$_2$
0.5% N-Laurylsarkosin
+ protease inhibitors

PBS/Glycin/Triton:

125mM glycin
0.1% Triton in PBS

PBT:

1x PBS
0.1% Triton-X-100

PBT/BSA:

PBT + 0.2% BSA

PBT/BSA/NGS:

PBT/BSA
+ 0.2% Normal Goat Serum
(NGS)

PBTween:

1xPBS
0.1% Tween 20

1000x Pepstatin:

10 mg/ml in DMSO

Protease inhibitors:

1x Aprotinin
1x Leupeptin
1x Pepstatin
1x PMSF

RIPA-buffer (140mM NaCl):

140 mM NaCl
10 mM Tris-HCl pH 8.0
1 mM EDTA
1% Triton-X-100
0.1% SDS
0.1% sodium deoxycholate
1 mM PMSF (from 100 mM
stock in 2-propanol)

RIPA-buffer (500mM NaCl):

500 mM NaCl
10 mM Tris-HCl pH 8.0
1 mM EDTA
1% Triton-X-100
0.1% SDS
0.1% sodium deoxycholate
1 mM PMSF (from 100 mM
stock in 2-propanol)

Streptavidin/HRP:

A + B solutions
(Vector Laboratories)
1:100 in PBT/BSA each;
incubated 1h at r/t prior to use

TE-buffer: 10 mM Tris-HCl, pH 8.0
                                    1 mM EDTA

HybA: 25 ml Formamide
                                    12.5 ml 20X SSC pH 5.0
                                    0.515 ml salmon sperm
                                          (9.7mg/ml)
                                    0.05 ml Tween 20
                                    ad 50 ml DEPC H_2O

HybB: 25 ml Formamide
                                    12.5 ml 20X SSC pH 5.0
                                    12.5 ml DEPC H_2O

3.1.9 Fly lines

Genotype	Source/Generated in the lab of
$\frac{wt}{Y};\frac{wt}{wt};\frac{ttk^{D2-50}}{TM3};\frac{wt}{wt}$	C. Klämbt
$\frac{wt}{Y};\frac{wt}{wt};\frac{Df(3R)awd-KRB}{TM3};\frac{wt}{wt}$	Bloomington stock centre
$\frac{wt}{Y};\frac{wt}{wt};\frac{ttk^1}{ttk^1};\frac{wt}{wt}$	A. Travers
$\frac{rP298-lacZ}{Y};\frac{wt}{wt};\frac{wt}{wt};\frac{wt}{wt}$	M. Ruiz-Gomez
$\frac{rP298-lacZ}{Y};\frac{wt}{wt};\frac{ttk^{D2-50}}{TM3};\frac{wt}{wt}$	this study
$\frac{wt}{Y};\frac{Adv}{CyO};\frac{UAS-Ttk69}{UAS-Ttk69};\frac{wt}{wt}$	A. Travers
$\frac{UAS-Ttk88}{Y};\frac{wt}{wt};\frac{wt}{wt};\frac{wt}{wt}$	A. Travers
$\frac{wt}{Y};\frac{wt}{wt};\frac{UAS-GFP}{UAS-GFP};\frac{wt}{wt}$	Bloomington stock centre
$\frac{rP298-Gal4}{Y};\frac{wt}{wt};\frac{wt}{wt};\frac{wt}{wt}$	S. Abmayr
$\frac{wt}{Y};\frac{sns-Gal4}{sns-Gal4};\frac{wt}{wt};\frac{wt}{wt}$	R. Renkawitz-Pohl
$\frac{wt}{Y};\frac{sns-Gal4}{sns-Gal4};\frac{wt}{wt};\frac{wt}{wt}$	S. Abmayr

Genotype	Source/Generated in the lab of
$\dfrac{wt}{Y};\dfrac{ttk-VME}{ttk-VME};\dfrac{wt}{wt};\dfrac{wt}{wt}$	E. Furlong
$\dfrac{wt}{Y};\dfrac{wt}{wt};\dfrac{Dr}{TM3,twi-Gal4,UAS-GFP};\dfrac{wt}{wt}$	Bloomington stock centre
$\dfrac{wt}{Y};\dfrac{wt}{wt};\dfrac{D}{TM3,AbdA-lacZ};\dfrac{wt}{wt}$	E. Bier
$\dfrac{wt}{Y};\dfrac{Sco}{CyO,wg-lacZ};\dfrac{wt}{wt};\dfrac{wt}{wt}$	Furlong lab stocks
$\dfrac{wt}{Y};\dfrac{Sco}{CyO,ftz-lacZ};\dfrac{wt}{wt};\dfrac{wt}{wt}$	Furlong lab stocks
$\dfrac{wt}{Y};\dfrac{sna^{10}}{CyO};\dfrac{wt}{wt};\dfrac{wt}{wt}$	M. Leptin
$\dfrac{wt}{Y};\dfrac{sna^{18}}{CyO};\dfrac{wt}{wt};\dfrac{wt}{wt}$	Bloomington stock centre
$\dfrac{wt}{Y};\dfrac{wt}{wt};\dfrac{MHC-\tau GFP}{MHC-\tau GFP};\dfrac{wt}{wt}$	E. Olson
$\dfrac{hs-FLP}{Y};\dfrac{wt}{wt};\dfrac{Dr}{TM3};\dfrac{wt}{wt}$	Bloomington stock centre
$\dfrac{wt}{Y};\dfrac{wt}{wt};\dfrac{PBac(RB)e03262}{TM3};\dfrac{wt}{wt}$	Exelixis (Harvard stock centre)
$\dfrac{wt}{wt};\dfrac{wt}{wt};\dfrac{PBac(RB)CG7375^{e02842}}{PBac(RB)CG7375^{e02842}};\dfrac{wt}{wt}$	Exelixis (Harvard stock centre)
$\dfrac{wt}{Y};\dfrac{wt}{wt};\dfrac{PBac(WH)CG17181^{f06749}}{TM3};\dfrac{wt}{wt}$	Exelixis (Harvard stock centre)
$\dfrac{wt}{Y};\dfrac{wt}{wt};\dfrac{PBac(XP)CG32345^{d05325}}{PBac(XP)CG32345^{d05325}};\dfrac{wt}{wt}$	Exelixis (Harvard stock centre)
74 UAS-RNAi lines (listed in Appendix 7.1)	Vienna Drosophila RNAi centre
67 deficiency and loss-of-function lines (listed in Table 4.1-1)	Bloomington stock centre

3.1.10 Web sites

Name	URL
Flybase	www.flybase.org
BDGP *in situ* database	www.fruitfly.org/cgi-bin/ex/insitu.pl
FRT deletion hunter	www.drosdel.org.uk/fdd/del_hunter.php
FlyTF database	www.flytf.org

3.2 METHODS

3.2.1 Molecular biology and biochemistry

3.2.1.1 Generation of transgenic reporter lines

To investigate the regulatory potential of genomic DNA regions, transgenic *Drosophila melanogaster* reporter lines were generated.

Regulatory regions identified in the Ttk69 ChIP-on-chip experiment were subcloned into the vector pH-pelican [195] and germ-line transformed into *Drosophila melanogast*er *white⁻* flies. Following are the coordinates of the subcloned regions (dm3/BDGP release 5.0, April 2006): chr3R: 6182898-6183568 (*jumuA*), chr3R: 6178511-6179367 (*jumuB*), chr2R: 15162910-15163336 (*rib*).

The elements assayed for Snail regulatory input were subcloned into the vector pDuo2n [196]. Germ-line transformation was then performed into *Drosophila melanogast*er flies with *attP* landing site mapped to the cytological band 51D on the second chromosome [197]. The coordinates of the subcloned regions are (dm3/BDGP release 5.0, April 2006): chr3R: 9118948-9119470 (*WntD-lacZ*), chr3R: 17205652-17206054 (*Tin B-374*), chr2R: 5819019-5819498 (*Mef2 I-D[L]*), chr2R: 8833920-8834301 (*Mdr49 early mesoderm*).

For all constructs, at least two independent transgenic lines were obtained and assayed.

3.2.1.2 Site-directed mutagenesis

Mutagenesis of sequence motifs was performed using the QuikChange Multi site-directed mutagenesis or QuikChange II site-directed mutagenesis kits, following the supplier's guidelines (Stratagene). In all reactions, the enhancers were mutated in pGL3 vectors and then were transferred to pDuo2n or pH-pelican vectors. All constructs were sequence-verified.

3.2.1.3 Luciferase reporter assays

Drosophila Kc cells were transiently transfected using Cellfectene (Invitrogen). Twist, Snail, and Dorsal were expressed from full-length ESTs (AT15089, RE58537,

and RE35237, respectively) in pAc5.1 vector. Successful expression of the proteins was verified by a Western blot. The enhancers (coordinates given above) were assayed in a pGL3 luciferase reporter vector with a *Heat-shock-protein-70* (*Hsp70*) minimal promoter and the luciferase activity was normalized to Renilla standard. The total amount of transfected DNA was kept constant by supplementing empty pAc5.1 vector. The measurements were performed according to the manufacturer's recommendations (Dual-Luciferase Reporter Assay, Promega) with a PerkinElmer 1420 Luminescence Counter.

3.2.1.4 Chromatin immunoprecipitations from *Drosophila* embryos and analysis on genomic tiling microarrays

ChIP-on-chip experiments were performed according to the previously described procedures [198].

Briefly, tightly staged *Drosophila* wild type embryos were formaldehyde crosslinked and extracted chromatin was sonicated to approximately 500 bp fragments. A chromatin aliquot was pre-incubated with the antibody or an equal volume of pre-immune serum and the complexes were then incubated with Protein A sepharose beads. The immunocomplexes bound to the beads were then thoroughly washed, RNA and proteins were digested and the crosslinking was reversed by heat. Phenol-chloroform extracted and precipitated DNA fragments were assayed by quantitative PCR for enrichment of positive over negative control regions.

Next, a linker consisting of a 20mer and 5' phosphorylated 24mer was annealed to the DNA fragments. Using the 20mer as a primer, the input material was amplified in two subsequent rounds of PCR. After purification, quality of the sample was again verified by quantitative PCR. Following DNAse I fragmentation the DNA was labelled using terminal deoxynucleotide transferase in the presence of biotinylated 2',3'-Dideoxyadenosine-5'-Triphosphate. Samples were then hybridized to Affymetrix GeneChip® *Drosophila* Tiling array 1.0R at 43°C for 16 hours. Washing and staining of arrays was performed according to the supplier's guidelines.

Obtained raw data were normalized and evaluated by Bartek Wilczynski as described previously [196].

3.2.2 Histological techniques

3.2.2.1 *In situ* hybridization and immunohistochemistry

In situ hybridizations were performed using standard protocols [199]. The following ESTs were used to generate digoxigenin, fluorescein, or biotin labelled probes: LD28689 (*ttk69*), RE35237 (*sna*), LD41072 (*row*), LD23630 (*CG8478*), GH01967 (*gem*), LD11946 (*CG9650*), HL07808 (*luna*), GH06338 (*CG30431*), SD10012 (*Mdr49*), LD47926 (*lmd*), AT15089 (*twi*), and SD06902 (*zfh1*).

The full-length cDNA clones of *Mef, tin, sns,* and *lacZ* were a gift from M. Taylor, M. Frasch, S. Abmayr, and U. Elling, respectively.

The probe against the *Green Fluorescent Protein* (*GFP*) transcript was made by amplifying the *GFP* sequence from the pH-Stinger vector.

The probes for *CG33936*, *dmrt11E*, *HmgZ,* and *htl* were generated from their largest exons which were cloned into the pCRII-TOPO vector.

3.2.3 Generation of *Drosophila* deletion lines by FRT-mediated recombination

The custom-made deletion line Δ*MED24* was generated according to the following crossing scheme [200]:

Cross	Progeny
$\dfrac{hs-FLP}{hs-FLP}; \dfrac{Dr}{TM3}$	$\dfrac{hs-FLP}{Y}; \dfrac{PBac(RB)e03262}{TM3}$
X	
$\dfrac{wt}{Y}; \dfrac{PBac(RB)e03262}{TM3}$	
$\dfrac{hs-FLP}{Y}; \dfrac{PBac(RB)e03262}{TM3}$	$\dfrac{hs-FLP}{wt}; \dfrac{PBac(RB)e03262}{PBac(RB)CG7375^{e02842}}$
X	During larval stages, subjected to a heat shock at 37ºC for 2 hours.
$\dfrac{wt}{wt}; \dfrac{PBac(RB)CG7375^{e02842}}{PBac(RB)CG7375^{e02842}}$	

$$\frac{hs - FLP}{wt} :: \frac{PBac(RB)e03262}{PBac(RB)CG7375^{e\,02842}}$$

X

$$\frac{wt}{Y} :: \frac{Dr}{TM3}$$

All progeny screened by complementation test against a deficiency line and by two-sided PCR.

Final stock

$$\frac{wt_{or}hs - FLP}{Y} :: \frac{PBac(RB)e03262 - PBac(RB)CG7375^{e\,02842}}{TM3}$$

The deletion line Δ*MED14* was generated according to the following crossing scheme [200]:

Cross	Progeny
$\frac{hs - FLP}{hs - FLP} :: \frac{Dr}{TM3}$ X $\frac{wt}{Y} :: \frac{PBac(WH)CG17181^{f\,06749}}{TM3}$	$\frac{hs - FLP}{Y} :: \frac{PBac(WH)CG17181^{f\,06749}}{TM3}$
$\frac{hs - FLP}{Y} :: \frac{PBac(WH)CG17181^{f\,06749}}{TM3}$ X $\frac{wt}{wt} :: \frac{PBac(XP)CG32345^{d\,05325}}{PBac(XP)CG32345^{d\,05325}}$	$\frac{hs - FLP}{wt} :: \frac{PBac(WH)CG17181^{f\,06749}}{PBac(XP)CG32345^{d\,05325}}$ Subjected to a heat shock at 37ºC for 2 hours during larval stages.
$\frac{hs - FLP}{wt} :: \frac{PBac(WH)CG17181^{f\,06749}}{PBac(XP)CG32345^{d\,05325}}$ X $\frac{wt}{Y} :: \frac{Dr}{TM3}$	White-eyed progeny was screened by complementation test against a deficiency line and by hybrid PCR.

Final stock

$$\frac{wt_{or}hs - FLP}{Y} :: \frac{PBac(WH)CG17181^{f\,06749} - PBac(XP)CG32345^{d\,05325}}{TM3}$$

4 RESULTS

4.1 A molecular screen for novel transcription factors essential for *Drosophila* mesoderm development

4.1.1 Experimental design of the screen

The main aim of the screen was to identify novel myogenic regulators that would fulfil two criteria:

o Expression in mesoderm: in order to play a regulatory role in muscle development the TF must be present in mesodermal cells.

o Essential role in mesoderm development: if the TF is involved in muscle development, in its absence the development of one or more muscle types should be perturbed.

The screen was therefore performed in two rounds; the first round focused on identifying novel mesodermally expressed TFs, and in the second round their requirement for myogenesis was systematically assessed.
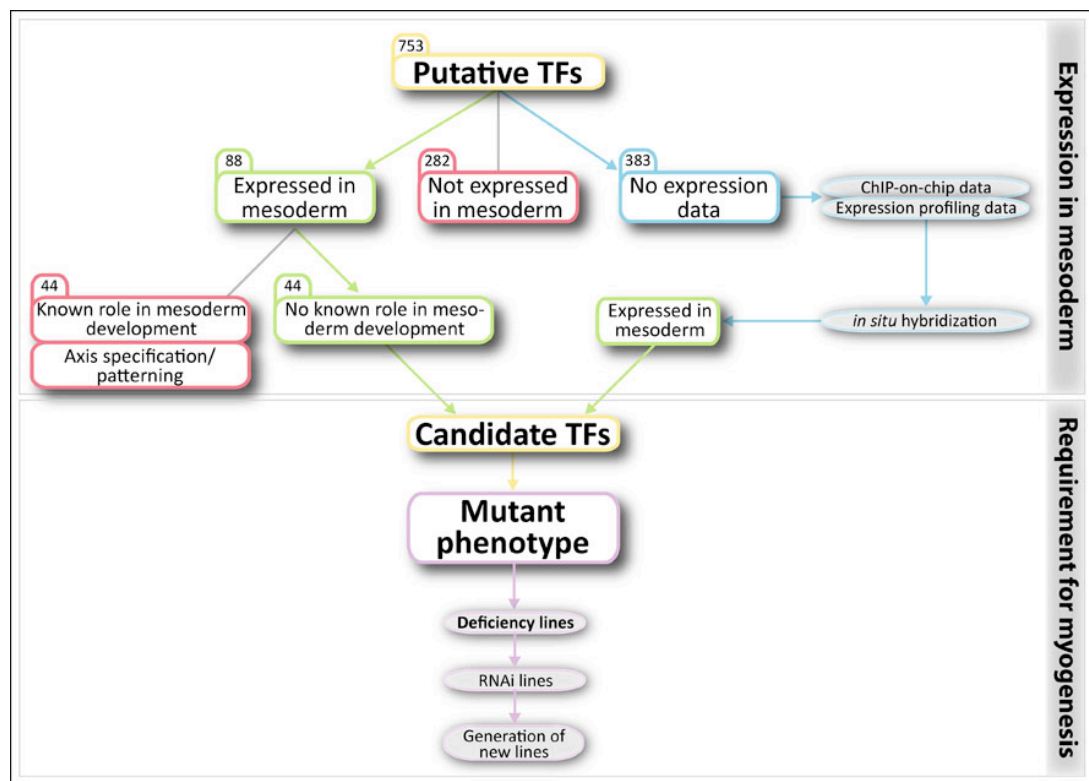

Figure 4.1-1: Experimental design of the molecular screen

The number of unique TFs in each class (in boxes) is based on Release 1 of FlyTF [201] and Release 2 of BDGP expression patterns database [202].

4.1.1.1 Selection of candidate mesodermal TFs

The screen was initiated by examining all putative *Drosophila* TFs which were manually curated based on their ability to bind DNA and/or regulate gene expression [201]. Altogether, there are 753 proteins which are annotated as being sequence-specific transcription factors. To refine the number of TFs for further analysis, the Berkeley Drosophila Genome Project (BDGP) *in situ* hybridization database was used as a source of annotated embryonic gene expression patterns [202]. Its Release 2 contained expression patterns of 370 TF-encoding genes, out of which 88 were annotated as mesodermally expressed during at least one time point up to stage 11. One half of these TFs have an already identified role in mesoderm development or are involved in the specification of embryonic axes. I therefore focused on the other half of the mesodermally expressed TFs whose functional roles in the mesoderm have not yet been investigated. This group of TFs therefore represents candidate TFs whose requirement for myogenesis was further assayed (Figure 4.1-1).

In addition to the TFs with known expression patterns, there were 383 TFs without annotated expression in the Release 2 of BDGP database. As many of these are likely to be present in mesoderm, available genome-wide data was used to prioritize them for further examination. More specifically, putative regulation of these genes by mesodermal TFs (using ChIP-on-chip binding data) and expression profiling of mesodermal mutants were used as indicators of their likelihood to be involved in myogenesis. Subsequently, colorimetric *in situ* hybridization of some candidates was performed and once their expression in mesoderm was confirmed, they were added to the group of candidate TFs (Figure 4.1-1).

4.1.1.2 Identification of novel mesodermally expressed TFs

Of the 383 TFs for which no information on their embryonic expression was available (Figure 4.1-1), *in situ* hybridization was performed based on their likelihood to be expressed in the mesoderm. This likelihood was reflected by a score that was derived from two types of previously generated genome-wide data. First, it was assumed that mesodermally expressed genes are activated by one or more key mesodermal transcriptional regulators. CRMs identified in 8 different ChIP-on-chip experiments with Twi, Tin, and Mef2 [196] were therefore assigned to target genes.

Each such target gene encoding a TF was allocated zero to eight points, depending on the number of experiments in which a CRM in its vicinity was detected. Altogether, 132 of the 753 TFs have at least one binding event in their proximity while about half of these TFs (53%) are regulated by at least two different mesodermal TFs. Second, if the potential mesodermal TFs are downstream of the known regulators, their expression levels should be reduced in their absence. Gene expression changes in mutants for *twi*, *tin*, and *Mef2* were previously quantified in genome-wide expression profiling experiments covering three to four different time windows [55, 62], contributing thereby 10 additional datasets. Again, one point was assigned if the expression of a TF-coding gene was significantly reduced ($\log_2 < -0.77$, $q < 0.05$) in a given dataset. The highest possible total score was therefore 18 as 8 ChIP-on-chip and 10 expression profiling datasets were included in the analysis (Figure 4.1-2 A).

As expected, the TF-encoding genes that are annotated by the BDGP as mesodermal have on average higher scores than those expressed in other tissues and conversely, genes not expressed in mesoderm obtain very low scores (Figure 4.1-2 B). The scores of genes without known expression pattern are distributed in between those of mesodermal and non-mesodermal genes. 39 of the 383 (10%) TFs that were not included in the BDGP database have a score of 4 or higher, compared to 3 of 282 (0.01%) non-mesodermal TFs, and 61 of 88 (70%) mesodermal TFs (Figure 4.1-2 B inset). As a consequence, this class of TFs without expression information is likely to contain mesodermal TFs and the assigned score should facilitate their systematic identification. I therefore selected genes from this class for *in situ* hybridization analysis.

By *in situ* hybridization I confirmed expression of three genes (*row*, *CG8478*, *gem*, Figure 4.1-2 C-E) that were annotated as mesodermal in BDGP database, yet their assigned score was low and the mesodermal aspect of their expression was not apparent from the images supplied by BDGP. In addition, six TFs with unknown expression pattern were identified as mesodermal. *Doublesex-Mab related 11E* (*dmrt11E*) is the third highest scoring gene in the class of TFs with unknown expression pattern and is expressed in the trunk mesoderm of stage 9 and 10 embryos (Figure 4.1-2 F). The only suggested role of *dmrt11E* is in sex differentiation [203]. *HMG protein Z* (*HmgZ*) encodes a member of the high mobility group of proteins that are associated with chromatin remodelling [204]. So far no direct link between this type of DNA-binding proteins and *Drosophila* muscle development has been

established, yet *HmgZ* mRNA could be detected in the trunk mesoderm of stage 10 embryos (Figure 4.1-2 H). *luna* encodes a zinc finger TF and its loss- and gain-of-function during embryogenesis interferes with viability [205]. The specific processes that are affected have however not been identified. *luna* mRNA can be detected in the primordia of somatic and visceral muscle of stage 11 embryos (Figure 4.1-2 I), suggesting that it might play a role in the development of these mesodermal derivatives. Finally, three TFs without any information on their developmental function have been identified as mesodermal. *CG9650* scores 8 points and shows panmesodermal expression at late stage 10 (Figure 4.1-2 G). *CG33936* can be detected in the somatic muscle primordium at stage 11 (Figure 4.1-2 J) and *CG30431* appears to be expressed in the mesoderm-derived fat body (Figure 4.1-2 K).

In summary, ChIP-on-chip and expression profiling experiments are predictive of mesodermal expression and thereby facilitated the identification of novel mesodermally expressed TFs. A number of low-scoring genes were also expressed in mesoderm (Figure 4.1-2 C-E), underscoring the fact that Twi, Tin, and Mef2 are not the only mesodermal regulators. As the Release 3 of BDGP expression patterns database was anticipated to cover expression of most of *Drosophila* TFs, instead of analyzing embryonic expression of additional genes, I rather proceeded to the second round of the screen, focusing on the phenotypic analysis of the candidate mesodermal TFs.
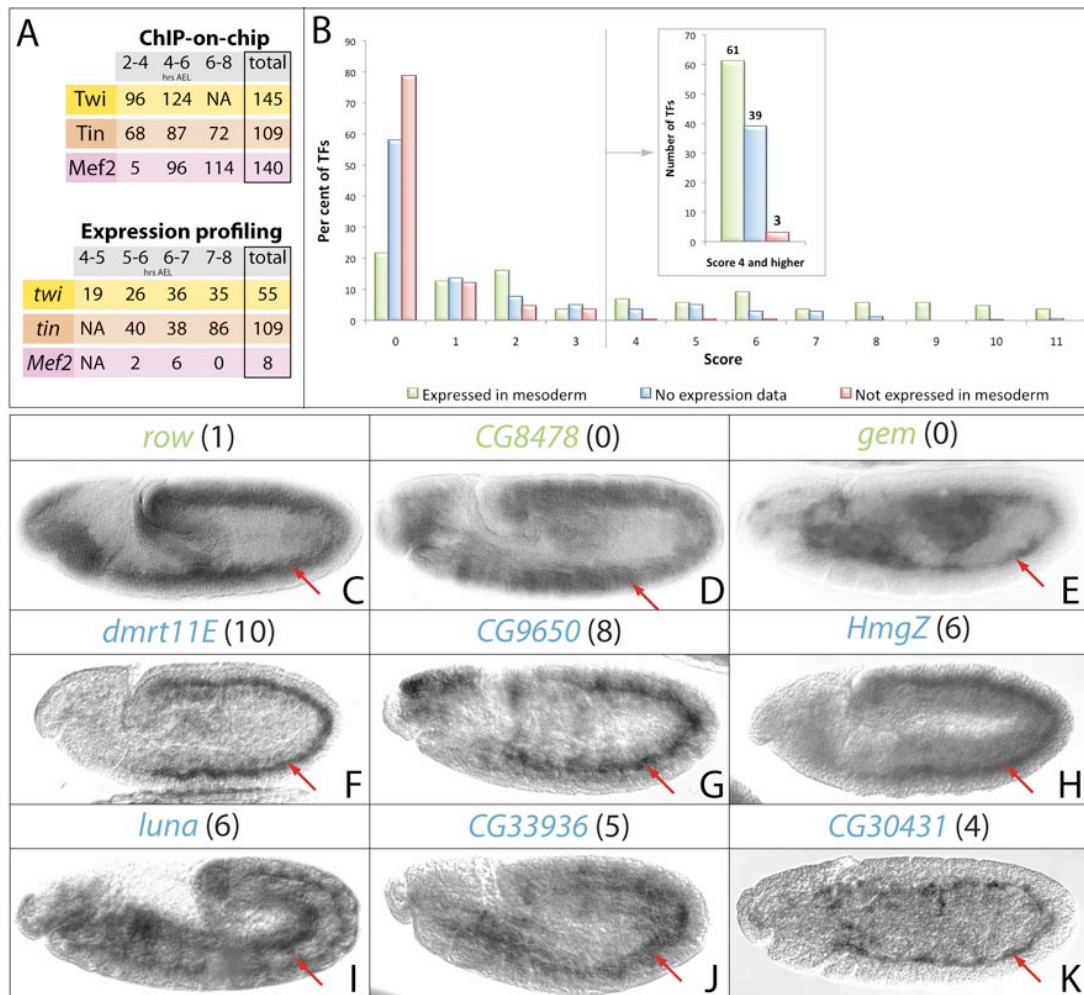
Figure 4.1-2: Identifying transcription factors expressed in mesoderm

(A)   Score assignment was used to reflect the likelihood of mesodermal expression. Eight different ChIP-on-chip and 10 different expression profiling datasets were included for calculation of the score. The numbers indicate all TFs (out of 753) identified in a given data set and the total number of TFs in the last column indicates unique TFs per a set of experiments (e.g. number of unique TFs regulated by Twi, Tin, or Tin across all considered time points).

(B)   Distribution of the score among the three expression classes; TFs expressed in mesoderm (green), TFs with unknown expression (blue), and TFs not expressed in mesoderm (red). The inset shows the number of TFs with a score equal to or higher than 4, the selected cut-off.

(C-K)   Colorimetric *in situ* hybridization to confirm mesodermal expression of low-scoring TFs annotated as mesodermal (C-E) and to identify novel mesodermally expressed TFs (F-K). In the brackets is a score assigned to each gene and the red arrows point to mesodermal expression.

4.1.1.3 Phenotypic analysis of new mesodermal TFs

To assess the importance of candidate mesodermal TFs for muscle development, their mutant phenotypes were analysed. First, available deficiency lines removing bigger parts of chromosomes, and thereby multiple genes, were used

individually or in combination. Deficiencies were selected so that the smallest possible deletion around the candidate TF was achieved. In most cases, two deficiency lines were combined so that the transheterozygous embryos were homozygous mutant only for their overlapping region. To enable unambiguous identification of the homozygous mutant embryos, all lines carried a *wg-lacZ* or *abd-A-lacZ* marked balancer chromosome, depending on whether the deleted region was located on the second or third chromosome, respectively.

Ideally, the candidate TF was situated close to the centre of the deleted region, especially in the case of deficiency lines whose breakpoints were derived only from polytene chromosome squashes and not by recently developed Polymerase Chain Reaction (PCR)-based protocols [200, 206]. Finally, deficiency lines deleting already known myogenic factors and/or patterning genes were excluded as their phenotype would obscure the mutant phenotype of the candidate TF.

When muscle-specific defects were observed in the deficiency lines, expression of the removed, mesodermally expressed genes was knocked down specifically in the mesoderm using RNAi lines (Appendix 7.1). Altogether, 74 different RNAi lines directed against 48 genes were used. Out of these, 23 lines, corresponding to 18 genes, showed lethality at early larval to pupal stages when combined with a panmesodermal driver. However, as the knockdown is effective only relatively late in embryogenesis, the RNAi lines do not provide definitive information regarding the requirement of the gene for mesoderm specification (see Section 5.1.2.1). Therefore, smaller deletion lines removing the candidate TFs with the most interesting phenotypes were generated.

4.1.2 Classes of identified muscle phenotypes

Altogether, 67 different deficiency and loss-of-function lines were used individually or in combination to delete 46 distinct genetic loci on the X, second, and third chromosomes. In 25 cases, the phenotype was either wild type or obscured by pleiotropic, general embryonic defects, whereas in the 21 other cases muscle-specific defects were identified. 20 of these occurred within somatic musculature and they were further subdivided into three phenotypic classes (Figure 4.1-3).
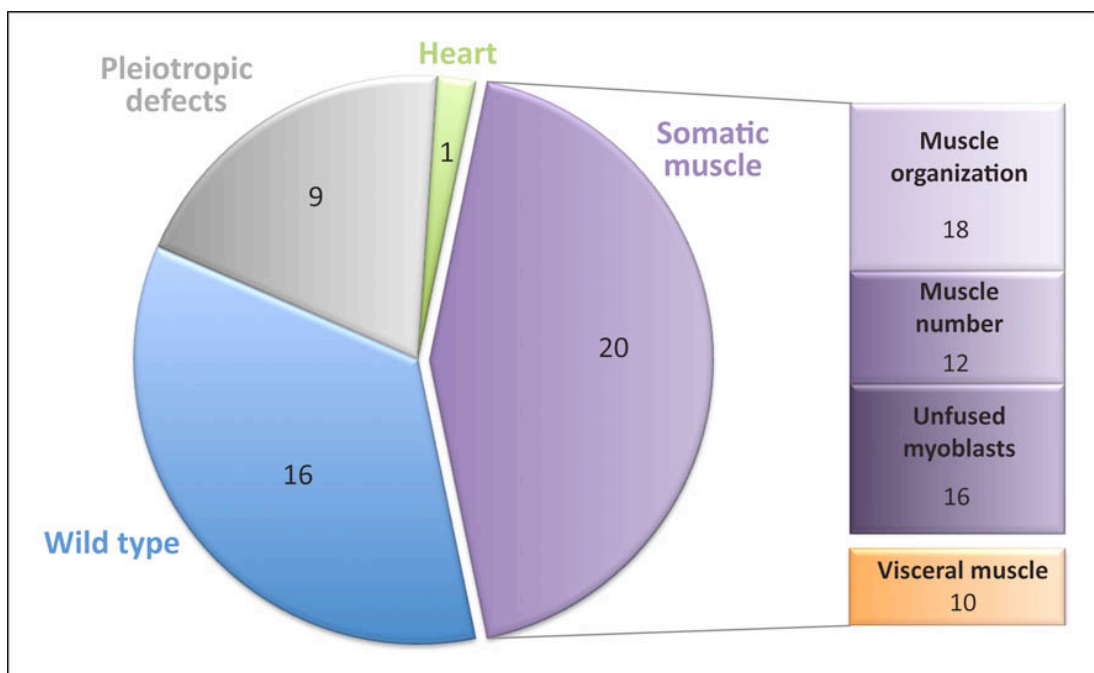
Figure 4.1-3: Distribution of identified phenotypes
Phenotypes of 46 genetic loci deleted by deficiency and loss-of-function lines were subdivided into 4 major groups based on the overall phenotype. The phenotypes in somatic musculature were further classified into three different subgroups reflecting the particular processes that were affected. The number of unique loci in each group is indicated (the somatic muscle and visceral muscle phenotypes are non-exclusive).

The genetic loci exhibiting abnormal somatic muscle development were divided into three loosely defined groups (classes III-V) based on the complexity of the observed somatic muscle phenotypes, while the wild type phenotype and pleiotropic defects represent classes I and II, respectively (Table 4.1-1).

Genotype	Wild type (16)	Pleiotropic defects (9)	Morphology and organization (18)	Muscle number (12)	Mononucleated myoblasts (16)	Heart (1)	Visceral muscle (10)	Class
			Somatic muscle					
Df(3R)Exel7309	■							I (16)
Df(2L)PM4/Df(2L)ED929	■							
Df(3L)Exel6084/Df(3L)Exel6083	■							
Df(2L)Exel7038	■							
Df(2L)Exel6048/Df(2L)TW1	■							
chn^{ECJI}/Df(2R)XTE-18	■							
Df(2R)Exel6059	■							
EcR^{V559fs}/Df(2R)20B	■							
Df(2R)ED3923	■							
Df(2R)Exel7153	■							
Df(2R)Exel6059/Df(2R)ED2155	■							
Df(1)r-D17	■							
Df(1)Exel6239	■							
noc^{TE35B}/Df(2L)A266	■							
Df(1)ED7355	■							
Df(2R)ED3610/Df(2R)ED3683	■							
Df(2R)Exel6069/Df(2R)Bsc26		■						II (9)
Df(2R)Pcl11B/Df(2R)14H10W-35		■						
Df(2L)pr-A20/Df(2L)Exel6044		■						
Df(2L)ED695/Df(2L)Exel6024		■						
Df(1)ED7147		■						
Df(1)Sxl-ra		■						
Df(1)Exel6253		■						
Df(2R)XTE-58		■						
Df(3R)T-61		■						
Df(3L)XG10			■					III (12)
Df(3L)vin5/Df(3L)BSC14			■				■	
Df(2L)A266/Df(2L)A260			■					
Df(2R)ED1552/Df(2R)42			■				■	
Df(2R)stan2/Df(2R)X1			■				■	
Df(3L)B71/ Df(3L)ED201			■					
Df(1)RR79			■				■	
Df(1)ED7067			■				■	
Df(1)m259-4			■				■	
Df(3L)7C/Df(3L)ED4177			■				■	
Df(3L)ZP1/Df(3L)Exel6112			■				■	
Df(1)Exel6245			■	■			■	
Df(2R)Exel7135			■					IV (6)
Df(3R)by416			■					
Df(3L)ED4978			■	■				
Df(3L)BK10			■					
Df(2L)Exel7042/Df(2L)ED680			■					
Df(3R)Exel6202			■				■	
Df(3L)Exel6134					■			V (2)
Df(3L)ED4475/Df(3L)iro-2					■			
Df(2R)Dll-MP						■		

Table 4.1-1: Analyzed genetic loci and the classification of their phenotypes

Summary of all analyzed deficiency and loss-of-function lines and their phenotypes. Each row represents a single genotype and each column a phenotype. The genetic loci were grouped into 5 classes (I-V) based on their overall and somatic muscle phenotypes.

4.1.2.1 Class I: Embryos with no observable phenotype

To assess the muscle phenotype of all analysed lines, an antibody directed against the cytoskeletal component β3 tubulin was used [207]. This antibody labels all mesodermally-derived cell types and tissues from stage 11 until the end of embryogenesis. After stage 16, when the somatic musculature has fully differentiated, the staining is prominent in dorsal, dorso-lateral (Figure 4.1-4 A-B'), lateral, and ventral muscle fibres (Figure 4.1-4 B-D'). β3 tubulin is present in four of the six cardioblasts per hemisegment (Figure 4.1-4 A, A'), where its expression is controlled by the cardiac regulator Tinman [208]. The antibody also labels the trunk visceral mesoderm which constricts and loops after stage 15 (Figure 4.1-4 E-H).
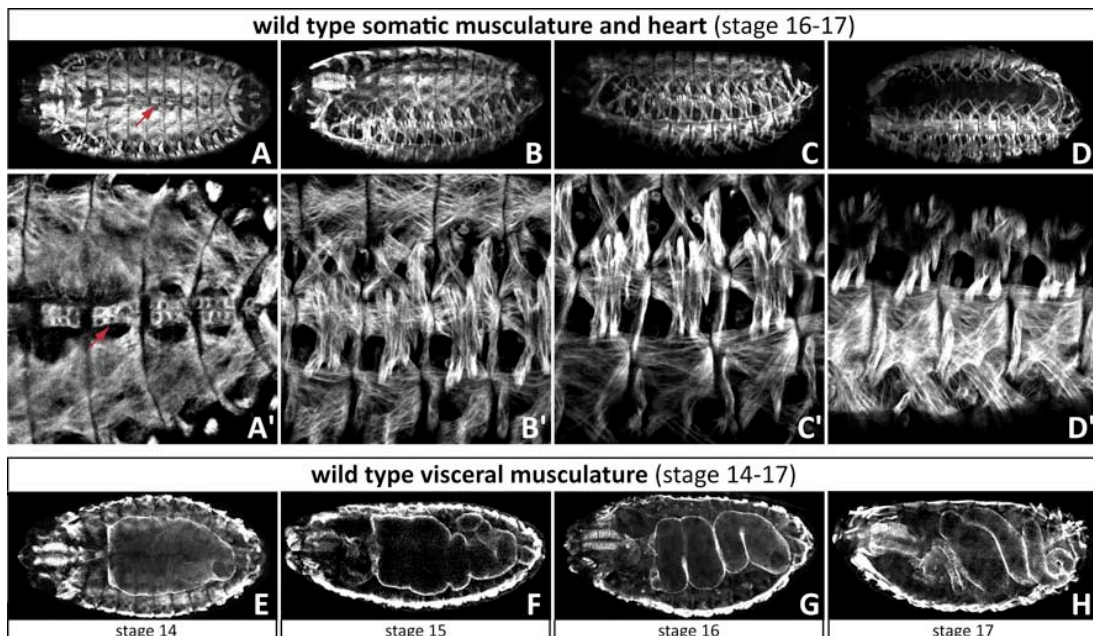
Figure 4.1-4: Somatic, cardiac, and visceral musculature in wild type embryos

All three muscle types in *Drosophila* wild type embryos are visible using an anti-β3 tubulin antibody.

(A-A') Dorsal view of a stage 16 embryo showing dorsal somatic muscle and the heart (red arrow).

(B-B') Dorsolateral view of a stage 17 embryos. All four groups of somatic muscles (dorsal, dorsolateral, lateral, and ventral) are visible.

(C-C') Lateral view of a stage 16 embryos with lateral and ventral muscles displayed.

(D-D') Ventral view of a stage 17 embryos showing ventral musculature.

(E-H) Development of visceral musculature in wild type embryos from stage 14 (E) to the end of embryogenesis (H).

Deficiency lines deleting 16 genetic loci exhibited no obvious defects in formation of the three muscle types (Table 4.1-1, data not shown). The absence of obvious phenotypes upon deletion of mesodermal TFs could have technical (mis-mapping of deficiency breakpoints) as well as biological reasons (discussed in the Section 5.1.1).

4.1.2.2 Class II: Pleiotropic defects

In the second phenotypic class, encompassing nine genetic loci, specific muscle defects could not be assessed as the embryos suffered from pleiotropic defects (Figure 4.1-5). These include a failure in morphological events, such as gastrulation or germ-band extension, that would lead to developmental arrest prior to muscle specification and differentiation. As the subdivision of the presumptive mesoderm into the three muscle primordia requires an intact overlying ectoderm as a signal source (see Section 1.3.1), any patterning or axis specification defects prevent analysis of muscle-specific phenotypes as well. The pleiotropic defects could also be caused by deleted gene(s) other than the candidate TF and the inability to assess the phenotypes due to pleiotropic defects is one of the greatest limitations of deficiency lines.
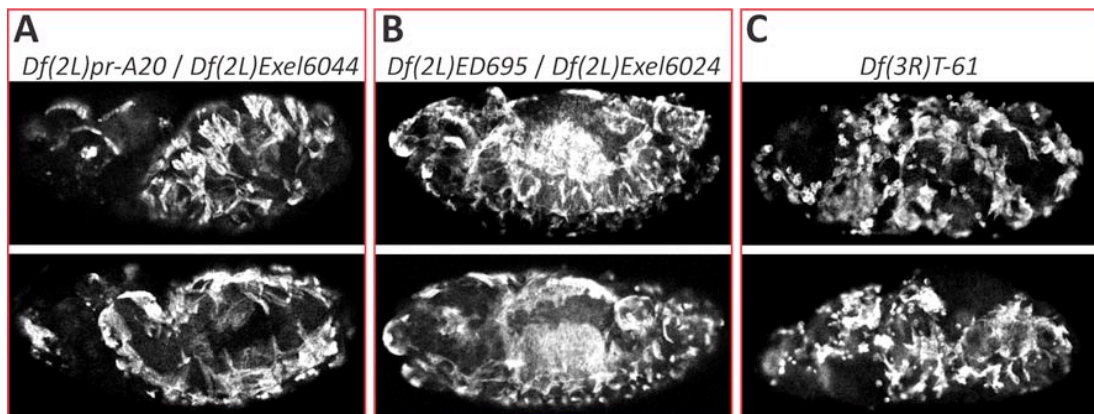
Figure 4.1-5: Phenotypic class II: Pleiotropic defects
Example of three (out of nine, see Table 4.1-1, class II) genetic loci which upon deletion give rise to pleiotropic defects. Embryos were subjected to an immunostain with the anti-β3 tubulin antibody.

4.1.2.3 Class III: Somatic muscle specification and myoblast fusion defects

The largest class of identified somatic muscle phenotypes comprises twelve genetic loci. Their deletion gives rise to complex defects in somatic muscle number, morphology, and overall organization, indicative of abnormal muscle specification and/or guidance and attachment to the ectodermal tendon cells. In addition, the severe phenotypes were often accompanied by unfused myoblasts, a sign of myoblast fusion failure.

The deficiency line *Df(3L)XG10* deletes an approximately 300 kB genomic region and the homozygous mutant embryos exhibit a defect mainly in muscle morphology and attachment (Figure 4.1-6 A). The most affected seem to be the lateral transverse muscles which do not extend but are rounded. In some cases, they remain mononucleated, indicating impaired fusion. The ventral muscles display dramatic muscle guidance defects and appear also randomly positioned.

The exact physical limits of the deficiencies *Df(3L)vin5* and *Df(3L)BSC14* have not been mapped, but based on the cytological information from polytene chromosome squashes, we can estimate that about a 70 kB region covering 20 predicted genes should be absent in transheterozygous embryos. The embryonic somatic musculature of these embryos is affected mainly in the dorsal region where muscle fibres are either absent or have abnormal morphology (Figure 4.1-6 B).

Transheterozygous embryos of *Df(2L)A266* and *Df(2L)A260* remove 34 genes spreading over about 500 kB. The resulting embryos have defects in the development of distinct muscle fibres, such as the dorsal oblique 3 or ventral oblique 4 and 5 muscles (Figure 4.1-6 C). The candidate TF removed in this region, *no ocelli* (*noc*), is however not responsible for this phenotype as its amorphic allele *noc*^{TE35B} [209] is viable when combined with the line *Df(2L)A266*.

The left limit of the region shared by lines *Df(2R)ED1552* and *Df(2R)42* was approximated by polytene analysis, while the right limit was confirmed by a PCR-based method. The affected locus has a size of approximately 90 kB and it removes 21 annotated genes. In the dorsal somatic musculature of transheterozygous embryos, there are some muscle fibres missing and those that are present have aberrant arrangements (Figure 4.1-6 D).

Deletion of 43 genes within a 450 kB region by the deficiency lines *Df(2R)stan2* and *Df(2R)X1* results in a mild phenotype affecting the dorsal somatic musculature with a few absent muscle fibres (Figure 4.1-6 E).

The overlap of deficiency lines *Df(3L)B71* and *Df(3L)ED201* removes about 200 kB covering 42 genes. The main feature of the transheterozygous muscle phenotype is the occurrence of a high number of unfused myoblasts, whereas some myofibers are absent or have abnormal shape and/or size, suggesting a defect in myoblast fusion (Figure 4.1-6 F).

The X-chromosome located deficiency *Df(1)RR79* deletes 48 genes spanning 490 kB. Embryos deficient for this region lack some muscle fibres, for example the dorsal oblique muscles, and their ventral fibres cross the ventral midline, suggesting a defect in muscle guidance (Figure 4.1-6 G).

Deletion of 29 genes within the 211 kB region removed by the deficiency line *Df(1)ED7067* gives rise to a very interesting phenotype in muscle identity determination. The dorsolateral muscle fibres dorsal oblique 3 and 4, and dorsal acute 3 are missing while some lateral transverse muscles are duplicated (Figure 4.1-6 H). As there are more muscle fibres of one type and fewer of another, it is likely that an identity transformation has occurred. Alternatively, a defect in asymmetric cell division could also give rise to more instances of a particular muscle fibre at the expense of others.

The deficiency *Df(1)m259-4* deletes approximately 280 kB containing 36 genes. The somatic musculature exhibits a few missing muscle fibres, while mononucleated, unfused myoblasts are present in their positions. In addition, many myotubes fail to extend towards their attachment sites and instead form spheroid-like structures instead (Figure 4.1-6 I).

Finally, the breakpoints of the deficiency *Df(1)Exel6245* were precisely mapped; it deletes 15 genes within a 92 kB region. In some of the mutant embryos, lateral transverse muscles are duplicated and their overall morphology is perturbed (Figure 4.1-6 J).

Altogether, the ten described deficiencies give rise to a range of phenotypes in several groups of muscle fibres. All show specification defects where either certain muscle fibres are missing or duplicated, accompanied by abnormal muscle morphology and arrangement.
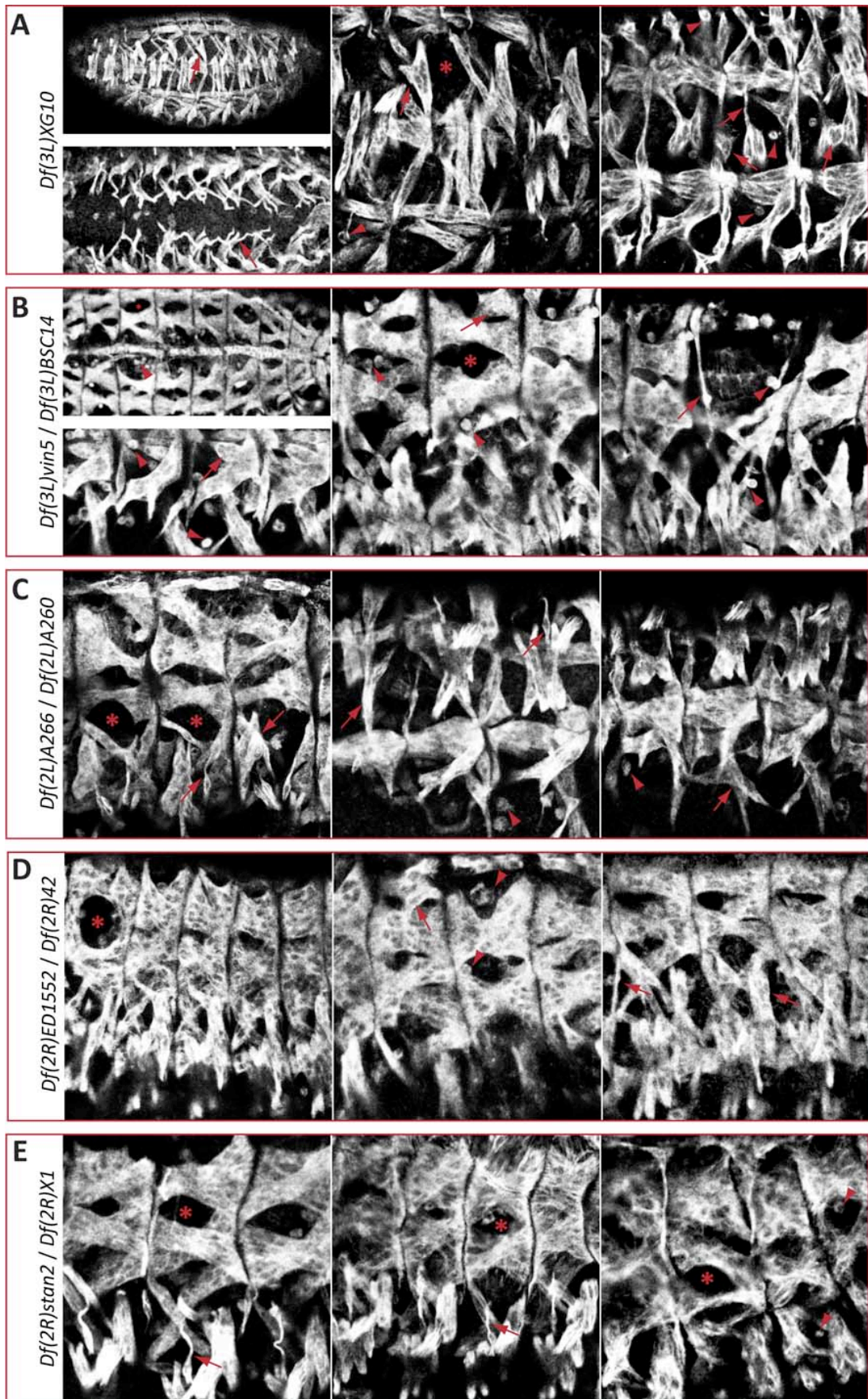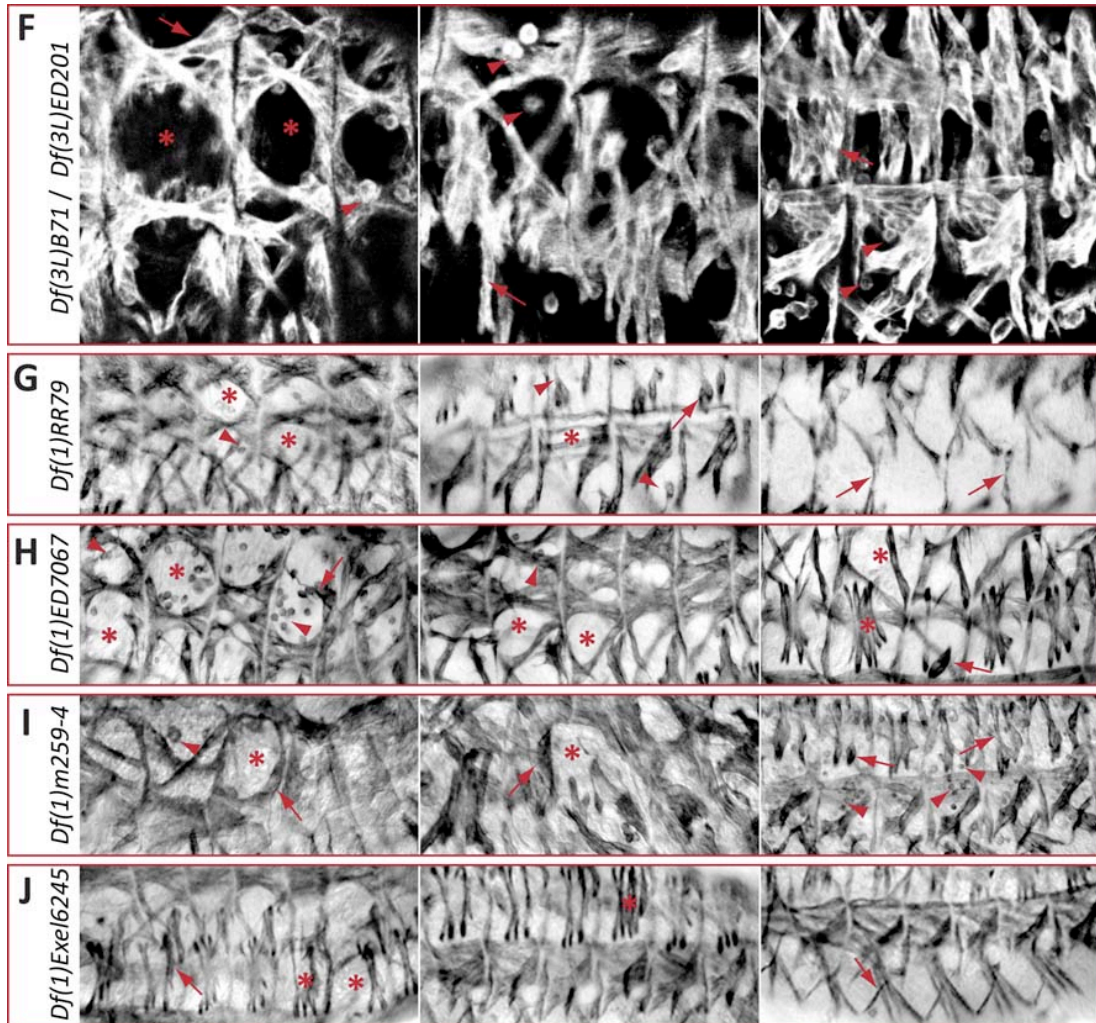
Figure 4.1-6: Phenotypic class III: Aberrant somatic muscle number and/or organization

Embryos of all genotypes were subjected to either fluorescent (A-F) or colorimetric (G-J) immunostaining with anti-β3 tubulin antibody. Red arrows indicate abnormal muscle morphology, red arrowheads point to unfused myoblasts, and red asterisks indicate absent or duplicated muscle fibres. At least 3 different stage 16 or 17 embryos are shown per each genotype.

4.1.2.4 Class IV: Somatic muscle morphology and organization

The fourth class of identified phenotypes contains six unique genetic loci whose deletion results in abnormal muscle arrangement or morphology while the number of muscle fibres remains unaffected.

The line *Df(2R)Exel7135* deletes a 133 kB locus which removes 20 annotated protein coding genes. Some of the somatic muscle fibres have an abnormal shape (Figure 4.1-7 A), suggesting that they fail to respond to attractant cues secreted by the ectoderm (or, alternatively, these cues are absent).

The breakpoints of the sequence deleted by *Df(3R)by416* is not precisely known as it was derived from polytene chromosome squashes. The largest possible region that it might delete encompasses 380 kB and 67 genes. Similar to *Df(2R)Exel7135*, lateral transverse muscles do not extend towards their attachment sites, but in addition, the ventral muscle fibres migrate over the ventral midline (Figure 4.1-7 B).

In the case of the line *Df(3L)ED4978*, 380 kB of genomic sequence and 57 genes are removed. In the somatic musculature of homozygous embryos, many unfused myoblasts are present, predominantly in ventral regions (Figure 4.1-7 C). Moreover, some myofibres are thinner than normal and exhibit an unusual morphology.

The endpoints of the deficiency *Df(3L)BK10* were derived from polytene analysis. A region estimated of almost 600 kB, covering 67 genes, is deleted. Although the region partially overlaps with another analysed deficiency line, *Df(3L)XG10* (Figure 4.1-6 A), the phenotype of *Df(3L)BK10* is milder as only a few muscle fibres remain rounded and unextended (Figure 4.1-7 D).

The overlap of the deficiencies *Df(2L)Exel7042* and *Df(2L)ED680* deletes 11 genes spread over almost 60 kB. The resultant somatic muscle phenotype is mild as only the lateral transverse muscles exhibit a thinner shape than normal (Figure 4.1-7 E).

The sixth member of class IV, *Df(3R)Exel6202*, removes approximately 220 kB encompassing 35 genes. In the dorsolateral somatic musculature of mutant embryos a few mononucleated cells can be found and the muscle fibre ventral acute 3 is thinner than in the wild type embryos (Figure 4.1-7 F).

The shared feature of the phenotypes in class IV is a failure in muscle guidance and attachment to the tendon cells. Whether this is due to muscle intrinsic or extrinsic defects would need to be determined by further experiments. In addition, in many cases mononucleated myoblasts in combination with thinner muscle fibres were observed. Even though these point to a defect in myoblast fusion, only very few myoblasts were usually affected, arguing against a general failure of the fusion process.
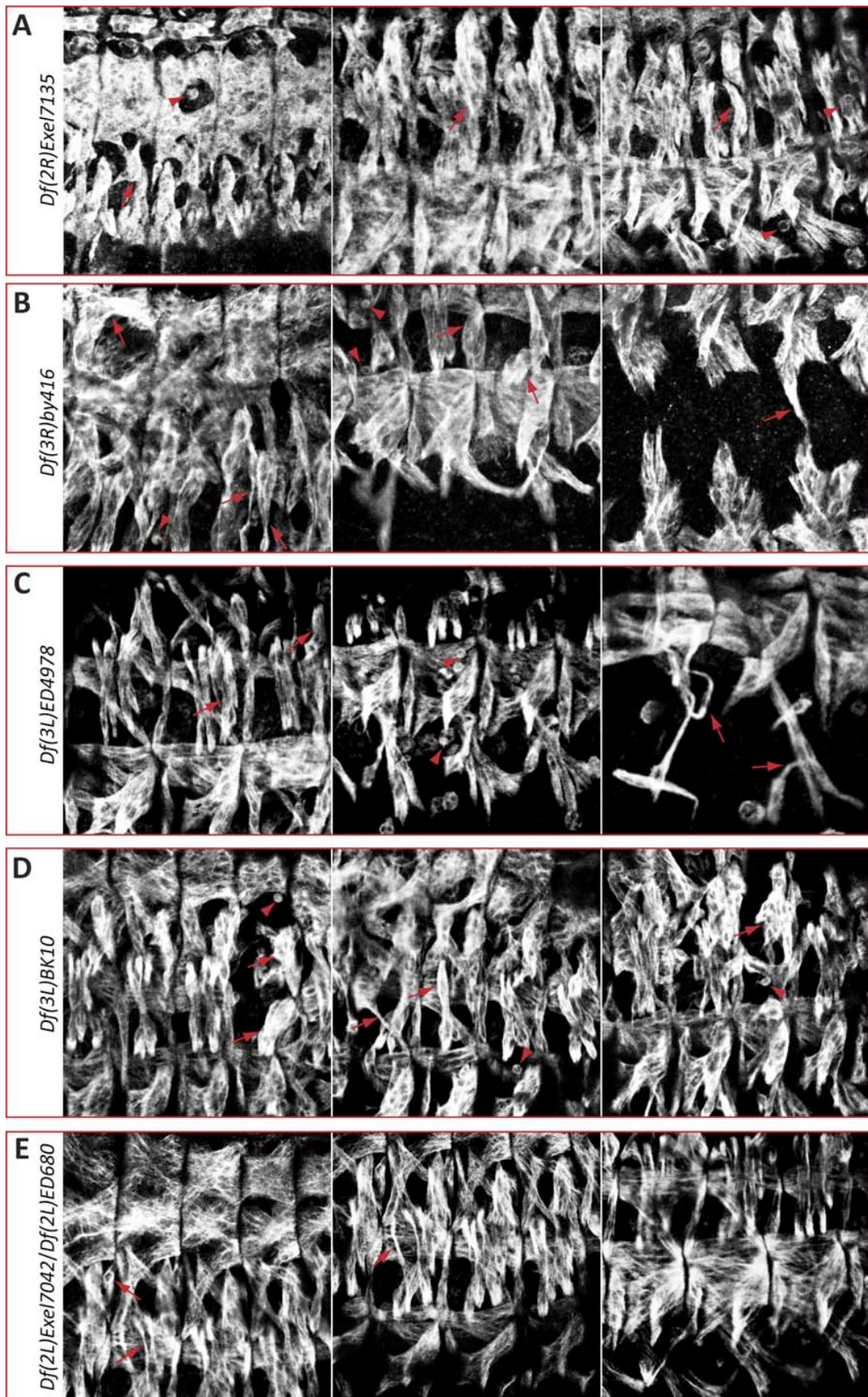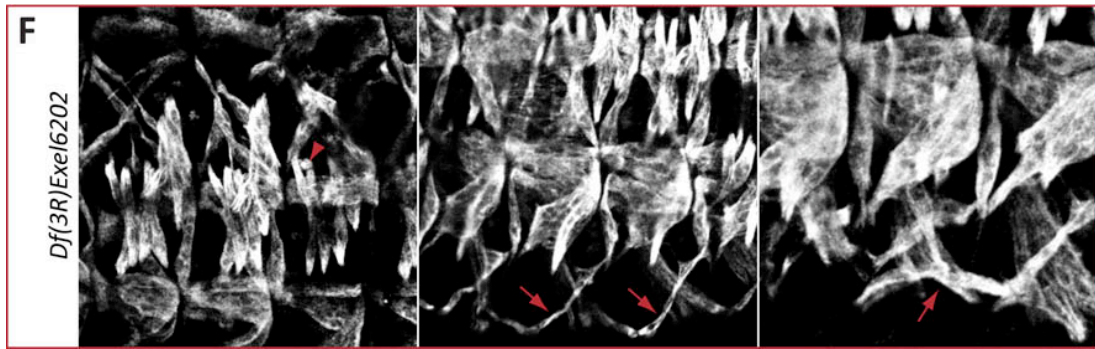
Figure 4.1-7: Phenotypic class IV: Abnormal somatic muscle guidance and attachment
Anti-β3 tubulin antibody was used to reveal somatic musculature pattern in stage 16 to 17 embryos with 6 different genetic loci individually deleted. Red arrows indicate abnormal muscle morphology and arrowheads unfused myoblasts.
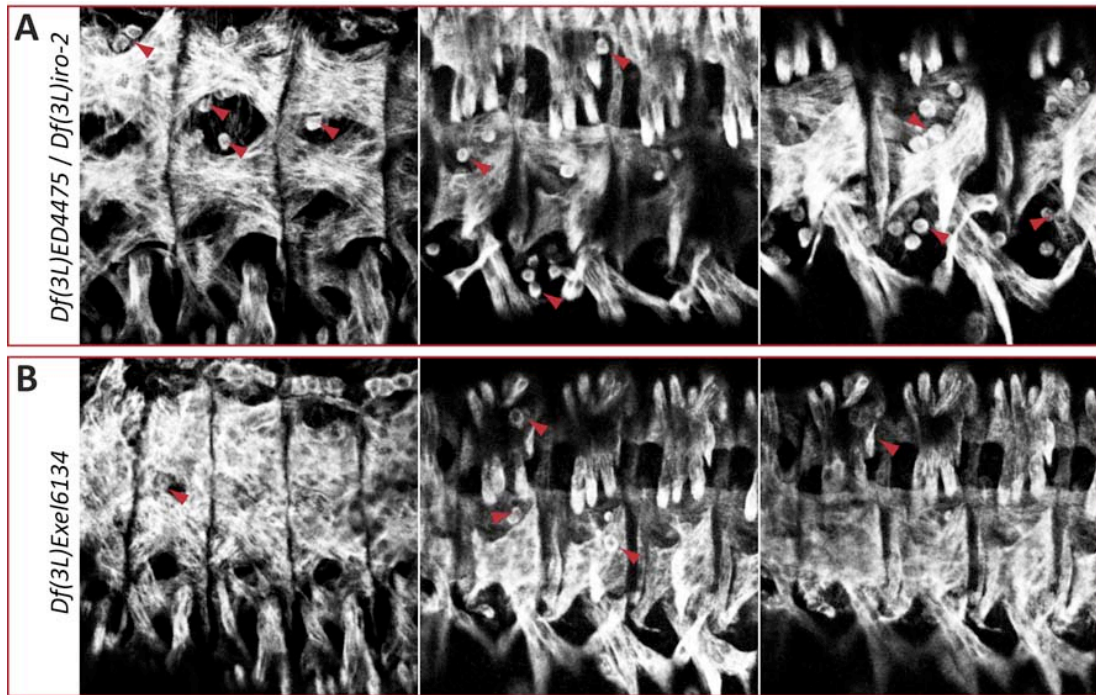
4.1.2.5 Class V: Mononucleated somatic myoblasts

In the last phenotypic class with two representatives, both the number and pattern of somatic musculature were normal, yet a high number of mononucleated, unincorporated cells recognized by the anti-β3 tubulin was revealed.

The overlap of the deficiencies *Df(3L)ED4475* and *Df(3L)iro-2* delete a 116 kB large region containing only 8 genes. Whereas the overall arrangement of the somatic muscle fibres was comparable to wild type embryos, numerous mononucleated myoblasts were present in the proximity of all muscle fibres and in all abdominal segments (Figure 4.1-8 A).

The breakpoints of *Df(3L)Exel6134* were precisely mapped to delete a 200 kB large region containing 19 genes. In the mutant embryos, the muscle number and organization seemed unaffected, but there were a few unfused myoblasts present (Figure 4.1-8 B). The phenotype is therefore similar to that of *Df(3L)ED4475/Df(3L)iro-2*, although it is substantially weaker.

These phenotypes are rather unusual as the overall muscle organization appears wild type. In contrast to more typical phenotypes seen when there is a block in myoblast fusion, the somatic muscle fibres are thick, multinucleated, and with normal morphology. The excessive unfused myoblasts could result from a defect in cell proliferation or cell death, processes whose contribution to muscle development has not been studied in detail. Alternatively, the cells might have undergone a cell fate transformation and originated from another tissue.

Figure 4.1-8: Phenotypic class V: Excessive mononucleated myoblasts in the somatic musculature

Red arrowheads point to free, unfused myoblasts labelled by anti-β3 tubulin antibody. The embryos are at the end of embryogenesis when the myoblast fusion has been completed and the muscle fibres should be fully differentiated.

4.1.2.6 Defects in heart formation

In the case of one deficiency line, a defect in heart development was also identified. The line *Df(2R)DlI-MP* deletes 34 genes in a 190 kB region and the homozygous mutant embryos form wild type somatic and visceral musculature. However, the heart tube is not properly assembled, although the cardioblasts are present (Figure 4.1-9). The precise number of cardioblasts was not assessed and so the possibility that there are more or fewer cardiac cells than in wild type embryos cannot be excluded. Consequently, the phenotype could arise from a defect in specification and/or morphogenesis.
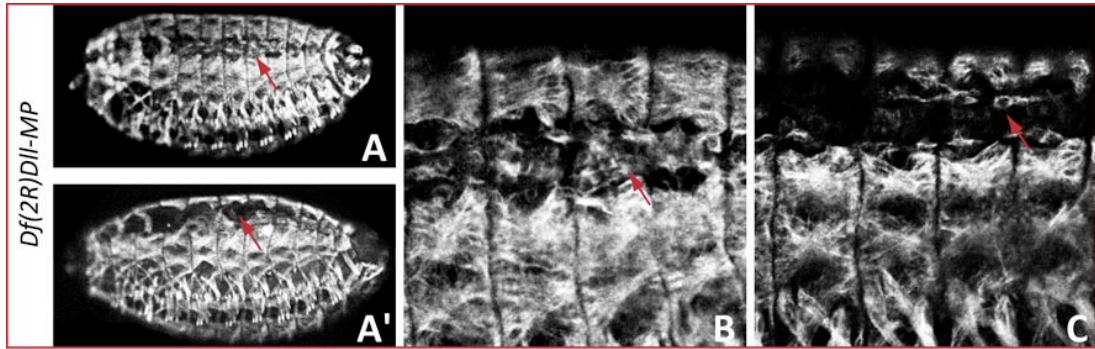
Figure 4.1-9: Defects in the heart development

(A-A') Whole-embryo views of two different stage 16 embryos showing wild type somatic musculature but absence of heart (red arrow).

(B-C) High magnification views of four hemisegments showing loose, disorganized cardioblasts in the position where heart tube should have been formed (red arrows).

4.1.2.7 Abnormal visceral muscle morphology

During the differentiation stages (15 to 17), the midgut is separated into four chambers by the formation of three constrictions. Deletion of the genomic regions of 10 deficiency lines led to abnormal gut morphology, in addition to somatic muscle defects (Table 4.1-1, data not shown), where the regular arrangement of the four chambers at stage 16 (Figure 4.1-4 G) was never observed. The specification of visceral mesoderm however seemed unaffected as the cells were clearly recognized by the anti-β3 tubulin antibody and no gaps in the tissue were visible. Since the development of midgut, originating from endoderm, and visceral mesoderm are interdependent processes [210, 211], I currently cannot conclude which tissue contributes to the phenotype without the use of additional markers.

### 4.1.3 Role of the Mediator complex in muscle development

The Mediator is a multi-subunit general transcriptional co-activator complex that is thought to link sequence specific transcription factors to RNA polymerase (see Section 1.1.3). A number of experimental observations in *Drosophila* as well as vertebrates have associated the basal transcriptional machinery with different aspects of myogenesis (see Section 1.1.3.4). However, the role of the Mediator complex in *Drosophila* muscle development has not been directly addressed. I therefore investigated the requirement of the Mediator complex, concentrating on its two subunits MED24 and MED14, for myogenesis in more detail.

4.1.3.1 Deletion of *MED14* and *MED24* by deficiency lines leads to severe somatic muscle phenotype

From the set of analyzed deficiency lines, deletion of two genetic loci resulted in a remarkably similar, severe phenotype in somatic muscle development (Figure 4.1-10). A high number of gaps in the musculature suggested an absence of muscle fibres, and many mononucleated, unfused myoblasts were detected. Many muscles failed to extend and attach to ectoderm and remained rounded. As a result of the specification and myoblast fusion defects, the size and morphology of the individual muscle fibres as well as the whole muscle organization were abnormal.

The overlap of the deficiencies *Df(3L)ZP1/Df(3L)Exel6112* (Figure 4.1-10 A) deletes 48 genes, including *MED24*. The specification and myoblast fusion defects are phenocopied, albeit to a slightly lesser extent, by *Df(3L)7C/Df(3L)ED4177* (Figure 4.1-10 B), which removes 45 annotated protein-coding genes. One of the deleted genes is *MED14*, encoding another subunit of the Mediator complex.

Mesoderm-specific knockdown of either *MED24* or *MED14* leads to lethality (Appendix 7.1). In addition, embryonic injection of double-stranded (dsRNA) against *MED24* was reported, but not demonstrated, to cause a muscle phenotype [52]. Transcripts of both subunits can be detected in the mesoderm of stage 8 to 11 embryos (Appendix 7.2). Based on these observations, I generated smaller deletion lines for *MED24* and *MED14* that would enable an examination of their role in muscle development.
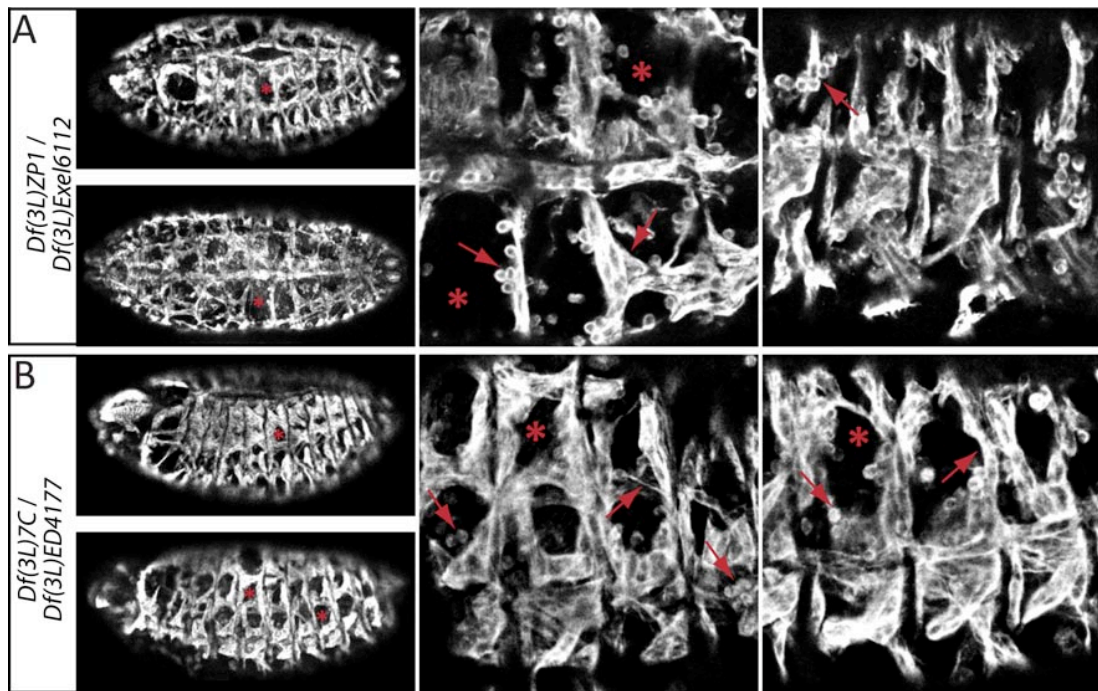
Figure 4.1-10: Deficiencies removing *MED24* and *MED14* show a similar somatic muscle phenotype

(A) The genetic locus removed by the deficiencies *Df(3L)ZP1/Df(3L)Exel6112* includes *MED24* and results in absent myofibres (asterisks), abnormal muscle fibre morphology, and presence of unfused myoblasts (arrows).

(B) The pair of deficiencies *Df(3L)7C/Df(3L)ED4177* deletes, among other genes, *MED14,* and the homozygous mutant embryos suffer from absence of myofibres (asterisks), abnormal muscle fibre morphology, as well as unfused myoblasts (arrows).

4.1.3.2 Generation of *MED24* and *MED14* deletion lines using FRT-mediated recombination

The recently established collection of lines with Flippase Recognition Target (FRT) sites randomly transposed throughout the genome enables creation of custom-made deletion lines without the need for transgenesis [200, 212]. When two lines with a FRT site are placed *in trans* and expression of the recombinase Flippase (FLP) is activated through a heat-shock inducible promoter, the region in between the FRT sites can be deleted with very high efficiency (see Section 3.2.3 for the crossing schemes). As the exact position of the FRT sites is known, successful deletions can be identified by different PCR-based methods, depending on the types of transposable elements that were used to generate the original FRT lines.

The *MED24* locus is located on the left arm of the third chromosome and the nearest available FRT sites delete a 6.3 kB region that also includes two other genes,

CG7387 and *CG7375* (Figure 4.1-11 A). Generation of two independent deletion lines was confirmed by a two-sided PCR where bands corresponding to both transposon elements were detected (Figure 4.1-11 B). Single-embryo based PCR mapping then confirmed deletion of *MED24*, *CG7837*, and *CG7375*, while the coding regions of *Autophagy-specific gene 18* (*Atg18*) and *CG8005*, the closest genes outside of the removed segment, seem to be unaffected (Figure 4.1-11 C).

The smallest possible deletion of the *MED14* locus removes an approximately 31 kB region that includes three other genes (Figure 4.1-11 D). Since the original FRT lines were created using different transposon elements, the deletion resulted in the generation of a new hybrid element. This can be then detected by PCR only in ΔMED14 lines, whereas when either of the original lines is used as a template, no product is amplified (Figure 4.1-11 E).

Finally, as the two newly generated lines are both located on the same chromosome arm, I recombined them onto the same chromosome to examine whether *MED24* and *MED14* are genetically linked. The successfully recombined lines therefore carry all the transposon elements as the original deletion lines; two-sided PCR amplifies both regions from ΔMED24 and hybrid PCR detects the hybrid transposon element present in ΔMED14 (Figure 4.1-11 F).
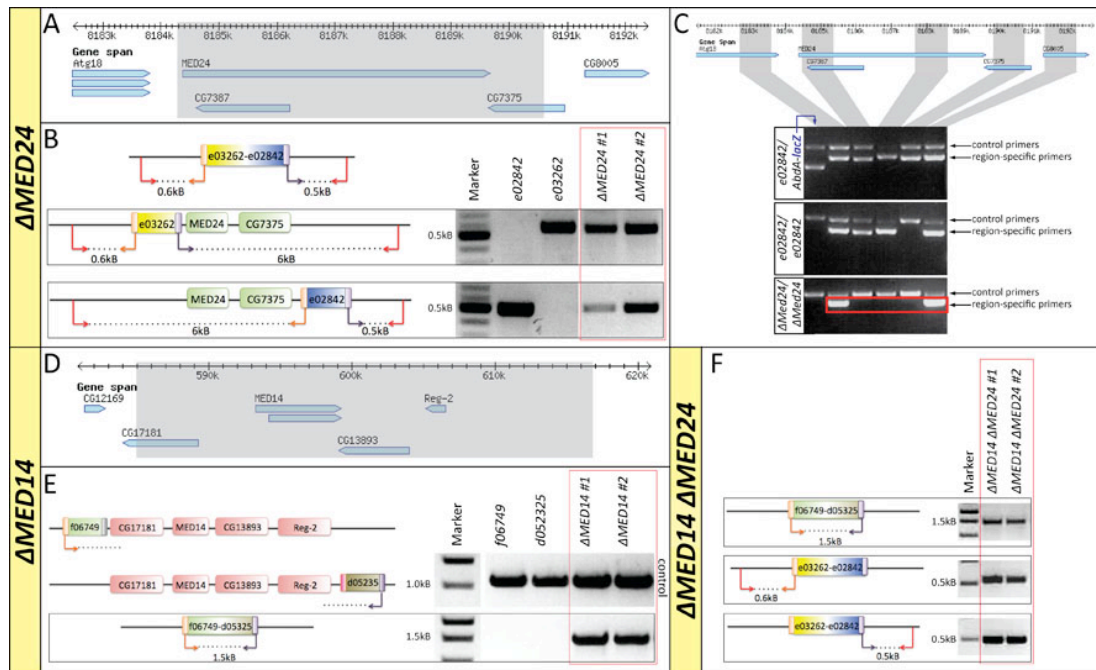

Figure 4.1-11: Generation of *MED24*, *MED14*, and *MED14 MED24* deletion lines
(A) Genomic locus of *MED24* (generated with GBrowse [154]). The region in grey is deleted in ΔMED24.

(B) Schematic representation of the final line, ∆*MED24*, and the two original lines, *e02842* and *e03262*. Two pairs of primers, one of which aligns to the transposon backbone (orange, purple) and one to the genomic sequence (red), were used to confirm successful deletion within two lines. While each original line gives rise to a single PCR product, only ∆*MED24* results in both.

(C) Single-embryo based PCR mapping of the sequences deleted in ∆*MED24*. Primers directed against *lacZ* sequence were used to identify homozygous mutant embryos. Embryo with one or two copies of the original allele *e02842* contains all the regions (with the exception of *CG7375* in the fifth lane of homozygous *e02842* embryo as it falls within the transposon insertion site). A ∆*MED24* homozygous embryo lacks only the regions that were predicted to be deleted (grey in A).

(D) Genomic locus of *MED14* (generated with GBrowse [154]) and the region (in grey) deleted in ∆*MED14*.

(E) Hybrid PCR confirms successful recombination in two different ∆*MED14* lines. As the two original lines *f06749* and *d05235* carried different transposon backbone (unlike in ∆*MED24*), deletion of the intermediate region gave rise to a hybrid element that is present only in the deletion line and in neither of the two original lines.

(F) Confirmation of successful recombination of ∆*MED24* and ∆*MED14* onto the same chromosome. Two candidate lines bear all the elements of the individual lines and could be therefore detected by the PCR methods used in (B) and (E).

4.1.3.3 Deletion of *MED24* results in the formation of excessive myoblasts

While the deficiencies *Df(3L)ZP1* and *Df(3L)Exel6112* combined *in trans* led to severe muscle specification and myoblast fusion defects (Figure 4.1-10 A), embryos homozygous for the small deletion line of *MED24* exhibit normal somatic musculature patterning (Figure 4.1-12 A). However, presence of unfused myoblasts, a hallmark of *Df(3L)ZP1/Df(3L)Exel6112*, is partially phenocopied in ∆*MED24* embryos. The number of mononucleated, β3 tubulin labelled cells within the somatic musculature was manually counted in 30 hemisegments of at least 10 different embryos. While the wild type embryos occasionally contain free cells, in ∆*MED24* their number is significantly higher (Figure 4.1-12 D).

Deletion of *MED14* does not cause any apparent defects in the development of any muscle type (Figure 4.1-12 B), which is in contrast with the severe phenotype of the deficiency line (Figure 4.1-10 B). It also does not seem to genetically interact with *MED24* as a double mutant line removing both *MED14* and *MED24* shows a phenotype comparable to that of ∆*MED24*, although a weak genetic interaction especially in the dorsal muscles cannot be excluded (Figure 4.1-12 C, D).

It therefore appears that the original phenotypes observed in the deficiency lines were caused by genes other then *MED24* and *MED14*. Nevertheless, *MED24* seems to be involved in the control of cell number as in ∆*MED24* embryos, the number of

myoblasts is significantly higher compared to wild type embryos (Figure 4.1-12 D) whereas the number and morphology of the muscle fibres remain intact (Figure 4.1-12 A). The suggested role for *MED24* in controlling the cell number is further supported by its recently identified requirement for cell death in the salivary gland cells [213].
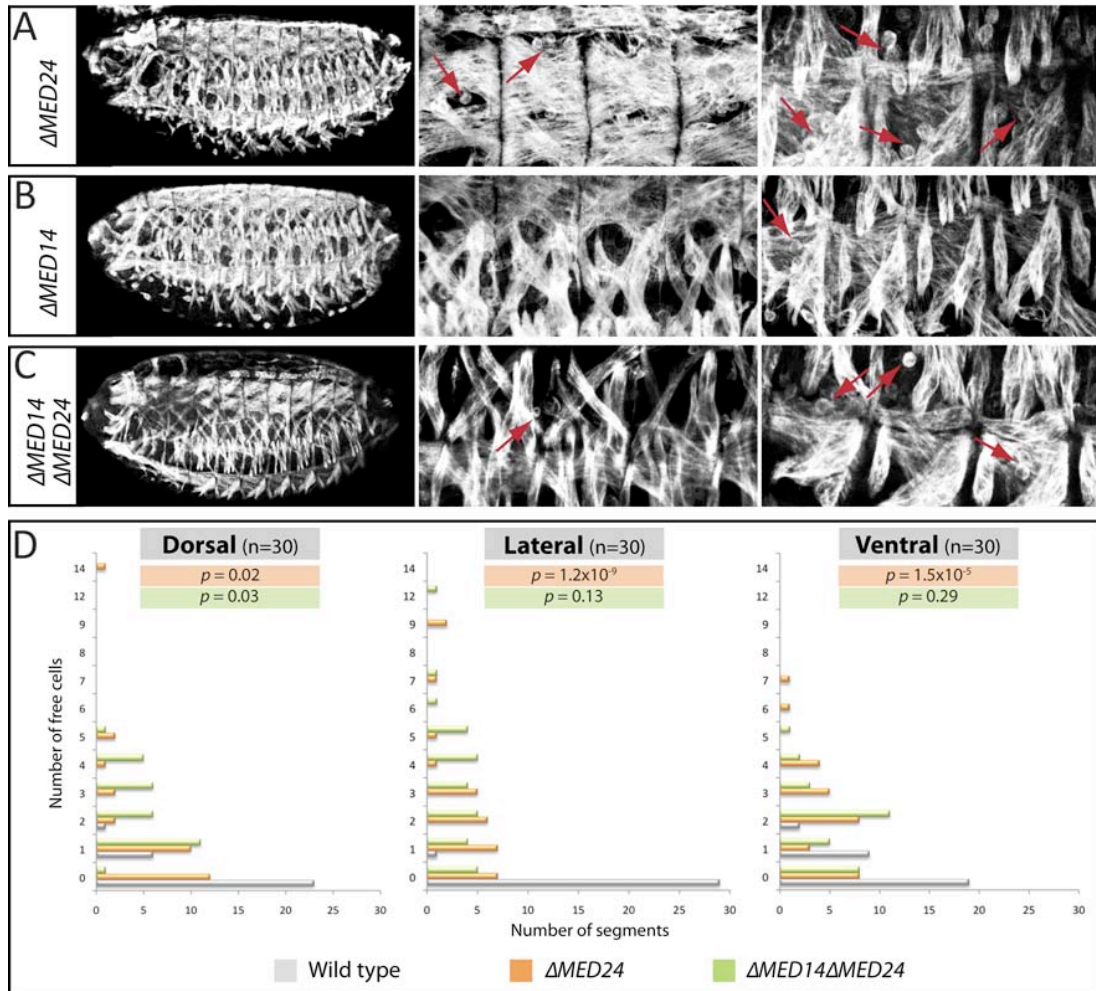
Figure 4.1-12: Muscle phenotypes in small deletion lines removing *MED14*, *MED24*, and both genes together

(A) The small deletion line Δ*MED24* displays a high number of mononucleated myoblasts formed (arrows) while the number and organization of muscle fibres are intact.

(B) Deletion of *MED14* by Δ*MED14* does not seem to affect muscle development, with the exception of sporadically observed unfused myoblasts (arrows).

(C) Embryos mutant for both *MED24* and *MED14* show many unfused myoblasts (arrows).

(D) Quantification of the free, mononucleated β3 tubulin-positive cells in wild type (grey), Δ*MED24* (orange), and Δ*MED14* Δ*MED24* (green) embryos. The cells were manually counted in three different somatic muscle areas and in 30 hemisegments (within at least 10 different embryos per genotype). The significance values comparing Δ*MED24* to wild type (orange) and Δ*MED14*Δ*MED24* to Δ*MED24* (green) were calculated using exact Wilcoxon rank sum test.

4.2 The role of Tramtrack in *Drosophila* muscle development

The transcriptional repressor Tramtrack was one of the novel mesodermal regulators identified within the molecular screen. The contribution of Tramtrack to *Drosophila* muscle development was characterized in more detail due to the following reasons:

o Multiple CRMs in the vicinity of *ttk* locus are bound by key mesodermal TFs (Twi, Mef2, Tin, Bin, Bap) and at least two of them have mesodermal *in vivo* activity [108, 214].

o Putative Ttk69 binding site is significantly enriched in early mesodermal CRMs.

o Ttk69 (but not Ttk88) mutants are recessive lethal at embryonic stage and a neuronal marker was detected in the somatic musculature of amorphs [155].

o Ttk is a well-described TF that is potentially expressed in mesoderm.

o Ttk is a dedicated transcriptional repressor and the current knowledge on repression within the myogenic network is very limited.

4.2.1 Analysis of Tramtrack expression and its loss-of-function and gain-of-function phenotypes

### 4.2.1.1 Expression of Tramtrack69 mRNA and protein in the mesoderm

As the reports on the expression of Ttk69 in the mesoderm were incoherent [150, 152], I analyzed the embryonic expression of both Ttk69 mRNA and protein (Figure 4.2-1). The expression in mesoderm initiates at early stage 11 when *ttk69* mRNA has detectable expression in the visceral muscle primordium (Figure 4.2-1 A-A"). The expression of the Ttk69 protein lags behind as it is present in the visceral muscle only from late stage 11 (Figure 4.2-1 E-E"). Both Ttk69 mRNA (Figure 4.2-1 D-D") and protein (Figure 4.2-1 F-F") are maintained in the visceral musculature until the very end of embryogenesis. In addition, the Ttk69-specific antibody identifies Ttk69 protein in the midgut endoderm primordia during multiple stages of development (Figure 4.2-1 D-F).

Expression of Ttk69 in the somatic muscle primordium is, on the other hand, very transient and therefore also very difficult to analyze. *ttk69* mRNA is present in a

subset of somatic muscle precursor cells at stage 11 (Figure 4.2-1 B-B"), corresponding to the time period of progenitor selection and founder cell specification. However, already at stage 12 (Figure 4.2-1 C-C") there is no *ttk69* mRNA detectable in the somatic mesoderm.

An additional tool to asses the expression of Ttk69 is a previously characterized enhancer *ttk-VME* [108], located upstream of the gene locus. At stage 11, the *in vivo* activity of this enhancer recapitulates the expression of endogenous Ttk69 protein in mesoderm and trachea (Figure 4.2-1 G-G") and it co-localizes with Mef2 protein in both somatic and visceral mesoderm (Figure 4.2-1 H-H").

The very transient expression of Ttk69 in the somatic mesoderm primordium at the stages of specification suggests its involvement in this process and predicts, that misregulation of *ttk69* in the mesoderm could be detrimental for its development.
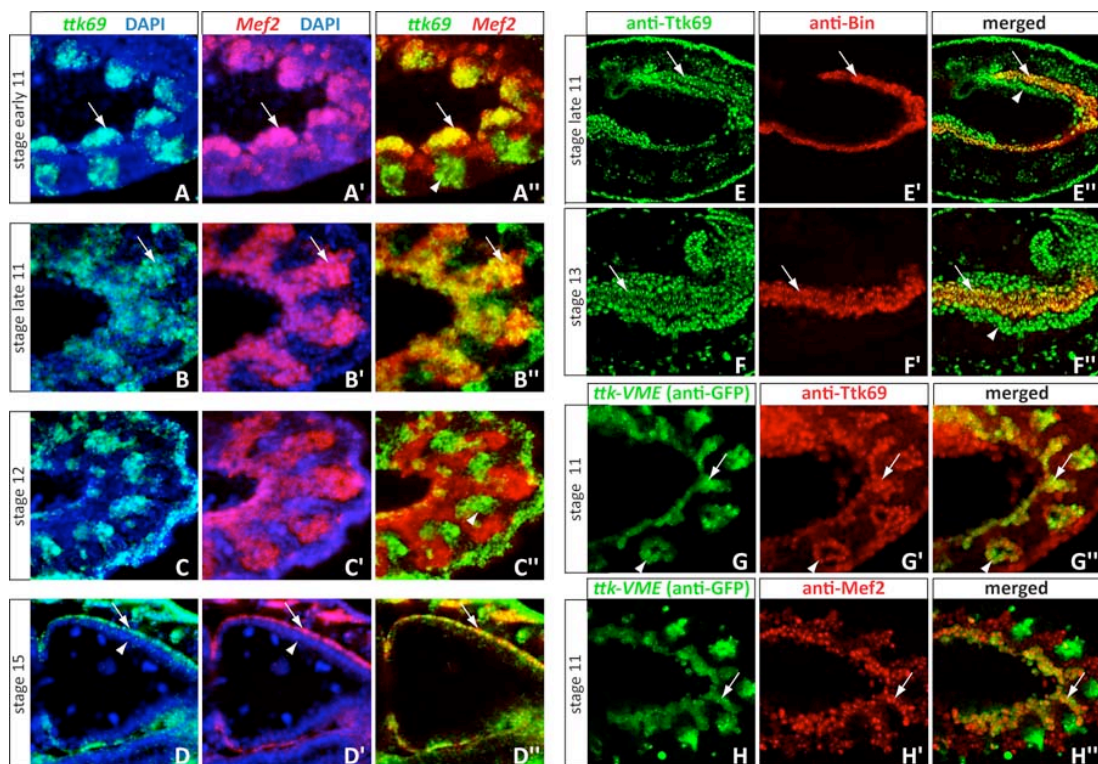
Figure 4.2-1: Expression of Ttk69 mRNA, protein and enhancer

(A-D) *In situ* hybridization against *ttk69* (green) and *Mef2* (red) mRNA in wild type embryos. During stage 11, *ttk69* co-localizes with *Mef2* in the visceral mesoderm (arrows in A-A") and subset of somatic mesoderm (arrows in B-B"). At stage 12 (C-C") there is no overlap between *ttk69* and *Mef2* as *ttk69* expression is maintained only in the trachea (arrowheads in A" and C"). At stage 15, *ttk69* and *Mef2* continue to be co-expressed in the visceral muscle (arrows in D-D") lining the gut (arrowheads in D and D'). At all stages 4',6-diamidino-2-phenylindole (DAPI) was used to follow the nuclei. Each image is an average intensity Z-projection of several focal planes.

(E-F)   Ttk69 protein (green) co-localizes with the marker of visceral mesoderm, Bin (red), at stage 11 (arrows in E-E'') and 13 (arrows in F-F''). Ttk69 protein is also expressed in the gut (arrowheads in E'' and F'').

(G-H)   Previously characterized *ttk* enhancer line *ttk-VME* [108] was subjected to fluorescent double immunostain. GFP expression (green) recapitulates endogenous Ttk69 expression (red) in mesoderm (arrows in G-G'') and tracheal placodes (arrowheads in G-G'') and co-localizes with Mef2 protein in the somatic and visceral muscle primordia (arrows in H-H'') of stage 11 embryos.

4.2.1.2 *ttk* loss-of-function phenotype

The amorphic allele *ttk^{D2-50}* was isolated in a genetic ethyl methanesulfonate (EMS) screen for mutants with defects in embryonic CNS development [215]. The precise lesion has not been mapped, but based on the severity of the CNS phenotype, *ttk^{D2-50}* was classified among the strongest *ttk* mutant alleles [155]. I confirmed that *ttk^{D2-50}* is a protein null allele by immunostaining the embryos with an antibody raised against the zinc fingers of the Ttk69 protein (data not shown). As the CNS phenotype in *ttk^{D2-50}* is identical to Ttk69-specific allele *ttk^{1e11}* [155, 178], *ttk^{D2-50}* is a null allele of Ttk69, but potentially affecting expression of the Ttk88 isoform as well. The Ttk88-specific allele *ttk^{1}* [165] is however homozygous viable and embryos develop wild type musculature (not shown). *ttk^{D2-50}* allele is hence predominantly a loss-of-function allele for Ttk69, although contribution of possible Ttk88 mutation to the observed phenotypes cannot be excluded.

Homozygous *ttk^{D2-50}* mutant embryos were unambiguously identified by the absence of *lacZ* or *gfp*, which was placed on the balancer chromosome, as well as the previously described defects in head involution and dorsal closure [163]. The muscle defects in *ttk^{D2-50}* homozygous embryos were indistinguishable from embryos carrying *ttk^{D2-50}* allele in *trans* to deficiency *Df(3R)awd-KRB*, removing the whole locus, in agreement with the amorphic nature of the allele. Transheterozygous embryos were stained with an anti-β3 tubulin antibody that marks all three muscle types of the *Drosophila* embryo ([207], Figure 4.1-4).

The most severely affected muscle tissue is the somatic musculature which shows abnormalities from stage 14 (Figure 4.2-2 A-B''). While in wild type embryos myoblast fusion is underway at this stage, in *ttk^{D2-50}*/*Df(3R)awd-KRB* transheterozygous embryos myoblast fusion is impaired and most myoblasts remain round and mononucleated. At later stages the cells do undergo partial fusion, however rather than forming extended myotubes they are organized in large round

clusters (Figure 4.2-2 C-D"). These clusters of myoblasts do not migrate or attach to ectodermal cells, making their identity unclear. The observed defects are rather atypical as normally a failure in myoblast fusion is accompanied by unfused, round shaped fusion competent myoblasts (FCMs) that undergo apoptosis and thin myofibres with correct muscle attachments, generated from the founder cells (FCs). In the case of *ttk* mutants, however, the unfused FCMs do not seem to apoptose but rather aggregate around the FCs, which in turn fail to differentiate into correctly positioned and attached myofibres.

Another perturbed tissue in *ttk*-deficient embryos is the gut which at stage 16 has an abnormal morphology and fails to form the characteristic midgut constrictions (Figure 4.2-2 D'). As Ttk69 is expressed in both the visceral muscle and the gut primordia (Figure 4.2-1 D-F), this defect may be caused by its function in either, or both, tissues. An immunostain with anti-Fasciclin III antibody revealed that the number as well as the morphology of visceral myoblasts are normal (not shown), suggesting that specification of the visceral mesoderm is not affected. Specification of heart precursors also occurs normally (not shown), although a proper dorsal vessel is never formed in ttk^{D2-50} homozygous embryos as a secondary consequence of a dorsal closure failure.

In multiple developmental contexts, Ttk69 negatively regulates cell cycle and in *ttk69* mutants excessive rounds of mitosis occur (see Table 1.4-2). Since both subpopulations of fusing cells, FCs and FCMs, undergo mitosis during stages 11 and 12 [137], the observed myoblast fusion defect in *ttk69* mutants could be explained by an aberrant cell cycle. To assess this, I used an antibody directed against the phosphorylation of Serine 10 residue on histone H3, a widely used marker of mitosis [216]. At stage 11 and 12, the number of muscle cells marked by anti-phosphorylated histone H3 antibody in wild type and *ttk* mutant embryos is comparable (Figure 4.2-2 E-H), suggesting that the cell cycle is not perturbed in *ttk69*-deficient somatic myoblasts.

Previous studies on the nervous system development reported that the neuronal marker 22C10 is aberrantly expressed in the somatic muscle in *ttk69* loss-of-function mutant alleles [155, 156]. To verify this finding, I repeated the experiment and double immunostained ttk^{D2-50} embryos with an anti-Mef2 antibody, labelling all three muscle types. As expected, the microtubule-associated protein Futsch, the antigen of 22C10, and the nuclear Mef2 are co-expressed in a subset of somatic

muscle fibres and also in the visceral mesoderm (Figure 4.2-2 I, I'). Futsch is also misexpressed in the heart, which is surprising considering that there is no detectable expression of Ttk69 in these cells (Figure 4.2-2 I''). As 22C10 is a marker of differentiated neurons, its derepression in *ttk* mutants could be due to partial transformation of muscle to a neuronal identity. To investigate this possibility, I analyzed the protein expression of key neuronal regulators (Asense, Achaete, Repo) in *ttk* mutants, but none of these proteins seem to be derepressed in the musculature (not shown). Therefore, it cannot be concluded that myoblasts devoid of Ttk69 have a partial neuronal cell fate. Nevertheless, the presence of Futsch alone might interfere with myogenesis as it is involved in axonal pathfinding [217, 218], a process that is at the molecular level parallel to myotube guidance (see Section 1.3.6).
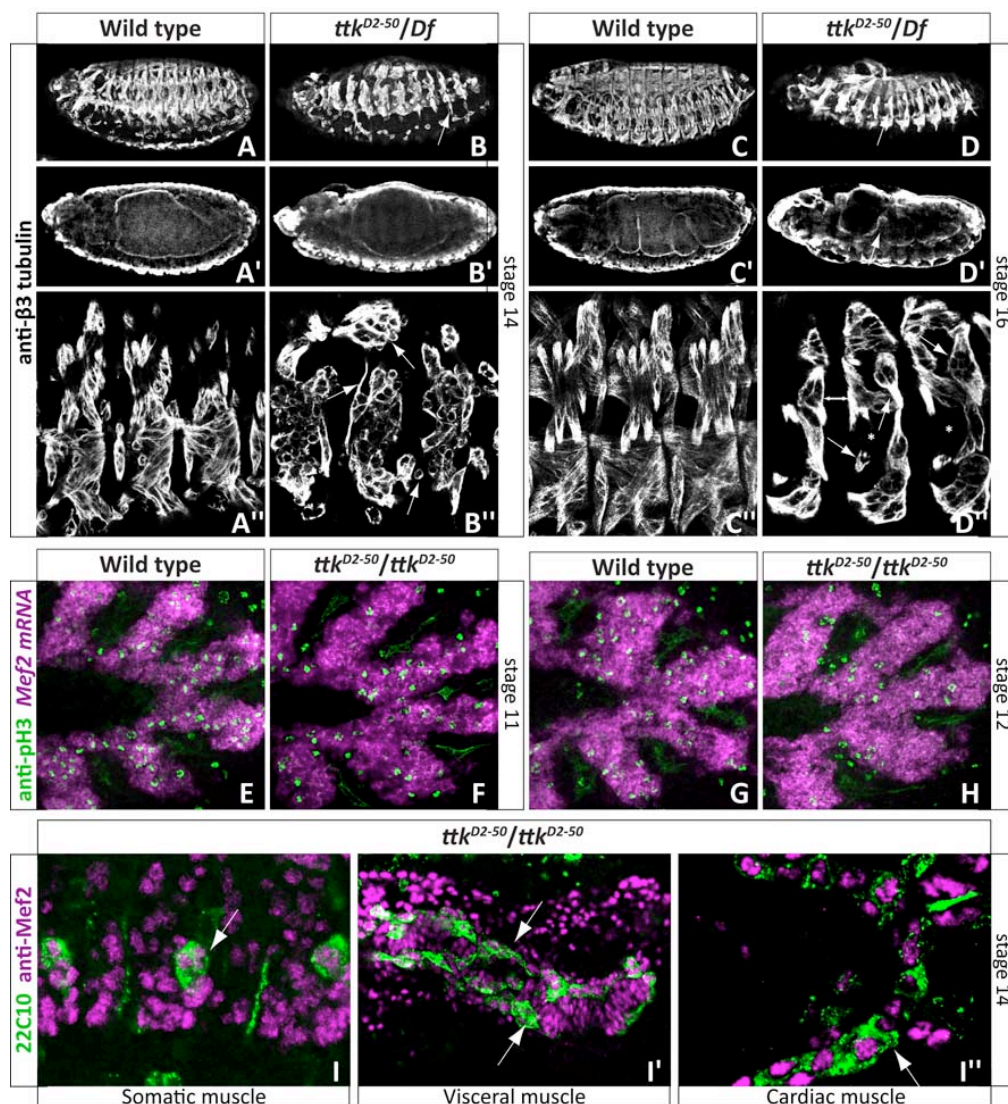
**Figure 4.2-2: Ttk is essential for normal somatic muscle development**

(A-D) Immunostain of wild type and *ttk*^D2-50^/*Df(3R)awd-KRB* embryos with anti-β3 tubulin antibody to visualize all three muscle tissues. At stage 14 (A-B), *ttk*-deficient

myoblasts fail to fuse and instead remain mononucleated, form aggregates, and have aberrant shape (arrows in B and B"). In stage 16 embryos (C-D), *ttk* mutants have abnormal gut morphology (arrow in D') and myofibres do not attach to tendon cells (increased distance between two hemisegments in D"). Arrows in D and D" point to fibres with abnormal morphology and asterisks mark areas normally filled with myofibres.

(E-H) Immunostain against phosphorylated histone H3 (pH3, green), a mitosis marker, combined with *in situ* hybridization against *Mef2* (magenta) in wild type (E, G) and *ttk*$^{D2-50}$ homozygous embryos (F, H). No apparent differences in the number of pH3 and *Mef2* double-stained cells could be observed at stage 11 (E-F) or 12 (G-H).

(I) In stage 14, *ttk*$^{D2-50}$ homozygous embryos, the cytoplasmic protein Futsch (22C10 antibody, green) is co-expressed with nuclear Mef2 (magenta) in the cells of somatic (I), visceral (I'), as well as cardiac (I") musculature.

As *ttk* mutants exhibit impaired myoblast fusion (Figure 4.2-2 B"), I analyzed the expression of genes essential for this process by *in situ* hybridization. The number of cells expressing FCM-specific genes *lame duck* (*lmd*) and *sticks and stones* (*sns*) is substantially reduced in *ttk*$^{D2-50}$ homozygous embryos compared to their wild type counterparts (Figure 4.2-3).
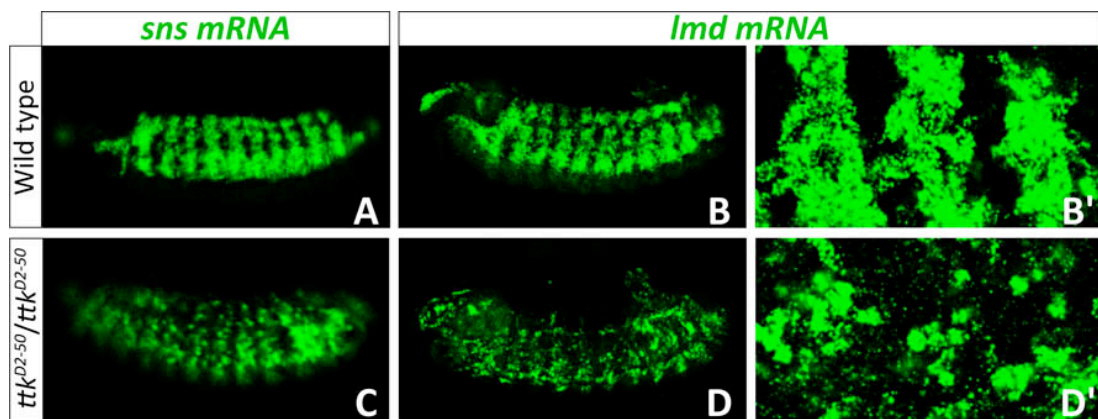
Figure 4.2-3: Reduced expression of two important FCM-specific genes in *ttk* mutants

(A-B) In wild type stage 13 embryos, strong levels of *sns* (A) and *lmd* (B, B') mRNA can be detected by fluorescent *in situ* hybridization

(C-D) In *ttk*$^{D2-50}$ homozygous embryos (C-D'), *sns* (C) and *lmd* (D, D') are expressed in fewer cells compared to wild type embryos.
High power images of 3 segments of stage 13 embryos show ventral and lateral somatic musculature (B', D').

This was surprising as there is no evidence that Ttk69 can directly activate transcription. One possible explanation for the reduced expression of FCM-restricted genes is a decrease in the number of FCMs at the expense of FCs. To assess this, I characterized the number of FCs by recombining the *ttk*$^{D2-50}$ allele with rP298-lacZ, a transgenic *Drosophila* line isolated in a *P*-element enhancer trap screen [219, 220]. The insertion point of the *P*-element has been mapped to the upstream region of the

gene *dumbfounded* (*duf*), encoding an FC-specific Ig-like transmembrane protein mediating recognition of FCMs [221]. The *lacZ* reporter expression is restricted to the FCs within somatic and visceral muscle from stage 11 [222] and rP298-lacZ is thus a widely used marker of FCs.

Consistent with the decreased expression of FCM genes, the FC-specific marker rP298-lacZ is misexpressed in ttk^{D2-50} homozygous embryos, being detected in the majority of cells in both the somatic and visceral mesoderm (Figure 4.2-4 B-B"). Moreover, it is also misexpressed in additional cells outside the mesoderm (revealed by a lack of β3 tubulin staining), such as midgut (Figure 4.2-4 B') and ectoderm (Figure 4.2-4 B"), while in wild type embryos its expression is restricted to muscle founder cells (Figure 4.2-4 A-A").

The strong expression of the rP298-lacZ reporter outside the mesoderm impedes a direct visual evaluation of its expression in myoblasts. To address this, I quantified the number of *lacZ*-positive cells only within the somatic mesoderm. For this purpose, ttk^{D2-50} heterozygous and homozygous embryos were triple immunostained with antibodies against β-Galactosidase (the product of rP298-lacZ), β3 tubulin (a general muscle marker), and GFP (visualizing the balancer chromosome). A series of images at multiple focal planes were then taken within the abdominal segments A3 to A6 in 13 different stage 12 to 13 embryos.

In wild type embryos, the number of rP298-lacZ marked nuclei is highly dynamic [222]. Initially it increases as the progenitors divide to form the FCs and then its expression is acquired by FCMs upon fusion with FCs. Before the onset of fusion, the number of rP298-expressing cells hence should not exceed 30 as there are 30 distinct muscle fibres, each originating from a single FC. Consistent with this, in ttk^{D2-50} heterozygous embryos, the number of counted FCs ranged from 15 to 30 and did not vary significantly across different segments (Figure 4.2-4 C). In ttk^{D2-50} homozygous embryos, the number of rP298-lacZ and β3 tubulin positive somatic myoblasts was also comparable between individual segments. However, in contrast to the heterozygous embryos, the number of rP298-lacZ expressing somatic myoblasts in ttk^{D2-50} homozygous embryos was significantly higher, in the range from 31 to 52 (Figure 4.2-4 C). This result confirms that there are more rP298-lacZ positive somatic myoblasts in embryos devoid of Ttk69. Even though they are marked by a

FC-specific reporter, their exact developmental identity is still not clear, hence they are referred to as FC-like cells.
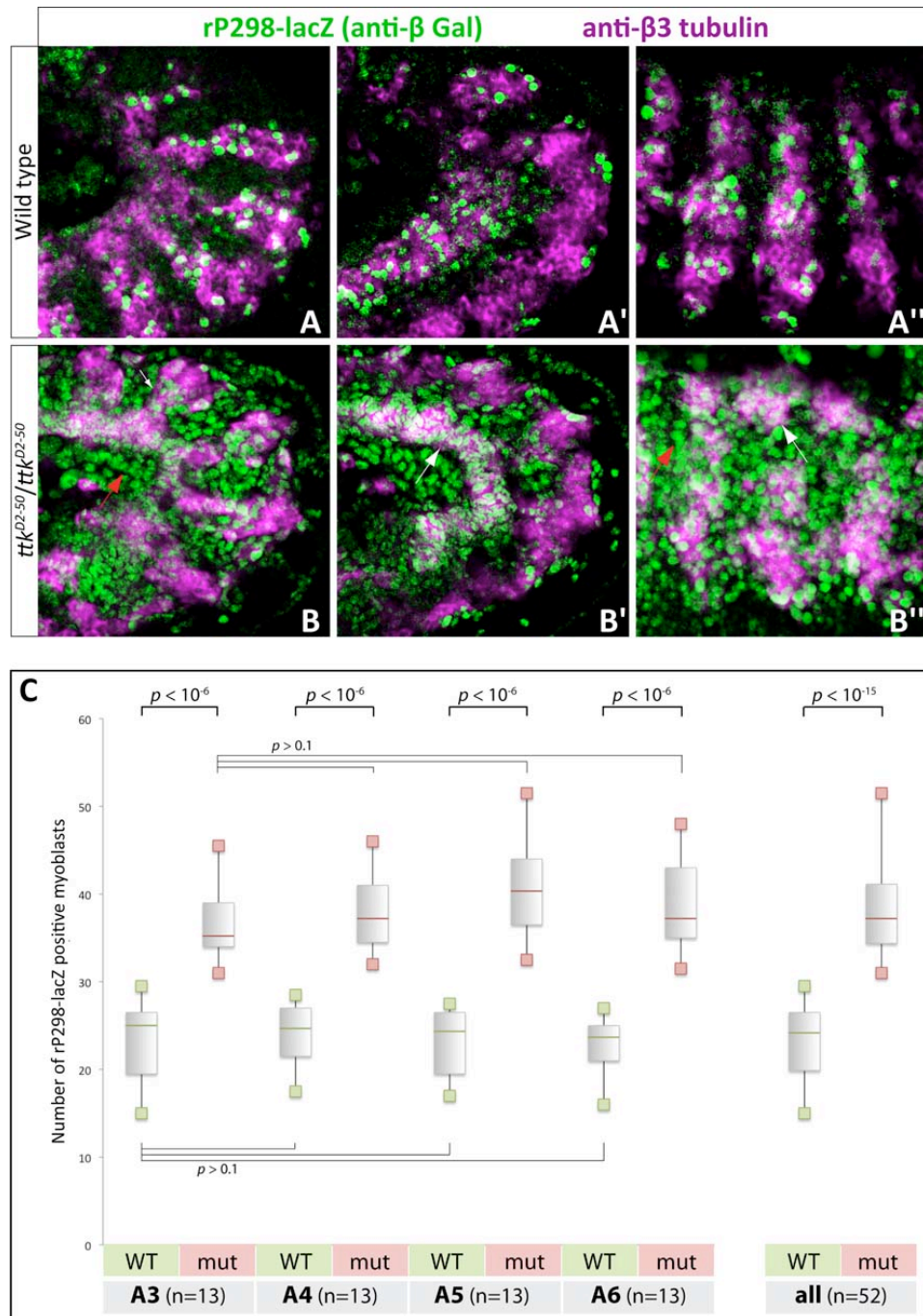
Figure 4.2-4: Ectopic founder cell-like cells in *ttk* mutants

(A-B) Compared to wild type embryos (A-A"), ttk^{D2-50} homozygous embryos (B-B") display increased levels of founder cell marker, rP298, visualized by an antibody against β-Galactosidase, both within (white arrows) and outside (red arrows) of mesoderm at stage 12 (A-A', B-B') and 13 (A", B"). Anti-β3 tubulin antibody (magenta) outlines all somatic (A, A", B, B") and visceral (A', B') mesodermal cells.

(C) Quantification of rP298-lacZ positive myoblasts in 4 segments of stage 12 to 13 ttk^{D2-50} heterozygous (*WT*) and homozygous (*mut*) embryos. Significance *p* values were calculated using exact Wilcoxon rank sum test. In total, 52 segments in 13 different embryos per genotype were analysed (*all*).
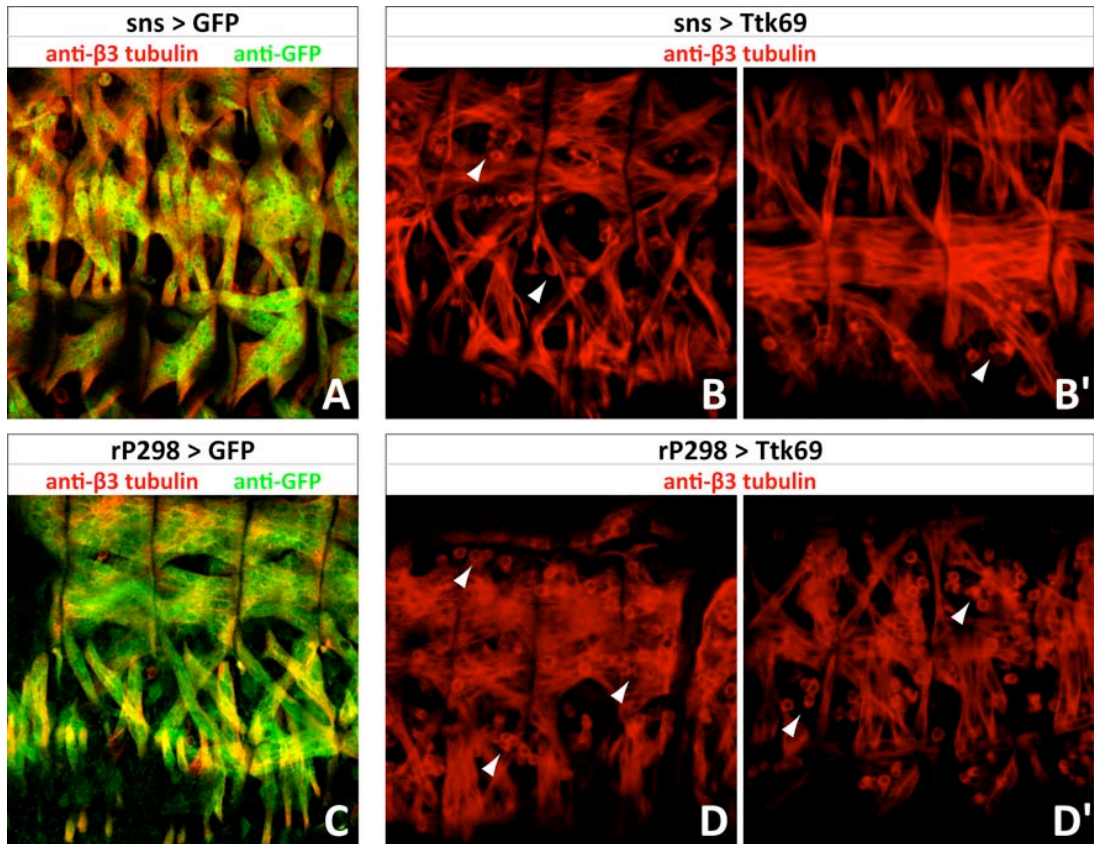
4.2.1.3 Ectopic expression of Ttk69 in founder cells and fusion competent myoblasts

To further explore the role of Ttk69 in myogenesis, I utilized the yeast Gal4-UAS system [223] to induce tissue-specific expression of Ttk69. Expression of Ttk69 in FCMs was achieved with two different FCM-specific sns-Gal4 driver lines [224, 225]. Combination of the lines with UAS-GFP resulted in expression of GFP protein in the visceral and somatic mesoderm from stage 13 and 14, respectively, and had no effect on muscle development (Figure 4.2-5 A). However, specific expression of Ttk69 in FCMs [155] with one of the sns-Gal4 driver lines [225] led to pupal lethality and a mild myoblast fusion phenotype in some, but not all, embryos (not shown). Embryos in which Ttk69 expression was induced with the other sns-Gal4 line [224] died at the end of embryogenesis and exhibited impaired myoblast fusion (Figure 4.2-5 B, B'). The defect was relatively mild as only some unfused myoblasts were observed and the size of the myotubes did not appear substantially reduced compared to wild type embryos (Figure 4.2-5 A-B').

A FC-specific driver line (rP298-Gal4) was generated from the FC-specific reporter line rP298-lacZ (described above) by replacing the original *P*-element transposon with a *P*-element carrying Gal4 [226, 227]. rP298-Gal4 drives expression of GFP protein in the somatic and visceral mesoderm from stage 12 (Figure 4.2-5 C). Ectopic expression of Ttk69 in FCs gives rise to embryonic lethal phenotype with a severe myoblast fusion defect (Figure 4.2-5 D, D'). The FCs do not fuse with surrounding FCMs and so form very thin, but correctly attached, mini-muscles, while the FCMs remain mononucleated and later undergo apoptosis. The thin muscle fibres maintain their correct shape, orientation, and ectodermal attachments, indicating that ectopic Ttk69 does not interfere with the expression of founder cell identity genes. However, due to the time delay introduced by the UAS-Gal4 system, ectopic Ttk69 may only be functional after muscle identity has been specified.

Consequently, it appears that Ttk69 is more detrimental when it is expressed in the FCs, while its forced expression in FCMs has a less dramatic effect. Given the decreased number of FCMs and ectopic FC-like cells in *ttk* mutants (Figures 4.2-2 and 4.2-3), the normal function of Ttk69 may be to repress FC-specific genes within FCMs. This model would explain why ectopic expression of Ttk69 in FCMs would result in only minor defects, which are most likely caused by prolonged expression of

Ttk69 in somatic mesoderm after stage 11, compared to wild type embryos (Figure 4.2-1 B).


Figure 4.2-5: Ectopic expression of Ttk69 in founder cells and fusion competent myoblasts

(A) Expression of GFP was driven in the FCMs with an sns-Gal4 driver [224] and collected embryos were double immunostained with anti-β3 tubulin and anti-GFP antibodies. As expected, somatic musculature develops normally without any apparent abnormalities.

(B) Sns-Gal4 driver line [224] was used to induce expression of Ttk69 in FCMs. All embryos developed normal heart and visceral muscle (not shown). The somatic muscle fibres were correctly specified but myoblast fusion partially failed as there were some unfused myoblasts present (arrowheads).

(C) Embryos with GFP driven by the FC-specific Gal4 driver rP298 show normal somatic musculature pattern and strong GFP expression in the muscle fibres.

(D) When Ttk69 is ectopically expressed in FCs, only limited myoblast fusion occurs as most of the somatic muscle cells remain unfused (arrowheads). The heart and the visceral muscle are unaffected (not shown).

4.2.1.4 Ectopic expression of Ttk69 in presumptive mesoderm

As the phenotype of ttk^{D2-50} embryos cannot be entirely explained by a myoblast fusion defect, Ttk69 was ectopically expressed in unspecified mesoderm from stage 9 with a combined twi, 24B-Gal4 driver [81, 223] (Figure 4.2-6). Such

embryos suffer from severe specification defects in all three muscle tissues. In most embryos only very few, or even no, cardioblasts are present and there are also gaps in the visceral mesoderm, as revealed by anti-Fasciclin III antibody (Figure 4.2-6 B-D). As a result, there are no midgut constrictions formed and it remains a one-chambered organ (Figure 4.2-6 A). Within the somatic musculature, there are hardly any correctly specified and differentiated myofibres. Instead, there are many unfused, mononucleated cells and a few mini-muscles without a clear shape, size, or attachment (Figure 4.2-6 A', A"). These observations suggest that premature Ttk69 expression counteracts specification of all three muscle identities and hence very tight control of its wild type expression is essential.
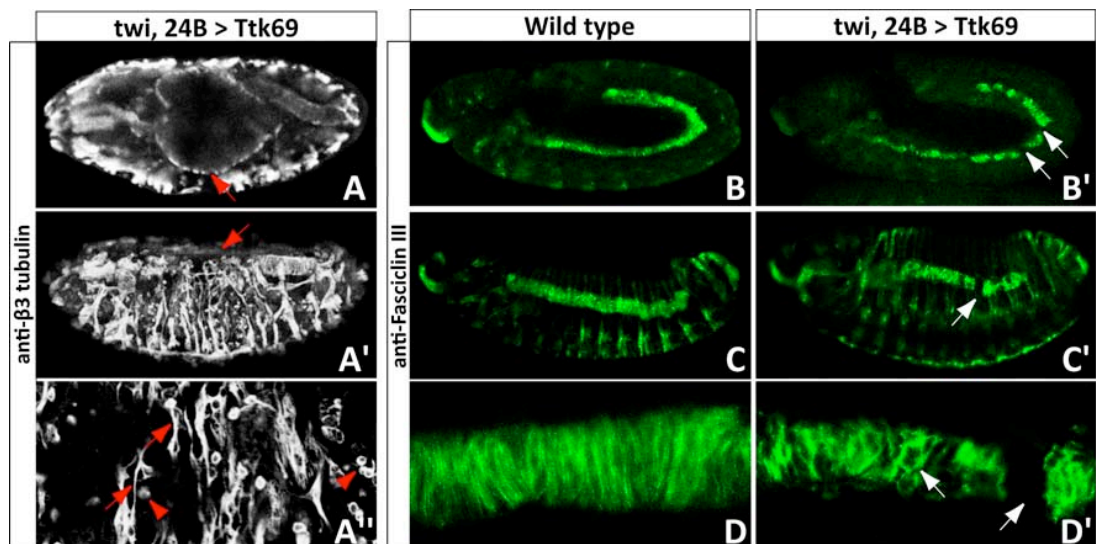
Figure 4.2-6: Panmesodermal Ttk69 interferes with specification of all three muscle types
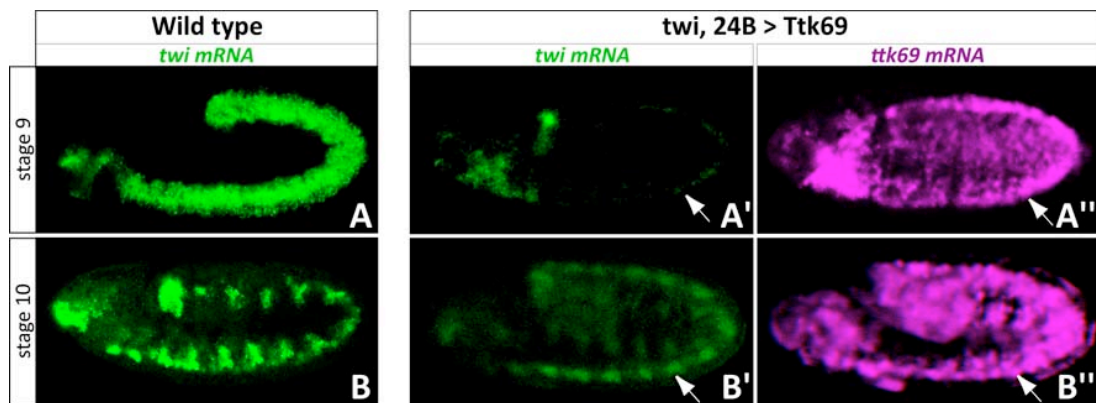
Ttk69 expression was driven with a panmesodermal driver twi, 24B - Gal4 [81, 223].

(A)   Stage 16 embryos were immunostained with anti-β3 tubulin antibody. Almost all embryos with ectopic Ttk69 have abnormal visceral musculature with many gaps (arrow in A), lack heart (A'), and form only a few thin fibres without a clear identity (arrows in A") while many myoblasts remain mononucleated (arrowheads in A").

(B-D) Antibody against Fasciclin III was used to reveal gaps in visceral mesoderm in stage 11 (B') and stage 13 (C', D') embryos with ectopic, panmesodermally induced Ttk69 when compared to the wild type embryos (B, C, D).

The severity and extent of the somatic musculature phenotype can be attributed to a reduction in *twist* (*twi*) levels compared to wild type embryos (Figure 4.2-7). In a wild type situation, *twi* is strongly expressed throughout the mesoderm at stage 9 (Figure 4.2-7 A) and during stage 10 its levels are modulated into high and

low expression domains (Figure 4.2-7 B) through the action of Notch signalling [120]. When Ttk69 is ectopically expressed from stage 8, the levels of *twi* mRNA at stage 9 decline dramatically (Figure 4.2-7 A'). Even though *twi* expression remains reduced, at stage 10 its modulation into the different domains is preserved (Figure 4.2-7 B').

High levels of Twist protein have been demonstrated to be fundamental for the development of somatic muscle [81]. However, it has been also reported that if Twist activity is genetically reduced after gastrulation using a temperature-sensitive allelic combination, visceral mesoderm and heart develop normally [81]. Therefore the defects in specification of these mesoderm derivatives upon expression of ectopic Ttk69 are probably due to the misexpression of other factors specific to the developmental programs of visceral and cardiac mesoderm.

Figure 4.2-7: Ectopically expressed Ttk69 represses *twist* expression

Ttk69 expression was driven with a panmesodermal driver twi, 24B - Gal4 [81, 223].

(A) In wild type embryos, *twi* is strongly expressed in all mesodermal cells at stage 9 (A). If panmesodermal expression of Ttk69 is induced (arrow in A''), *twi* levels are reduced significantly (arrow in A').

(B) At late stage 10, expression of *twist* is segmented by Notch signalling into high and low Twist domains [120]. Ectopic expression of Ttk69 (arrow in B'') results in an overall decreased level of *twi* mRNA expression but its modulation seems to still occur (arrow in B'). Note that the microscope settings in (B') were changed compared to (A') to demonstrate the patterned expression.

4.2.2 Deciphering the molecular function of Ttk69

4.2.2.1 ChIP-on-chip assay identifies direct targets of Tramtrack69

To understand the dramatic muscle phenotypes of *ttk69* loss-of-function and gain-of-function mutants, I performed a ChIP-on-chip experiment to identify its direct targets. For this purpose, two different polyclonal antibodies specific against the Ttk69 isoform [194, 228] were used to immunoprecipitate chromatin extracted from 6-8 hour old wild type embryos. As a reference to account for non-specific binding, another aliquot of the same chromatin was immunoprecipitated with rabbit pre-immune serum. Purified, amplified, and labelled DNA fragments were hybridized to high-resolution Affymetrix GeneChip® Drosophila Tiling 1.0R Arrays covering the whole genome. The resultant data were processed as described previously [196] and as significant binding were considered regions above the TileMap posterior score [229] of 5.5, corresponding to a false discovery rate (FDR) estimate of approximately 2 %. *Cis*-regulatory modules (CRMs) were defined within 100 bp from each direction of a detected peak. By these criteria, there are 2,084 CRMs bound by Ttk69, the relatively high number of which possibly reflects its diverse developmental roles and widespread expression. There is a significant overlap between Ttk69-bound and mesodermal CRMs [196], suggesting that a large portion of Ttk69 binding is in mesodermal cells and that this may contribute to the mesodermal regulatory circuitry.

4.2.2.2 Validation of Ttk69 ChIP-on-chip data

In order to verify the acquired dataset of global Ttk69 occupancy, I utilized previously experimentally validated Ttk69-bound regions as well as genes known to be downstream of Ttk69 in the embryonic nervous system (see Appendix section 7.3). The identified CRMs cover the previously mapped binding sites of Ttk69 in the promoter regions of *tailless* [153], *even-skipped* [181], and *string* [177] (Appendix 7.3 A-C). Previous experiments demonstrated that two Ttk69 sites in the promoter region of *eve* are strongly bound when a nuclear extract from 5-10 hour old embryos is used as a source of Ttk69 protein [230]. Consistent with this, both sites are covered by a single CRM identified in the ChIP-on-chip experiment (Appendix 7.3 B).

A number of neuronal genes and cell-cycle regulator *CycE* are derepressed in *ttk* mutant embryos (see Table 1.4-2), but whether these genes are directly repressed

by Ttk69 was unknown. Based on our ChIP-on-chip data, Ttk69 occupies multiple regions in the vicinity of *dpn*, *sc*, *scrt,* and *CycE* (Appendix 7.3 D-G), suggesting that it regulates these genes directly.

4.2.2.3 ChIP-on-chip based validation of computationally predicted Ttk69 binding

Independent of my genetic analysis of *ttk*, a computational biologist in our group, B. Wilczynski, had also indicated a role for Ttk69 within the regulatory network driving mesoderm development. The core mesodermal TFs Twi, Tin, and Mef2 co-bind to a *ttk* CRM [108] at 2-4 and 4-6 hours of embryogenesis. Using a position weight matrix of Ttk69, Bartek discovered a significant enrichment of putative Ttk69 binding motifs in additional Twi-Tin-Mef2 co-bound CRMs, compared to general mesodermal CRMs [196]. Since Ttk69 itself is targeted by these TFs, its regulatory input into the Twi-Tin-Mef2 combinatorially bound CRMs would implicate its involvement in extensive feed-forward regulation (Figure 4.2-8 A).

To test the model, the ChIP-on-chip dataset was used as an accurate source of Ttk69 *in vivo* binding events. The binding of Ttk69 within all mesodermal CRMs is significantly enriched compared to randomized regions, reflecting the high overlap between Ttk69-bound and mesodermal CRMs (Figure 4.2-8 B). Even more striking, however, is the enrichment of Ttk69 binding within the Twi-Tin-Mef2 class of CRMs; over 36% of these CRMs are bound by Ttk69 as compared to 13% of all mesodermal CRMs (Figure 4.2-8 B). The Ttk69 occupancy data provides global validation for the predicted model and implicates an involvement of Ttk69, being a repressor, in incoherent type of feed-forward regulatory loops (reviewed in [57]). As the Twi-Tin-Mef2 class of CRMs is bound by the mesodermal TFs earlier in development, one possible function of Ttk69 binding could be a need to terminate transcription of genes which are expressed in presumptive mesoderm and do not contribute to its later subdivision and differentiation.
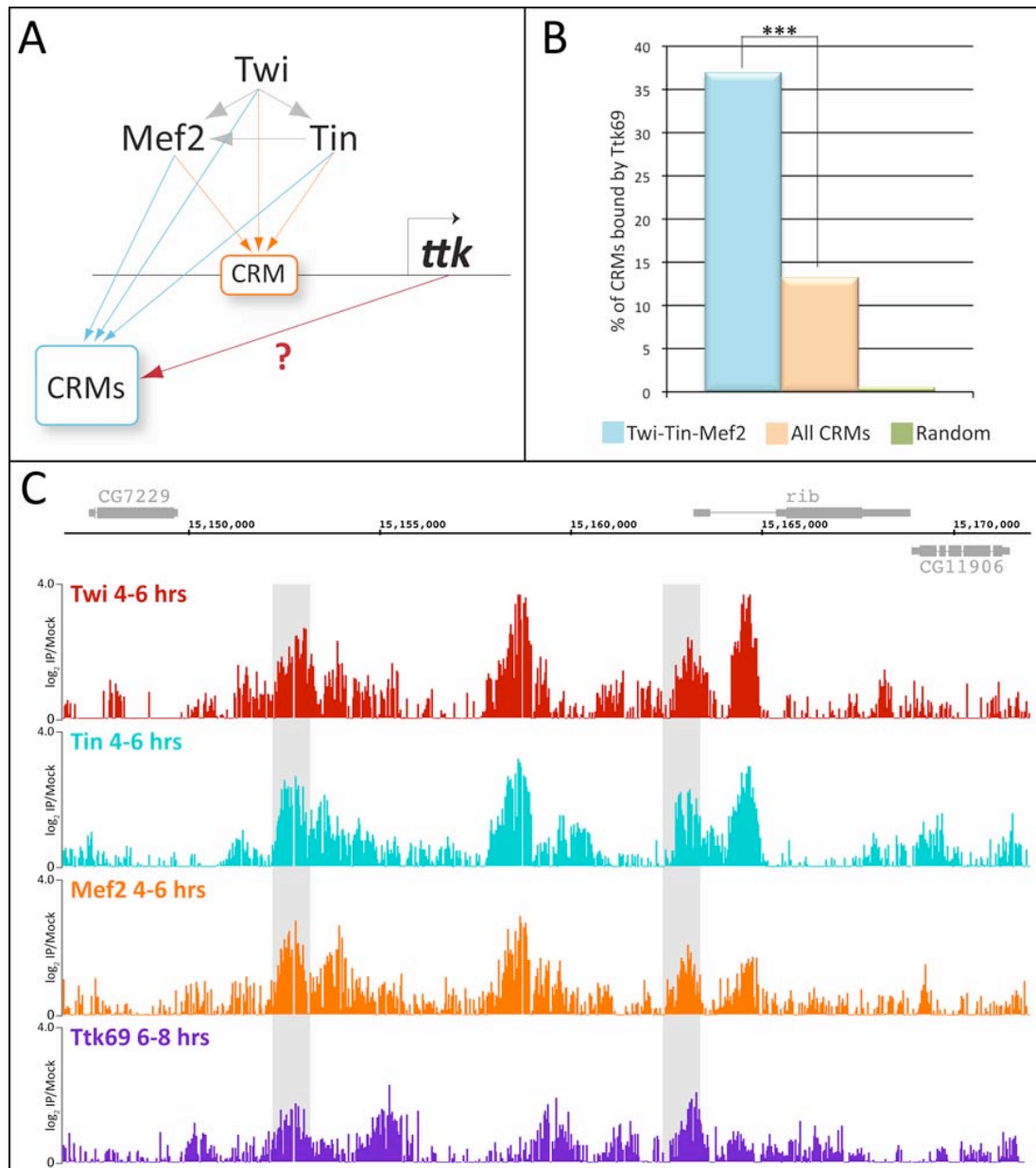
Figure 4.2-8: Ttk69 binds to CRMs that are co-bound by Twi, Tin, and Mef2 during mesoderm development

(A) A predicted model in which Ttk69 occupies CRMs that are co-bound at earlier time point by Twi, Tin, and Mef2. As Ttk69 is a repressor that is itself targeted by these TFs, it would imply incoherent feed-forward regulation. Arrows pointing to CRMs indicate regulatory input rather than transcriptional activation. The question mark indicates the tested hypothesis.

(B) Based on ChIP-on-chip data, Ttk69 binding is significantly enriched in CRMs co-bound by Twi, Tin, and Mef2 (blue) compared to all mesodermal CRMs (orange) or random genomic sequences (green). *** $p < 0.01$. Analysis was performed by B. Wilczynski.

(C) ChIP-on-chip profiles of Twi (red), Tin (blue), and Mef2 (orange) at 4-6 hours, and Ttk69 at 6-8 hours (purple). In grey are highlighted regions that are bound at all conditions.

4.2.2.4 *In vivo* activity of Ttk69-bound CRMs

In order to assay the regulatory input of Ttk69 into the CRMs identified in the ChIP-on-chip experiment, I analyzed the *in vivo* activity of three selected CRMs. Transgenic reporter assays were performed to assess the activity of wild type CRMs as well as CRMs where several or all putative Ttk69 binding sites were mutated. All three CRMs are co-bound by Twi, Tin, and Mef2 at 4-6 hours after egg laying (AEL) and therefore fall within the set of CRMs predicted to be targeted by the Ttk69 incoherent feed-forward loop (see Section 4.2.2.3). In addition, they are associated with genes encoding TFs with developmental functions related to Ttk69 [231, 232].

The tested CRMs were subcloned into a *P*-element vector upstream of a *lacZ* reporter with a *Hsp70* minimal promoter [195]. For each CRM, at least three different transgenic lines were generated to account for spurious activity due to integration events as a result of *P*-element transposition.

Two selected Ttk69 CRMs fall into the intronic region of the gene *jumeau* (*jumu*) (Figure 4.2-9 A), coding for a TF implicated in asymmetric cell division in the nervous system [232]. Activity of the CRM *jumuA* initiates at early stage 11 when it is expressed strongly in tracheal placodes and weakly in the visceral mesoderm primordium (Figure 4.2-9 C'). At stage 11, strong *lacZ* signal comes from the cells of caudal visceral mesoderm that will eventually form the longitudinal visceral muscle (Figure 4.2-9 C"). Expression in the trachea as well as visceral mesoderm persists until the very end of embryogenesis (not shown). There are 10 putative Ttk69 binding sites in the CRM, out of which 3 sites were mutated (see Appendix section 7.4 for the sequences). Interestingly, the mutated version of the enhancer exhibits *lacZ* derepression at stage 5 (Figure 4.2-9 D). While the wild type enhancer is inactive at this stage (Figure 4.2-9 C), *jumuA-Ttkmut* shows expression in transverse ectodermal stripes (Figure 4.2-9 D). In addition, the expression in tracheal placodes is reduced and in caudal visceral mesoderm it is completely lost (Figure 4.2-9 D"). However, the weak expression in the trunk visceral mesoderm is unaffected (Figure 4.2-9 D').

The second intronic *jumu* CRM, *jumuB*, shows ubiquitous activity during early stages (Figure 4.2-9 E). At stage 11, it drives a complex *lacZ* expression in the tracheal primordia, discrete cells of the central nervous system, and caudal visceral mesoderm (Figure 4.2-9 E', E"). When 9 out of the 15 Ttk69 motifs were mutated, the

early and tracheal placode CRM activities are unaffected (Figure 4.2-9 F, F'), but the number of *lacZ*-positive CNS cells is decreased and the signal from caudal visceral mesoderm is completely lost (Figure 4.2-9 F'').

Finally, the CRM upstream of the *rib* locus (Figure 4.2-9 B) drives weak expression in the mesodermal cells at stage 5 and 6 (Figure 4.2-9 G). During stage 9 and 10, the *lacZ* reporter shows metameric expression in dorsal ectoderm (Figure 4.2-9 G') and at stage 11 it can be detected in the caudal visceral mesoderm (Figure 4.2-9 G'') and a subset of neuroectodermal cells. This CRM contains three putative Ttk69 binding sites and mutations in all of them affect only the caudal visceral mesoderm activity (Figure 4.2-9 H-H'').

Overall, the activity of the wild type as well as mutant CRMs provides unexpected insights into the regulatory potential of Ttk69 (discussed in Section 5.3.3.1). First, the CRMs drive complex and dynamic expression in tissues that mirror the expression of *ttk69* itself. Second, mutagenesis of putative Ttk69 binding motifs results in a loss of expression in several different tissues. Ttk69 binding seems to have a particular importance for activity in the caudal visceral mesoderm of all wild type CRMs as the expression is lost upon mutation of Ttk69 motifs. Finally, one mutated CRM, *jumuA-Ttkmut*, shows derepression at blastoderm stage, which is most likely mediated by maternally provided Ttk69 protein.
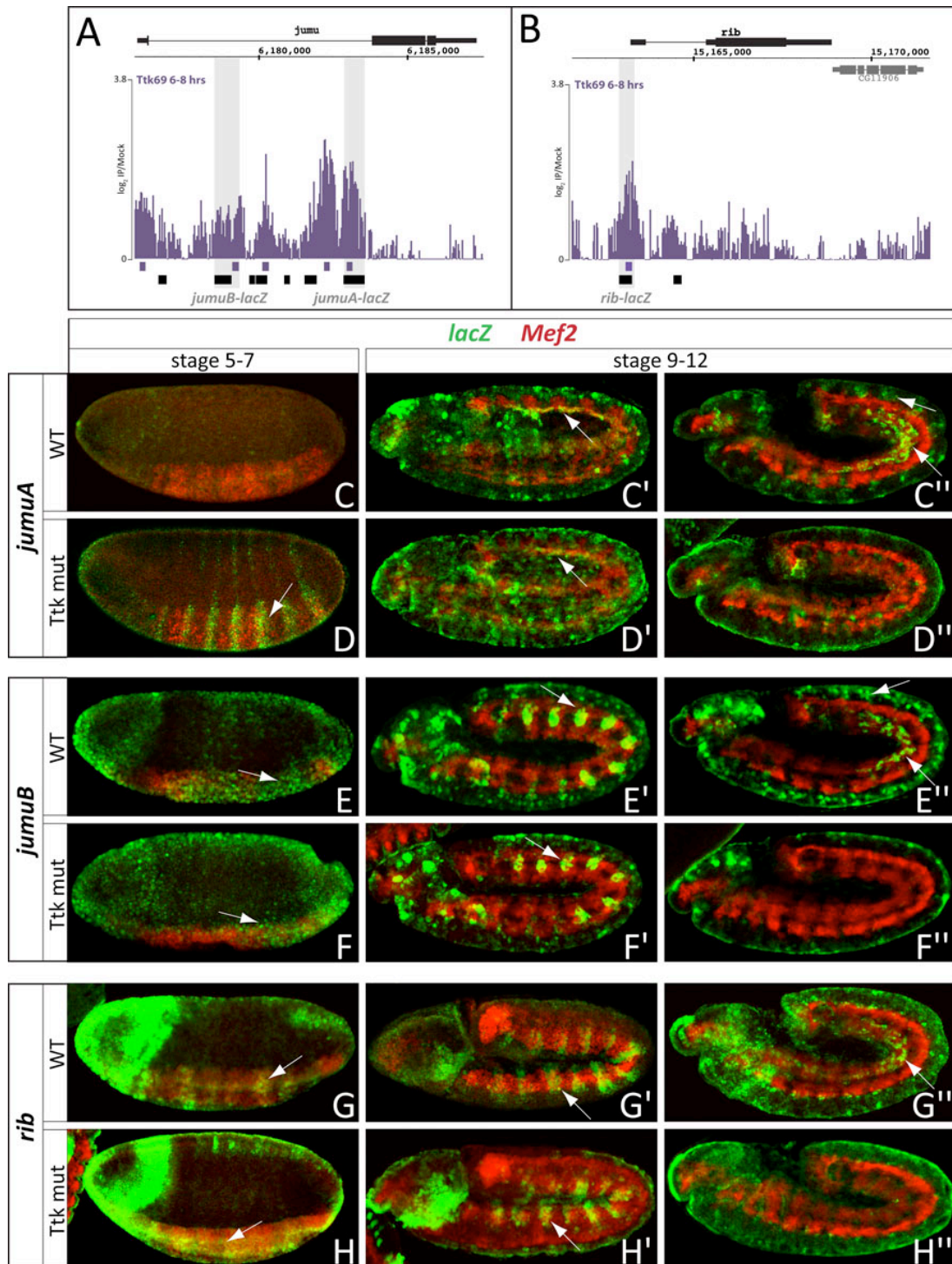
Figure 4.2-9: *In vivo* **activity of Ttk69-bound CRMs**

Activity of all CRMs was assayed by fluorescent *in situ* hybridization with a probe against *lacZ* (green) and *Mef2* (red), labelling the mesoderm.

(A-B) Loci of *jumu* (A) and *rib* (B) with Ttk69 binding profile in purple and mesodermal CRMs [196] in black. The sequences assayed in the transgenic reporter assays are highlighted in grey.

(C-D) The wild type *jumuA* CRM is not active at stage 5 (C) but is weakly expressed in the trunk visceral mesoderm (C') and strongly in trachea and caudal visceral mesoderm (C'') at stage 11 (arrows). When several Ttk69 sites are mutated, striped expression at stage 5 appears (arrow in D) while the activity in trachea is reduced (D'') and in the

caudal visceral mesoderm completely lost (D"). In the trunk visceral mesoderm it remains unaffected (arrow in D').

(E-F) The CRM *jumuB* drives *lacZ* expression ubiquitously at stage 5 and in the tracheal placodes (E'), the CNS and caudal visceral mesoderm (E") at stage 11. Its mutated version shows lack of the caudal visceral mesoderm expression and reduction in the CNS expression (F") while the other aspects seem unaffected (arrows in F, F').

(G-H) The CRM *rib* is expressed weakly in the mesoderm (arrows) and strongly in the anterior part of stage 5 embryos (G) and it shows a metameric distribution in dorsal ectoderm at stage 9 (arrows in G'). At stage 11 it is weakly expressed in the neuroectoderm and caudal visceral mesoderm (arrows in G"). Mutagenesis of the three putative Ttk69 sites does not affect its activity in the presumptive mesoderm (arrow in H) or dorsal ectoderm (arrow in H'), but the activity in the caudal visceral mesoderm is lost (H").

4.2.3 An integrated network of Lame duck and Tramtrack69 activity is required for FCM specification

One explanation for the observed *ttk69* loss- and gain-of-function phenotypes is its requirement for FCM specification, where it would repress FC-specific genes. There is currently only one other TF known to be essential for FCM specification; Lame duck (Lmd), a zinc finger TF that has been reported to act mainly as a transcriptional activator [133]. The only known direct target of Lmd is the differentiation factor *Mef2* [131] and so the knowledge about the regulatory subnetwork downstream of Lmd is extremely limited. To understand if and how Ttk69 contributes to FCM specification, it is essential to discern if Ttk69 has an overlapping or distinct function to Lmd. Consequently, to provide insights into the regulatory program driven by Lmd and its possible connection to Ttk69, I performed a ChIP-on-chip experiment against Lmd at the same stages that were used for Ttk69.

4.2.3.1 Genome-wide occupancy of Lame duck and overlap with Tramtrack69

ChIP-on-chip experiments with Lmd were performed under the same time points and conditions as Ttk69 (see Section 4.2.2.1). Two independent Lmd antibodies generated in rabbit were used to immunoprecipitate Lmd-bound genomic regions. The samples were hybridized to Affymetrix GeneChip® *Drosophila* Tiling

1.0R Arrays and data was processed as for Ttk69 and other mesodermal TFs [196]. A TileMap threshold score [229] was set to 5.5, equal to Ttk69, and corresponding to an FDR estimate of approximately 8%. Above this threshold there are 1,024 significantly enriched CRMs. Expression of Lmd is restricted only to FCMs and is absent from non-mesodermal tissues. Therefore the lower number of bound regions compared to Ttk69 (2,084) was expected.

To date, there is only one known direct target of Lmd, *Mef2*. The regulatory region bound by Lmd has been experimentally minimized to 20 base pairs [131]. The generated ChIP-on-chip dataset confirmed binding of Lmd to this sequence at the correct stages of development, demonstrating the sensitivity of the data (Figure 4.2-10 A). There is a substantial overlap between the Lmd-bound regions and mesodermal CRMs, as expected.

Importantly, there is also a high overlap between Ttk69 and Lmd-bound regions, which is highly significant ($p < 10^{-237}$) compared to randomly shuffled non-coding regions (Figure 4.2-10 B). There are altogether 191 CRMs that are co-occupied by Ttk69 and Lmd, and almost all of them (166) are also bound by at least one other mesodermal TF [196]. In order to obtain insights into the functional role of this binding, the co-bound CRMs were assigned to the nearest gene and available databases and literature were surveyed for the embryonic expression patterns of these genes (Figure 4.2-10 C). More than a half of the co-regulated genes lack information on their expression or are not expressed in spatially and temporally controlled manner. However, 64 genes show specific expression patterns during embryogenesis, the majority of which (46 genes) are specifically expressed in mesoderm or its derivatives. It therefore appears that Ttk69 and Lmd have a shared role in regulating a subset of mesodermal genes. However, both TFs also provide important regulatory input independently of each other.
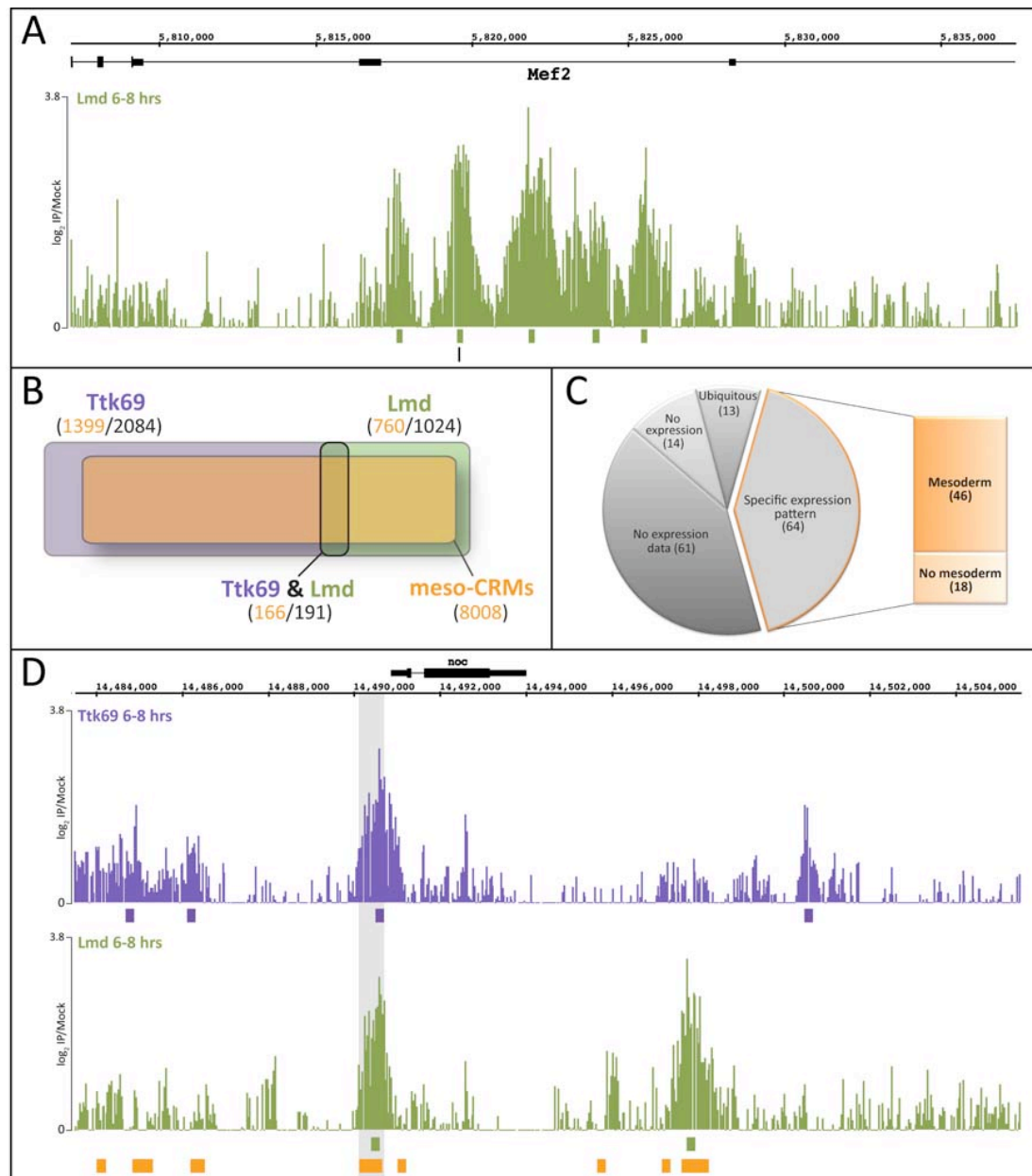
Figure 4.2-10: Overlap between Lmd and Ttk69 binding profiles

(A) ChIP-on-chip data successfully recovered the only experimentally validated binding site of Lame duck (in black). There are also additional Lmd-bound CRMs in the *Mef2* locus identified in the experiment (green bars below the signal).

(B) There is a significant overlap ($p < 10^{-10}$) between CRMs bound by Ttk69 (purple), Lmd (green), and mesodermal CRMs (orange). In brackets is a total number of CRMs in each class (black) and a number of CRMs that overlap with mesodermal CRMs (orange).

(C) Expression patterns (based on literature and BDGP *in situ* database [233]) of Ttk69 and Lmd co-regulated genes. In brackets are the numbers of unique genes in each group.

(D) Example of a mesodermal gene (*no ocelli, noc*) with a CRM co-bound by Ttk69 (purple), Lmd (green), and at least one another mesodermal TF (orange) highlighted in grey.

4.2.3.2 Preference of Ttk69 binding to CRMs of founder cell-specific genes

The genetic analyses suggested a model by which Ttk69 represses genes normally restricted to the FCs within FCMs. To assess if this is true at a molecular level, I used the genome-wide occupancy data in combination with a collection of genes systematically classified as expressed only in FCs (FC-specific), FCMs (FCM-specific), or both cell types [234]. Regions 1,500 bp upstream and downstream of the transcriptional start sites (TSSs) of these genes were scanned for Ttk69-bound CRMs. If such CRMs were identified, they would be further classified depending on whether they overlap with a mesodermal CRM or not. Consequently, each myoblast fusion gene falls into one of three possible categories: it either has no Ttk69 CRMs in the 3 kb region around its TSS, or it has at least one such CRM overlapping a mesodermal CRM, or it has one or more Ttk69-bound CRMs that do not overlap mesodermal CRMs. Classifying the CRMs based on their overlap with mesodermal CRMs was used to help restrict the analysis to binding events within mesodermal tissues.

In agreement with the phenotypic results, there is indeed a significant preference of Ttk69 binding to CRMs associated with FC-specific genes compared to genes expressed only in FCMs or in both FCs and FCMs (Figure 4.2-11 A). As most (90%) of these bound regions are in addition occupied by a mesodermal TF, it is likely that they have a potential to drive mesodermal expression.

When the same analysis was performed using the Lmd ChIP-on-chip dataset, the opposite trend was apparent. Binding of Lmd is clearly depleted at regulatory elements associated with FC-specific genes and enriched at the CRMs of genes expressed in the FCMs (Figure 4.2-11 B), consistent with its ability to activate transcription in FCMs. Ttk69 and Lmd are thus both involved in the specification of FCMs, but while Lmd does so by activating a FCM-specific differentiation program, Ttk69 represses genes that could interfere with FCM fate. In addition, both TFs co-occupy a considerable number of mesodermal CRMs (see Section 4.2.3.1), however the biological role of this co-occupancy remains to be determined.
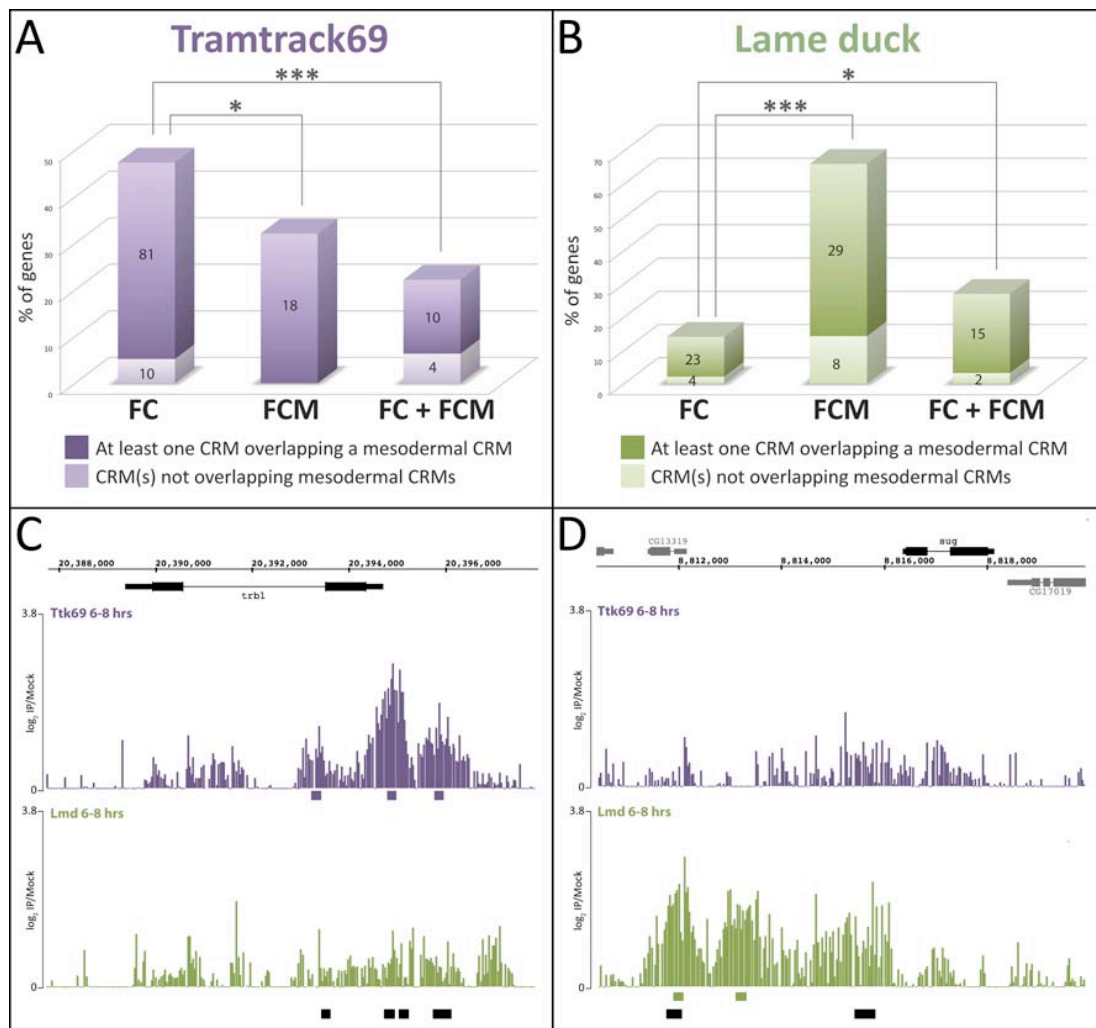
Figure 4.2-11: Occupancy of Ttk69 and Lame duck in the vicinity of genes essential for myoblast fusion

(A-B) Ttk69 binding is significantly enriched in the vicinity of genes expressed in FCs compared to genes expressed in FCMs or in both cell types (A). Lmd binding is, on the other hand, enriched in the vicinity of FCM-specific genes (B). Dark colour indicates that at least one assigned CRM overlaps a mesodermal CRM [196]. CRMs within 1.5 kb distance upstream and downstream of the TSS were considered. Numbers refer to the number of genes in each class. * $p < 0.05$, *** $p < 0.01$.

(C-D) Example of an FC-specific gene *tribbles* (*trbl*, C), an FCM-specific gene *sugarbabe* (*sug*, D), and corresponding ChIP-on-chip signal for Ttk69 (purple) and Lmd (green). In the track below the signal are the computed CRMs in the respective colour and in black are mesodermal CRMs bound by the five key myogenic TFs [196].

Note that there is only background signal for Lmd in the *trbl* locus and similarly background signal for Ttk69 in the *sug* locus.

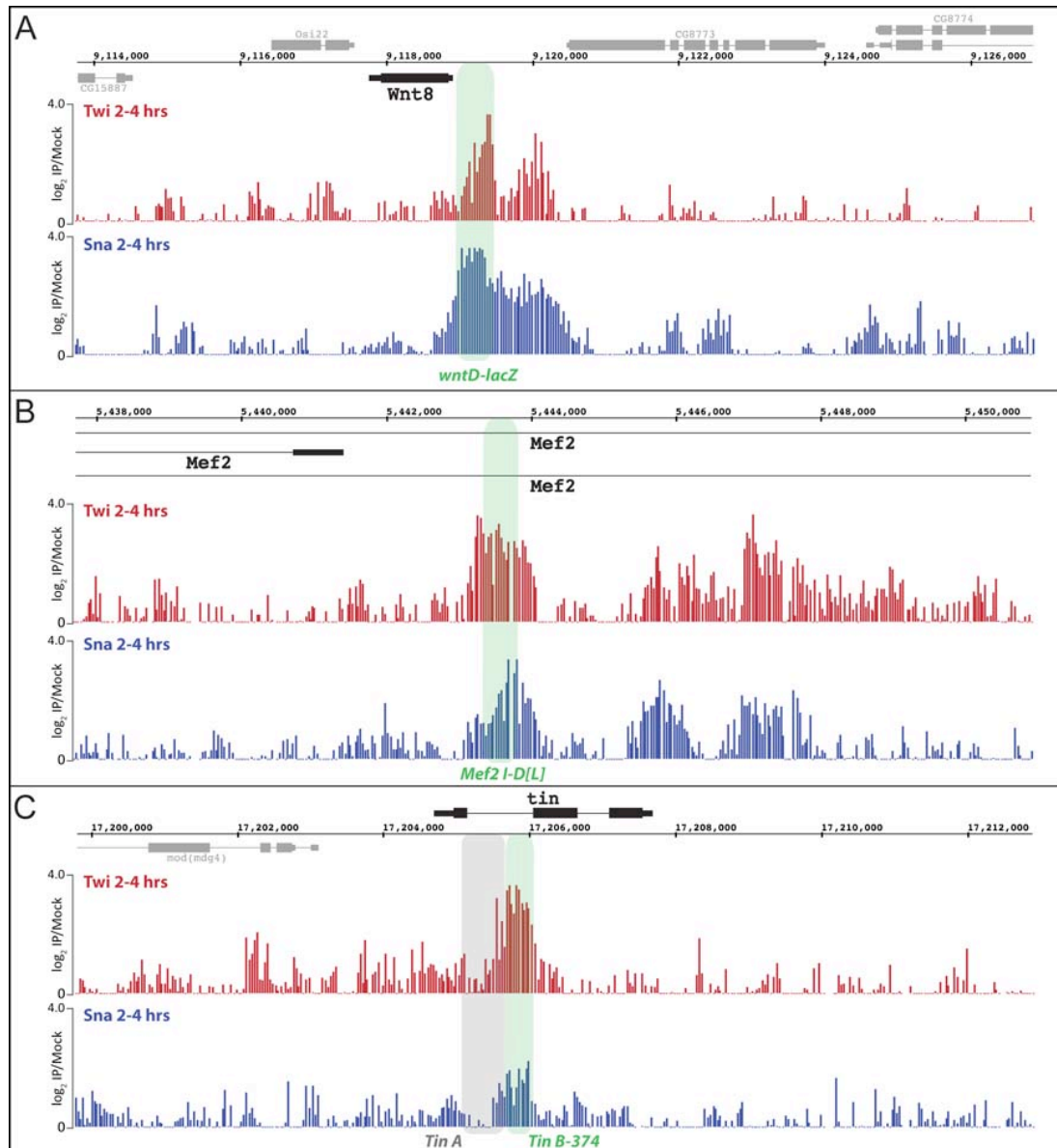# 4.3 Snail as a transcriptional activator in the presumptive mesoderm

Twist and Snail are two major regulators co-expressed in the presumptive mesoderm of blastoderm stage embryos. While Twist is a master regulator of mesodermal gene expression, Snail has been traditionally viewed as a repressor of neuroectodermal genes (see Section 1.2.2). It therefore came as a surprise to find that Snail binds to CRMs which are active in mesoderm at the stages of Snail occupancy (Figure 4.3-1). As multiple additional observations could be explained by Snail positively regulating gene expression (see Section 1.2.2.1), I directly addressed this possibility first *in vitro* and then *in vivo* using reporter assays.

## 4.3.1 Snail binds to enhancer regions of early mesodermal genes

As has been observed previously, there is a significant overlap in the binding profiles of the mesodermal activator Twist and the repressor Snail [56]. These observations were derived from ChIP-on-chip experiments in *Toll10B* embryos, in which, through a constitutively active Toll receptor, high nuclear levels of Dorsal protein activate the master mesoderm regulator Twist throughout the embryo [235]. As a result, the entire embryo is composed of mesoderm at the expense of neurogenic and dorsal ectoderm.

To confirm that these findings hold true in a wild type genetic background, Twist and Snail ChIP-on-chip experiments were performed in wild type 2-4 hours old embryos (by Robert Zinzen and Martina Rembold). This wild type occupancy data confirmed Twist and Snail co-binding to previously identified enhancers of genes repressed by Snail, such as *single-minded* (*sim*), rhomboid (*rho*), or the *wnt inhibitor of Dorsal* (*wntD*) (Figure 4.3-1 A). In addition, however, their binding was also detected near genes encoding key mesodermal regulators, such as *Mef2* (Figure 4.3-1 B) and *Tin* (Figure 4.3-1 C). Enhancers responsible for early mesodermal expression of both genes, *Mef2 I-D[L]* [236] and *Tin B-374* [60], respectively, are co-bound by Twist and Snail at 2-4 hours after egg laying (AEL) (Figure 4.3-1 B, C). An example of yet another mesodermal enhancer bound by both TFs is the *Mdr49 early mesoderm* enhancer (Figure 4.3-1 D) which directs expression during cellularization and early

gastrulation stages [56]. Two mesodermal genes whose expression is reduced in *sna* mutant embryos are *htl*, which encodes an FGF-receptor required for mesoderm migration [237], and *zfh1*, a transcriptional repressor important for later development of particular mesoderm derivatives [238]. Both Snail and Twist bind to intronic as well as upstream regions within the loci of both genes (Figure 4.3-1 E, F).
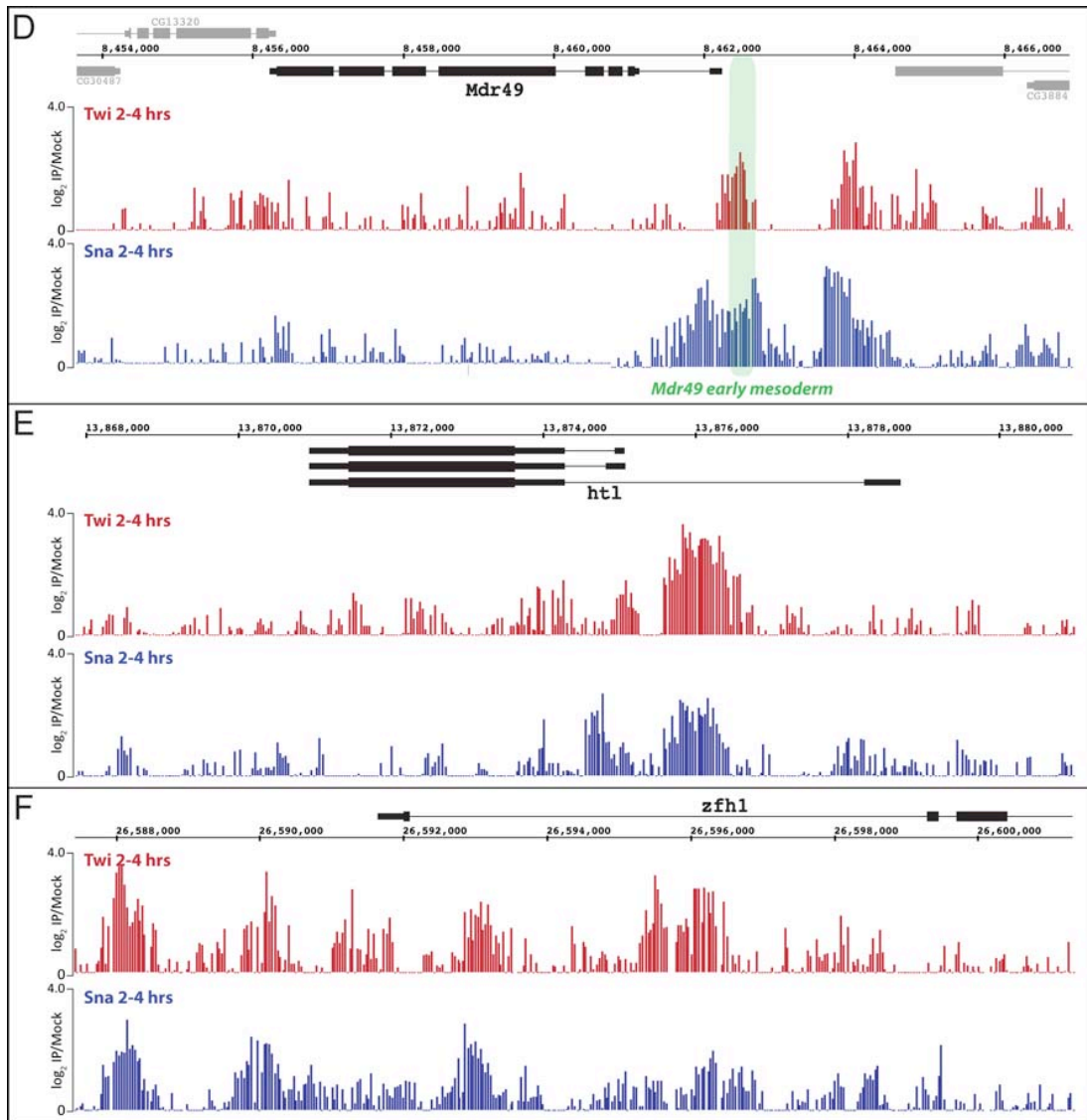
Figure 4.3-1: Twist and Snail co-occupy multiple regions in the vicinity of mesodermal genes

Results of a ChIP-on-chip experiment with 2-4 hours old wild type embryos and Twist or Snail antibodies. Both TFs co-bind to numerous regions in the vicinity of *wntD* (A), *Mef2* (B), *tin* (C), *Mdr49* (D), *zfh1* (E), and *htl* (F) genes (in black). Known enhancer regions that were selected for further analysis are shown in green. In grey are other relevant enhancer regions. Data kindly supplied by Robert Zinzen and Martina Rembold.
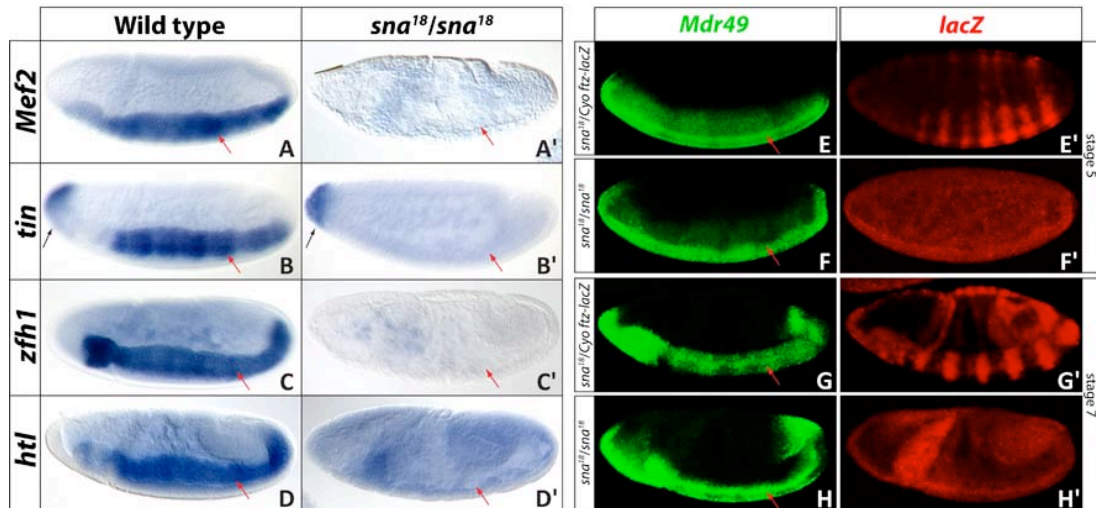
4.3.2 Endogenous expression of several mesodermal genes is dependent on Snail

Earlier studies noted that the mRNA transcripts of some early mesodermal genes are significantly reduced or completely absent in *sna*-deficient embryos [86-88, 90]. To confirm these findings, I performed *in situ* hybridizations in embryos carrying the *sna* loss-of-function null allele *sna[18]* [239, 240] and compared mesodermal gene expression to wild type embryos. Given the failure of gastrulation followed by additional morphological defects, expression of mesodermal genes in *sna* mutant embryos can only be judged during early embryonic stages. To facilitate the identification of mutant embryos, the balancer chromosome was marked with *ftz-lacZ* so that the absence of *lacZ* expression in ectodermal stripes indicated homozygous mutant embryos.

The earliest time point when *Mef2* mRNA is observed in wild type embryos is stage 6. However, in homozygous *sna[18]* embryos, no *Mef2* expression can be detected at any developmental stage (Figure 4.3-2 A, A'). Similarly, the expression of *tin* in the trunk mesoderm of *sna* mutant embryos is completely absent at stage 7 (Figure 4.3-2 B') and cannot be observed at any other time point. An interesting aspect of *tin* expression in *sna[18]* homozygous embryos is the normal expression of *tin* in the anterior head mesoderm (Figure 4.3-2 B, B'). Given that the complex expression pattern of *tin* is governed by independent modular CRMs [60] and that *snail* is not expressed in these cells, it is likely that Snail exerts its regulatory effect only on the CRM(s) responsible for *tin* expression in the trunk mesoderm.

Based on the TF occupancy data, *zfh1* and *htl* are additional mesodermal target genes that may be positively regulated by Snail. While both genes are highly expressed at stage 7 in wild type embryos (Figure 4.3-2 C, D), in *sna[18]* homozygous embryos their mRNA expression is absent (Figure 4.3-2 C', D').

Another candidate gene for activation by Snail is *Mdr49* whose expression commences with the onset of cellularization. However, neither the initiation nor the maintenance of *Mdr49* mRNA expression seem to be Snail-dependent as no apparent changes in its pattern or levels could be observed at stage 5 (Figure 4.3-2 E-F') or at stage 7 (Figure 4.3-2 G-H') in *sna[18]* homozygous embryos.

Figure 4.3-2: Expression of mesodermal genes in *snail* loss-of-function embryos

Endogenous expression of genes expressed in the presumptive mesoderm was detected by colorimetric or fluorescent *in situ* hybridization in stage 5 or 7 wild type and *sna*[18] homozygous embryos. The *sna* mutant embryos also carry a *ftz-lacZ* marked balancer chromosome. The absence of *lacZ* expression in the ectoderm was used to identify homozygous mutants.

(A-A') Expression of *Mef2* is completely lost in *sna* mutants compared to wild type.

(B-B') *tin* is absent in the trunk mesoderm of *sna*[18] homozygous embryos but it persists in the anterior tip of the head (arrow).

(C-D') In *sna* mutant embryos, levels of *zfh1* (C) and *htl* (D) mRNA are significantly reduced.

(E-H') Fluorescent double *in situ* hybridization in embryos generated by *sna*[18]/*CyO ftz-lacZ* flies. Probe against *lacZ* (red) facilitates discrimination between homozygous *sna* mutant embryos (F', H') and embryos carrying only one or none *sna*[18] allele (E', G'). There are no apparent changes in *Mdr49* expression (green) at stage 5 (E-F') and 7 (G-H') in *sna*[18] homozygous embryos (F, H).

In summary, there are several examples where Twist and Snail co-occupy active mesodermal enhancers and many of the co-bound enhancers are associated with genes whose expression is dependent on the presence of Snail. This observation contradicts the traditionally assumed role of Snail as a dedicated repressor. In an attempt to resolve this discrepancy, I tested the possibility that Snail could also positively regulate transcription.

4.3.3 Snail enhances the activation of mesodermal enhancers *in vitro*

4.3.3.1 Luciferase reporter assays in *Drosophila* Kc cells

As a first step in assessing whether the binding of Snail to mesodermal enhancers is functional, I performed luciferase reporter assays with three selected enhancers in *Drosophila* Kc cells. This cell line is of late embryonic origin [241] and does not express detectable levels of endogenous Twist, Snail, and Dorsal mRNA or protein ([242, 243], data not shown). In each experiment, the cells were transfected with a uniform amount of a firefly luciferase reporter vector with the luciferase gene linked to an upstream minimal *Hsp70* promoter and the enhancer to be assayed. In parallel, *Renilla* luciferase driven only by the *Hsp70* minimal promoter was co-transfected to account for variability in transfection efficiency. In each assay, increasing amounts of one or two different protein expression vectors were co-transfected and their quantity was equalized with an empty expression vector. To eliminate background effects, the detected changes in firefly luciferase activity were adjusted to the levels in cells transfected with the empty expression vector alone. Each condition was measured in triplicate per experiment and the final values of mesodermal enhancers activity were derived from three independent experiments.

I assayed the activity of the *wntD-lacZ* enhancer that is co-bound by Twist, Snail, and Dorsal at 2-4 hours AEL ([56], Figure 4.3-1 A) as a control for the experimental setup. Expression of the endogenous *wntD* gene is regulated positively by Twist and Dorsal and negatively by Snail [244]. Kc cells were co-transfected with the *wntD* construct and with varying amounts of Dorsal and/or Snail expression vectors. In agreement with *in vivo* observations, transfection of 100 ng of plasmid encoding Dorsal induces the activity of the enhancer approximately 80-fold when compared to the empty expression vector. When an equal amount of Dorsal is co-transfected with 100 ng of Snail, the relative expression levels of luciferase are reduced to approximately 50-fold (Figure 4.3-3 A). The results obtained from the cell culture system are therefore consistent with the observations made in embryos, indicating that, at least to a certain extent, the cell-based luciferase assays reflect the *in vivo* situation.

4.3.3.2 Snail acts to synergistically enhance *Mdr49 early mesoderm* and *Tin B-374* enhancer activity

The *Mdr49 early mesoderm* CRM recapitulates the early mesodermal expression of the endogenous *Mdr49* gene [56]. Similar to the *wntD-lacZ* CRM, it is also co-bound by Twist, Snail, and Dorsal [56]. However, the strong expression of the *Mdr49 early mesoderm* CRM in the early mesoderm suggests a positive input from Snail. Indeed, while Dorsal can activate the enhancer weakly, addition of Snail dramatically increases the average fold of luciferase activity (Figure 4.3-3 B). 100 ng of transfected Dorsal cDNA induces an average luciferase expression 4.6-fold, while addition of the same amount of Snail together with Dorsal raises the levels to 16-fold (Figure 4.3-3 B). Co-transfection of Snail therefore induces almost a four fold increase in Dorsal-mediated enhancer activation.

The expression of the endogenous *tin* mRNA is controlled by multiple modular enhancer elements with distinct spatio-temporal activities [60]. During early embryonic stages, *tin* expression is driven by two adjacent enhancers, *Tin B-374* in the presumptive trunk mesoderm and *Tin A* in the anterior mesoderm of the head [60]. While Twist and Snail co-bind to the *Tin B-374* enhancer at 2-4 hours AEL, they do not seem to occupy the *Tin A* enhancer (Figure 4.3-1 C). This is consistent with the normal head mesoderm expression of *tin* in *snail* loss-of-function mutant embryos (Figure 4.3-2 B'). The *Tin B-374* CRM is also bound by Twist *in vitro* and this binding is dependent on at least three different E-box sequences, which are required for the activity of the CRM [60].

As expected, the *Tin B-374* CRM responds to Twist in Kc cells, where 50 ng of transfected Twist induces luciferase activity to approximately 4-fold (Figure 4.3-3 C). Transfection of Snail alone does not affect the CRM activity, but addition of 50 ng of Twist in the presence of 100 ng of Snail increases the average luciferase activity to approximately 8-fold (Figure 4.3-3 C). There are four putative Snail binding sites in the *Tin B-374* CRM identified using a Position Weight Matrix (PWM) of Snail that was generated by a Systematic Evolution of Ligands by Exponential Enrichment (SELEX) [196] (Figure 4.3-3 E). However, as three Snail sites largely overlap with the Twist motifs, it was not possible to mutate these sites without also affecting Twist binding. The fourth site, named Sna4, is non-overlapping (Figure 4.3-3 E), yet its

mutation does not have any effect on the activity of the CRM *in vitro* or *in vivo* (data not shown).

4.3.3.3 The *Mef2 I-D[L]* enhancer is activated by Snail *in vitro*

In a transgenic reporter assay, the *Mef2 I-D[L]* CRM drives reporter expression in the presumptive mesoderm, but not during later stages of muscle specification and differentiation [236]. The CRM is bound by Twist protein *in vitro* [59] and *in vivo* (Figure 4.3-1 A) and its activity can be induced in ectodermal cells by ectopically expressing Twist [236]. As expected, Twist can activate this CRM to 15-fold in Kc cells (Figure 4.3-3 D). Transfection of Snail alone has no effect on CRM activity, however co-transfection of both Snail and Twist almost doubles the average luciferase activity (Figure 4.3-3 D). Similar to the *Tin B-374* CRM, Snail therefore appears to stimulate the activity of the *Mef2 I-D[L]* CRM in the presence of Twist, but it does not seem to be able to activate its expression alone.

There are three putative Snail binding motifs in the *Mef2 I-D[L]* CRM, one of which largely overlaps with a Twist site that is required for Twist binding *in vitro* [59]. Mutating this site in such a way as to preserve Twist binding, in combination with mutations in the other two sites (Figure 4.3-3 F), results in an enhancer (named *Mef2 I-D[L] Sna123*) that is still activated by Twist, albeit to a lower level (Figure 4.3-3 D). Importantly, the synergistic activation of Snail with Twist is now lost (Figure 4.3-3 D), demonstrating, that the positive effect of Snail is dependent on its ability to bind to DNA.

In summary, Snail significantly contributes to the activation of three well-characterized early mesodermal CRMs and in the case of *Mef2 I-D[L]* enhancer, the Snail binding motifs are required for this positive regulatory input.
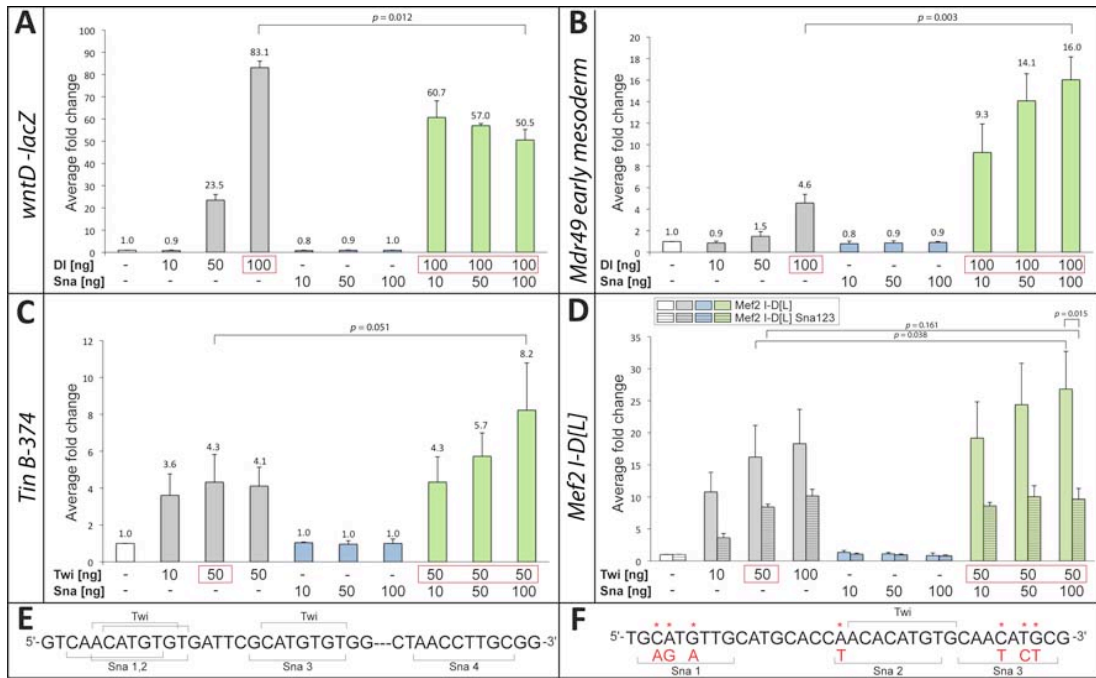
Figure 4.3-3: Snail acts to synergistically enhance CRM activity in combination with Dorsal or Twist

(A) Activity of the Dorsal-activated CRM *wntD-lacZ* is gradually suppressed with increasing amounts of Snail transfected (n=2).

(B) *Mdr49 early mesoderm* CRM is activated by Dorsal weakly, but co-transfection of Snail increases it activity almost four-fold (n=3).

(C) *Tin B-374* CRM is also activated by Twist alone, yet its activity is enhanced if Snail is co-transfected (n=3).

(D) The wild type *Mef2 I-D[L]* CRM (full bars) is activated by Twist, but upon addition of Snail the average luciferase activity is approximately doubled (n=3). If the putative Snail binding motifs are mutated (striped bars), addition of Snail does not increase the basal activity levels achieved with Twist (n=3).

(E) Putative Snail and Twist binding motifs in *Tin B-374* CRM.

(F) Putative Snail and Twist binding motifs in *Mef2 I-D[L]* CRM. Red asterisks and residues mark mutations in *Mef2 I-D[L] Sna123*.

4.3.4 *In vivo* activity of mesodermal enhancers further supports the ability of Snail to positively regulate transcription

In order to test whether Snail can contribute to the activation of mesodermal CRMs *in vivo*, I generated transgenic reporter lines where each CRM was linked to a *lacZ* reporter gene. To enable a direct comparison between the *lacZ* spatial expression and levels of expression in different lines, site-directed φC31-mediated transgenesis was used [197]. As all the discussed CRMs were targeted to the same landing site, their activity can be directly compared without the need to consider variation due to positional effects of different integration sites.

4.3.4.1 Snail binding sites are required for *Mef2 I-D[L]* expression *in vivo*

The expression of a *lacZ* reporter driven by the *Mef2 I-D[L]* CRM is initiated at early stage 6 in most, but not all, cells expressing the endogenous *Mef2* transcripts (Figure 4.3-4 A, A'). At stage 7, during germ band elongation, there is robust *lacZ* expression in the entire trunk mesoderm (Figure 4.3-4 B, B'). Up to stage 8, both enhancer-driven *lacZ* and the endogenous *Mef2* mRNA co-localize with endogenous *sna* expression (Figure 4.3-4 C-C"). At the end of stage 8, the expression of *sna* in the mesoderm diminishes and instead appears in cells of the neurogenic ectoderm [245]. *lacZ*-driven enhancer activity persists in the mesoderm until the end of stage 10 and ceases with the onset of segmentation and mesoderm subdivision at stage 11 (data not shown).

Mutagenesis of the putative Snail binding motifs in the *Mef2 I-D[L]* CRM results in loss of its activation by Snail in *Drosophila* cells (Figure 4.3-3 D, F). This enhancer can still be activated by Twist, although to a lesser extent. One would therefore predict that the mutated CRM is still, at least partially, activated by Twist and therefore expressed *in vivo*. When the activity of the *Mef2 I-D[L] Sna123* CRM was assayed in transgenic flies with a probe directed against *lacZ*, it indeed showed lower, but not entirely absent, expression. At stage 6, there are some *lacZ* positive cells (Figure 4.3-4 D), but their number is lower when compared to the original CRM (Figure 4.3-4 A). At later stages, during which *sna* mRNA is still present in the mesoderm (Figure 4.3-4 E", F"), the expression of *lacZ* is almost completely lost, with the exception of strong expression in a posterior patch of the trunk mesoderm,

corresponding to the precursors of the longitudinal visceral muscles (Figure 4.3-4 D, E).

These observations are consistent with the *in vitro* results; the *Mef2 I-D[L] Sna123* CRM shows only residual activity compared to the wild type CRM. The Snail binding motifs are therefore also required for the activation of the *Mef2 I-D[L]* CRM *in vivo*.
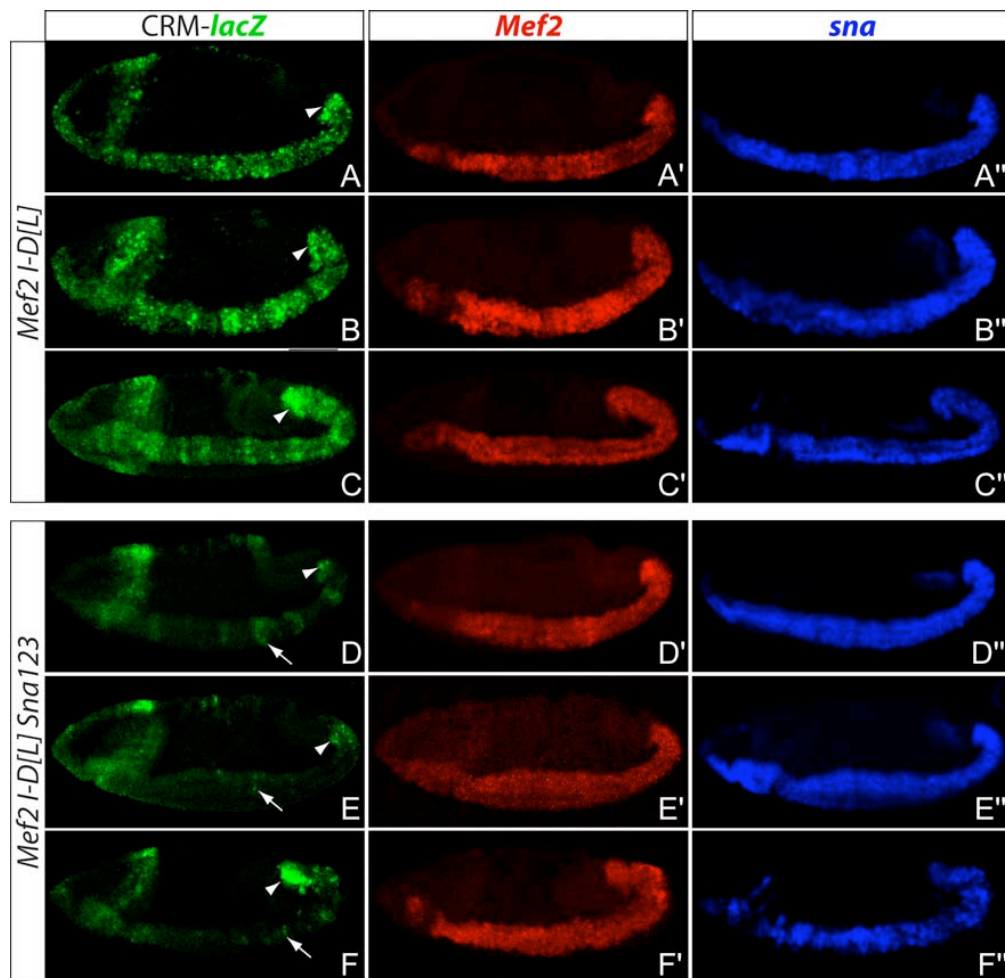
Figure 4.3-4: *In vivo* **activity of wild type and mutated** *Mef2 I-D[L]* **CRM**

The levels of CRM activity (green) as well as *Mef2* (red) and *sna* (blue) endogenous gene expression were detected by a fluorescent *in situ* hybridization.

(A-C) The wild type *Mef2 I-D[L]* CRM drives *lacZ* expression (green) in an overlapping domain to the endogenous *Mef2* gene (red). *sna* (blue) is also expressed in the mesodermal cells at stage 6 (A"), 7 (B"), and 8 (C").

(D-F) When the putative Snail sites in the *Mef2 I-D[L]* CRM are mutated, the expression of *lacZ* (green) in the trunk mesoderm is almost completely absent (arrow). However, expression of the endogenous *Mef2* gene (red) or *sna* (blue) is not affected in stage 7 (D, E) and 8 (F) embryos. Arrows indicate loss of mesodermal expression and arrowheads point to expression in the posterior-most mesodermal cells which will be contribute to longitudinal visceral muscles.

The *lacZ* expression in the anterior region is due to a non-specific transgenesis effect as it is also present in embryos transgenic for the empty vector (R. Zinzen, personal communication).

4.3.4.2 Expression of the *Mef2 I-D[L]* CRM is dependent on Snail activity *in vivo*

The expression of the *Mef2* gene is lost in embryos homozygous for a *sna* amorphic allele (Figure 4.3-2 A'). As *Mef2 I-D[L]* is the only CRM known to drive the early mesodermal expression of *Mef2* [59], I tested whether Snail is necessary for its activity *in vivo*. For this purpose, the CRM was placed into the genetic background of two different *sna* loss-of-function alleles. sna^{18} is an amorphic allele identified in an EMS screen and has been classified as one of the strongest *sna* mutant alleles that have been isolated to date [239, 240]. The precise lesion has not been mapped, but both Snail *mRNA* and protein are not expressed in homozygous mutant embryos [90]. These embryos misexpress neuroectodermal genes such as *sim* and *rho* in the mesoderm [90], while the expression of several mesodermal genes is lost (Figure 4.3-2). The mutation in the hypomorphic allele sna^{10} was characterized and mapped to a single base pair that leads to a conversion of a glycine to a glutamic acid [90]. The mutated residue is located between two of the four Snail zinc fingers [90] and it is currently not understood how this change in a single amino acid affects the DNA binding properties or the general activity of Snail. However, the effect of the mutation on the expression of Snail downstream genes is rather peculiar. Derepression of *sim* is even more pronounced compared to sna^{18} homozygous embryos, while the expression of mesodermal genes *zfh1* and *srp*, both absent in sna^{18} embryos, is not affected [90]. The differential response of neuroectodermal and mesodermal genes to the mutated Snail protein in sna^{10} may therefore provide clues regarding the mechanism behind the distinction of inhibitory and activatory functions of the Snail protein.

Embryos collected from either *Mef2 I-D[L]* sna^{18}/*CyO ftz-lacZ* or *Mef2 I-D[L]* sna^{10}/*CyO ftz-lacZ* adults were subjected to triple fluorescent *in situ* hybridization assays. The genotype of the embryos was determined with a *sim* probe, using derepression of *sim* in the mesoderm as a hallmark of both alleles. In addition, the probe directed against *lacZ* can distinguish the homozygous mutant embryos via an absence of *ftz-lacZ* expression on the balancer chromosome. Finally, a probe against the endogenous *Mef2* transcripts revealed whether the changes at the CRM and at the gene level are consistent.

In sna^{18} mutant embryos, *sim* is misexpressed in some mesodermal cells and the endogenous expression of *Mef2* is lost (Figure 4.3-5 A'-A", B'-B"). As expected, the

activity of the *Mef2 I-D[L]* CRM is also affected and no *lacZ* expression could be detected in mesodermal cells of stage 7 or 8 embryos (Figure 4.3-5 A, B). However, the situation is different in embryos homozygous for the *sna*[10] allele. The derepression of *sim* in the mesoderm is substantially more robust than in *sna*[18] embryos, as it is widely expressed in the mesoderm (Figure 4.3-5 C", D", E"). Both the distribution and the levels of *Mef2* mRNA, as well as the CRM-driven *lacZ*, seem to be comparable to wild type in most embryos (Figure 4.3-5 C-C', E-E'). Interestingly, in some *sna*[10] homozygous mutant embryos the CRM is inactive in the mesoderm with the exception of a posterior patch (Figure 4.3-5 D), whereas expression of the *Mef2* gene appears unaffected (Figure 4.3-5 D').

In summary, the expression of *Mef2* in the two different alleles is in agreement with that of the previously observed mesodermal genes *srp* and *zfh1* [90]; in *sna*[18] homozygous embryos it is completely absent while in *sna*[10] embryos it seems unaffected. In addition, these changes are recapitulated at the CRM level as *Mef2 I-D[L]* activity is lost in *sna*[18] but does not seem to be affected in most of the *sna*[10] homozygous embryos. Snail is therefore required for the activation of the *Mef2 I-D[L]* CRM *in vivo*. It is likely that the absence of endogenous *Mef2* mRNA in *sna*[18] embryos is due to the inactivity of the *Mef2 I-D[L]* enhancer. However, existence of an as yet unidentified early mesodermal *Mef2* CRM that would also be responsive to Snail cannot be excluded.
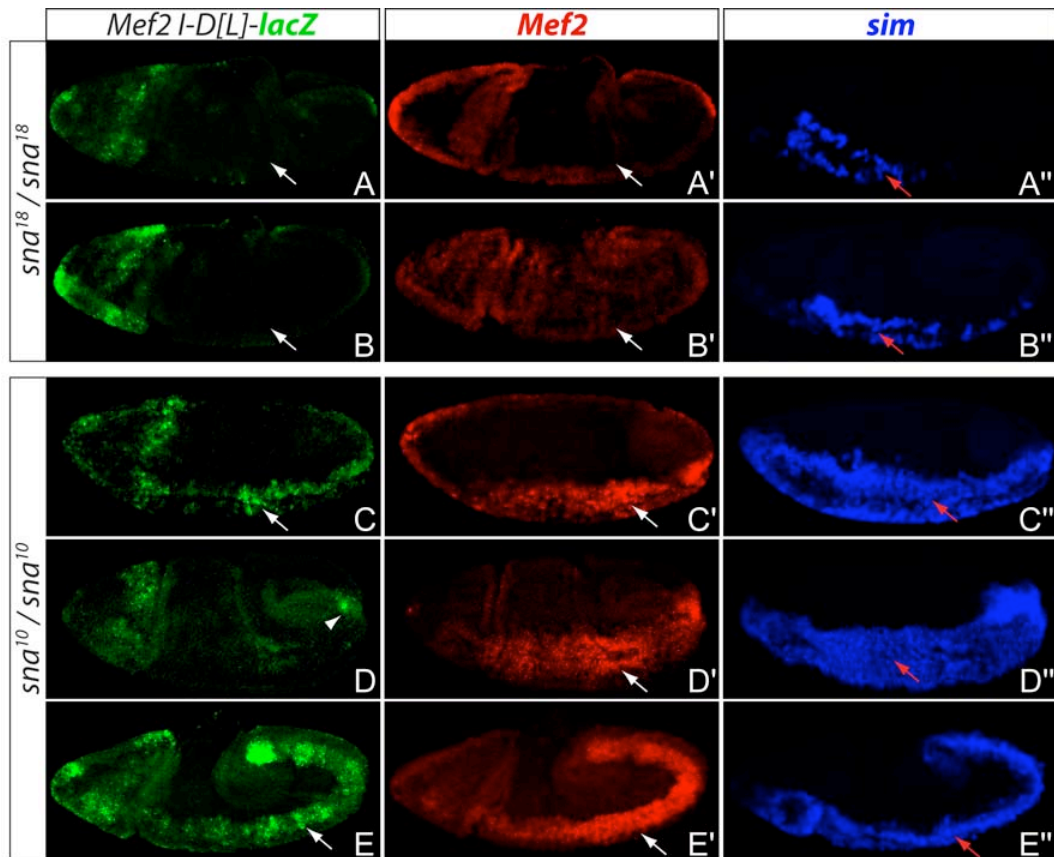
Figure 4.3-5: The *Mef2 I-D[L]* CRM is dependent on Snail activity *in vivo*

The *Mef2 I-D[L]* CRM was placed in the genetic background of two different *sna* alleles. Embryos were subjected to triple fluorescent *in situ* hybridization with anti-sense probes against *lacZ* (green), *Mef2* (red), and *sim* (blue). Absence of *lacZ* signal in the *ftz* expression domain and derepression of *sim* in the mesoderm (red arrows) are distinctive of *sna* homozygous mutant embryos.

(A-B'')   In *sna^18^* homozygous embryos of stage 7 (A-A'') and 8 (B-B''), no *lacZ* (green) or *Mef2* (red) can be detected in the presumptive mesoderm (white arrows).

(C-E'')   In *sna^10^* homozygous embryos of stage 6 (C-C''), 7 (D-D''), and 8 (E-E''), endogenous *Mef2* (red) appears to be expressed at wild type levels. *lacZ* driven by the *Mef2 I-D[L]* CRM is present in the mesoderm of most embryos (C, E) but is absent in some embryos (with the exception of patches of posterior-most cells, white arrowhead) (D).

4.3.4.3 *In vivo* activity of the *Mdr49 early mesoderm* and the *Tin B-374* CRMs

The *lacZ* reporter driven by the *Tin B-374* CRM recapitulates the pattern of the endogenous *tin* expression in the trunk mesoderm from stage 5 until stage 10. At stage 7 it is strongly expressed in the trunk mesoderm, while the head mesoderm expression of the endogenous *tin* expression is not covered by the CRM (Figure 4.3-6 A, A'). As reported previously, the CRM expression is patterned along the anterior-

posterior axis by Eve, which directly binds the CRM and suppresses its activation by Twist [60]. As a result, both the CRM and *tin* gene show striped expression with lower levels in the Eve-expressing domains (Figure 4.3-6 A, A').

In stage 6 embryos homozygous for the *sna*[18] allele, the activity of *Tin B-374* is almost completely absent, with the exception of a few mesodermal cells in the Eve-negative stripes (Figure 4.3-6 B). At stage 7, no *lacZ* signal can be observed in the mesoderm (Figure 4.3-6 C). In agreement with the colorimetric *in situ* hybridization assay (Figure 4.3-2 B, B'), *sna*[18] homozygous embryos express no endogenous *tin* in the presumptive mesoderm at stage 6 or 7 (Figure 4.3-6 B', C'). However, in the background of *sna*[10] allele, both the *Tin B-374* CRM and the *tin* gene show wild type expression (Figure 4.3-6 D-E").

The *Mdr49 early mesoderm* CRM drives expression of *lacZ* only transiently during stages 5 and 6 in the presumptive mesoderm (Figure 4.3-6 F, G). Similar to the endogenous gene, the expression of the CRM does not seem to be affected in *sna*[18] homozygous embryos (data not shown), suggesting that the mode of its regulation by Snail might be different from *Mef2* and *tin*.

Taken together, *tin* and *Mef2* show the same response to two different *sna* alleles at the gene, as well as at the CRM, level. Their activity is almost completely abolished in *sna*[18] homozygous mutant embryos in which the expression of other mesodermal genes such as *htl*, *zfh1*, or *srp* is also affected ([90], Figure 4.3-2). In contrast, the expression of these genes remains unchanged in *sna*[10] embryos where the Snail protein bears a single amino acid mutation. These experiments therefore reveal a new role for Snail as a transcriptional activator of mesodermal genes and in addition identify their Snail-responsive regulatory elements.
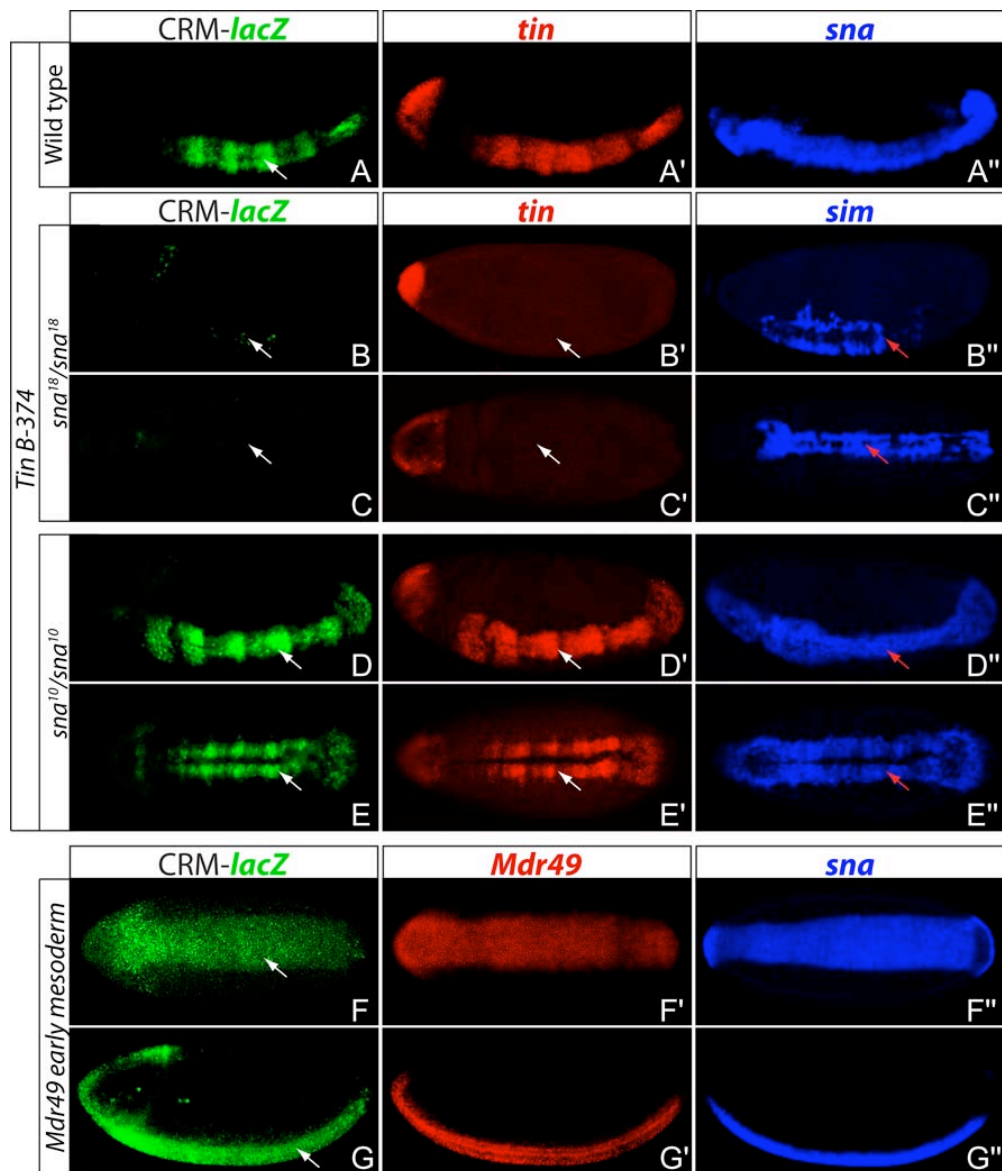
Figure 4.3-6: Expression of the *Tin B-374* and *Mdr49 early mesoderm* CRMs

In vivo activity of the CRMs was assayed by fluorescent *in situ* hybridization with a *lacZ* probe (green). In addition, the endogenous gene expression of *tin* or *Mdr49* is shown in red and the expression of *sna* in wild type or *sim* in mutant embryos is shown in blue.

(A-A'') *LacZ* driven by *Tin B-374* CRM (A) can be detected in the trunk mesoderm where it overlaps with *tin* (A') and *sna* (A'').

(B-C'') In embryos homozygous for the amorphic *sna^{18}* allele, the activity of *Tin B-374* is substantially reduced at stage 6 (B) and almost completely absent at stage 7 (C). The endogenous *tin* transcripts are completely absent in the trunk mesoderm at both stages (B', C'). Expression of *sim* mRNA (blue) is derepressed in some mesodermal cells (red arrow) and was therefore used to recognize homozygous mutant embryos. B-B'' is a lateral and C-C'' a ventral view.

(D-E'') In the genetic background of a hypomorphic allele *sna^{10}*, neither the *Tin B-374* CRM activity (D, E) nor the endogenous *tin* expression (D', E') is affected while the neuroectodermal gene *sim* is expressed throughout the mesoderm (D'', E'') during stages 6 (D-D'') and 7 (D-E''). D-D'' is a lateral and E-E'' a ventral view.

(F-G'') At stage 5, the *Mdr49 early mesoderm* CRM (F, G) is detected in the presumptive mesoderm where its expression overlaps with the endogenous *Mdr49* (F', G') and *sna* (F'', G'') mRNA. F-F'' is a ventral and G-G'' a lateral view.

5   DISCUSSION

5.1 Different approaches for forward genetic screening in *Drosophila*

In the traditional, forward genetic screens pioneered in *Drosophila* [246], a collection of randomly mutagenized lines was generated and then evaluated for a phenotypic trait of interest. Although these screens usually identified tens or even hundreds of new genes involved in a given process, they were exceptionally labour-intensive, depending on the phenotypes being assayed. Nowadays, there are diverse screening approaches and toolkits available and the selection of the most suitable one depends on the studied system and the objectives of the experiment.

Here, I used a combination of multiple gene deletion or silencing methods to identify novel regulators of *Drosophila* embryonic myogenesis. As my search for novel genes was restricted to transcription factors (TFs), I did not perform a random mutagenesis screen. Instead, based on prior knowledge, I defined a set of candidate genes whose requirement for muscle development was systematically tested in a phenotypic assay.

5.1.1   The advantages and limitations of deficiency lines

In the second round of the screen, the functional requirement of mesodermally expressed candidate TFs for muscle development was analyzed using deficiency lines. As there are hundreds of different lines available, the deficiencies offered an efficient way of identifying genetic loci required for myogenesis and also to eliminate genes encoding TFs. However, the lines delete large genomic segments that can cover up to hundreds of genes, therefore another round of analysis is required to define the genes responsible for the observed phenotypes.

The fact that the deficiency lines remove multiple genes can at the same time be convenient as it allows a rapid screening and narrowing down of the regions of interest. Another great advantage of the deficiency lines is that they are *bona fide* null alleles for the deleted genes. In contrast, knockdown with RNA interference (RNAi) or random mutagenesis can give rise to a series of phenotypes (see Section 5.1.2.2) of variable strength that can complicate the analysis.

On the other hand, deficiency lines have many limitations that were also encountered in the course of the screening. Some of them were circumvented by considerate selection of the lines, while others required additional experiments. The most commonly experienced difficulties were the following:

o   As mentioned, the foremost limitation of deficiency lines is the deletion of a large number of genes. The size of the deleted segments can be reduced by combining two partially overlapping lines *in trans*, yet tens of different gene loci are normally affected. Subsequently, secondary phenotypic analysis at a single-gene level is necessary to identify the gene of interest. Further dissection can however be elaborate as several of the deleted genes might contribute to the observed phenotype which is then the combination of multiple, more subtle defects.

o   Often, the candidate gene lies in the vicinity of a gene that is known to be involved in the assayed process. For example, the loci of some candidate mesodermal TFs, such as *CG12744* or *CG13424*, are near the locus of *Mef2* and all available deletions that cover the candidate genes also delete *Mef2*. As differentiation of all three muscle types fails in absence of *Mef2* [247], the contribution of the candidate genes to muscle development can only be assessed by approaches specifically targeting the individual genes.

o   A similar problem occurs when the gene of interest is nearby a gene whose deletion impairs analysis of the given phenotypic trait. Many times the presence of such a gene in the deleted region is not revealed before the actual analysis. Then its deletion might result in cell death before the actual phenotype can be assayed or it might interfere with developmental processes that are a prerequisite for the studied trait. Deficiency lines in this category form the phenotypic class II (Figure 4.1-5) as the muscle-specific defects in the embryos could not be evaluated. This problem can be avoided by choosing other deficiency lines that would not delete the more general genes (if their identity is known) or again by using single-gene level analysis.

o   The recent ability to systematically generate deletiom lines enables precise mapping of their endpoints by PCR-based methods. However, the breakpoints of older deficiency lines generated by X-ray mutagenesis were only estimated from the missing cytological bands in polytene chromosome squashes and by complementation analysis using other mapped lines. The uncertainty if the genes at the breakpoints of the removed region are deleted can be circumvented by preferentially choosing

deficiency lines with the candidate gene in the centre of the predicted removed segment. However, this was not always possible with the lines that were available.

o   Although there are hundreds of deficiency lines available, the genome is not yet fully saturated with deletions. For example, the candidate mesodermal TF encoded by *Misexpression suppressor of ras 4* (*MESR4*) was not covered by any deficiency line at the time of the screen. This problem was partially solved by the advent of tools for generation of custom-made deletion lines [200, 206]. As these are however dependent on random integration of transposition elements, some genomic loci might still be underrepresented among the available deficiency lines.

o   As deficiency lines remove only the zygotic component of a gene's expression, maternally expressed factors important for mesoderm specification would be missed. In addition, phenotypes of some of the mesodermal TFs whose deletion does not lead to any apparent muscle defects (class I) could be attenuated by their maternally loaded transcripts.

o   Another explanation for situations when removal of a mesodermal TF does not cause any apparent muscle abnormalities is its redundancy with another factor. Redundancy among the core myogenic regulators is a well-studied phenomenon in vertebrates, and recently also uncovered in yeast (reviewed in [248]), yet its contribution to *Drosophila* muscle development remains unknown. Although genetic interaction between two given genes can be readily tested, high-throughput screens for redundantly acting factors *in vivo* are currently not feasible.

5.1.2 Alternative approaches to large-scale phenotypic screening

The discovery that the introduction of double-stranded RNA (dsRNA) into cells leads to the silencing of gene expression [249] initiated a new era of screening techniques in numerous organisms, including *Drosophila*. Two broad categories of RNA interference (RNAi) experiments, differing in the mode of dsRNA delivery, can be used to knockdown endogenous expression of specific genes in *Drosophila* embryos.

5.1.2.1 Inducible gene silencing using the Gal4-UAS system

Libraries of transgenic fly lines, each carrying two inverted repeats of the target sequence placed downstream of the yeast UAS element, have been generated by both random, *P*-element mediated, as well as site-directed transgenesis ([250], unpublished). When a source of the Gal4 activator is supplied, the inverted repeats are transcribed and the RNA molecule folds into a hairpin of effective dsRNA [251]. The inducibility by the Gal4-UAS system [223] enables spatially and temporally controlled dsRNA expression, which in turn facilitates the assessment of the role of a gene specifically in the tissue and time of interest. It became therefore possible to address the role of genes essential for embryogenesis during larval to adult stages. Since in the remaining, Gal4-negative cells the target gene is expressed normally, developmental processes can be studied in isolation without the need to consider extrinsic, secondary effects.

At the same time, however, the indirect activation of the expression via Gal4 introduces a time delay between the induction of Gal4 expression and the actual dsRNA transcription and folding. Consequently, while the system works efficiently during later stages of the *Drosophila* life cycle, for the rapid process of embryogenesis the time delay is an impediment.

Nevertheless, in the second round of the screen I used 74 different RNAi lines to narrow down the number of candidate genes contributing to the most interesting deficiency phenotypes. With the earliest available mesodermal Gal4 driver, *twi, 24B-Gal4*, 23 lines did not produce viable offspring (see Appendix section 7.1). However, in all assayed cases, the lethality occurred after embryogenesis and the examined embryos developed wild type musculature (not shown). As UAS-RNAi knockdown of the key myogenic factors also does not affect embryogenesis (not shown), RNAi lines with the set of currently available drivers are not a suitable method for screening of embryonic muscle phenotypes.

An identical conclusion can be derived from a recent genome-wide screen which used a *Mef2-Gal4* line as a driver for almost 18,000 RNAi lines corresponding to about 10,000 different genes [252]. While knockdown of 1,969 genes led to lethality, only 31 cases (1.5 %) of embryonic lethality were reported. Gene ontology analysis in addition showed that most of the 31 genes encode structural muscle constituents rather than developmental regulators. 66 of the RNAi lines that I

analyzed were also included in the genome-wide screen and the obtained results are well-correlated. All the lines that were viable in our screen were viable in the large screen too, but 5 lines reported viable with the *Mef2-Gal4* driver are lethal in combination with the *twi, 24B-Gal4* line (see Appendix section 7.1). The differences in lethality hence underscore the importance of the driver line for timing, and probably also levels, of dsRNA expression.

5.1.2.2 Direct dsRNA injection into embryos

The time delay limitation inherent to the Gal4-UAS system can be circumvented by injecting the double-stranded RNA directly into blastoderm embryos. At the syncytium stage, when the nuclei are not yet separated by cell membranes, it can freely diffuse and spread across the embryo. When highly concentrated, *in vitro* synthesized dsRNA is injected, it is thought to also silence maternally loaded transcripts [52]. An obvious disadvantage (similar to a genetic null allele) is the ubiquitous distribution of the dsRNA resulting in a gene knockdown in all, or most embryonic cells. In addition, as it is supplied only in early embryogenesis, it may be used up or degraded before the actual process of interest occurs.

By far the most important drawback of dsRNA injections is the high variability and low reproducibility of the detected phenotypes. In a typical experiment, when about 100 embryos were injected with dsRNA against *Mef2*, a range of phenotypes was observed (Figure 5.1-1 A). Many embryos were damaged by the injection itself and some embryos exhibited wild type musculature. Only about one fifth of injected embryos displayed specific muscle defects. However, the severity of these varied greatly and resembled the series of hypomorphic *Mef2* alleles [253] (Figure 5.1-1B). This variability is likely the result of unequal volumes of dsRNA injected and its uneven distribution within the embryo. In addition, knockdown of highly expressed genes might be less efficient and the overall efficiency can vary with differently designed dsRNA sequences. To account for the variability, a high number of embryos need to be injected and evaluated to obtain a reliable result. This is fairly simple if dsRNA injections are used to confirm a known phenotype. However, if the expected phenotype is not known, as it would be in a screen, obtaining a definitive answer only from the dsRNA injections remains a major challenge.

Nevertheless, with the fast development of new technologies, automatized dsRNA injection and phenotype evaluation methods might become available. If the

variability as well as workload were reduced by robotics, high-throughput screens could be performed with the ease of the current cell culture screens. Eventually, co-injection of two or more different dsRNAs would be an efficient approach to study genetic redundancies.
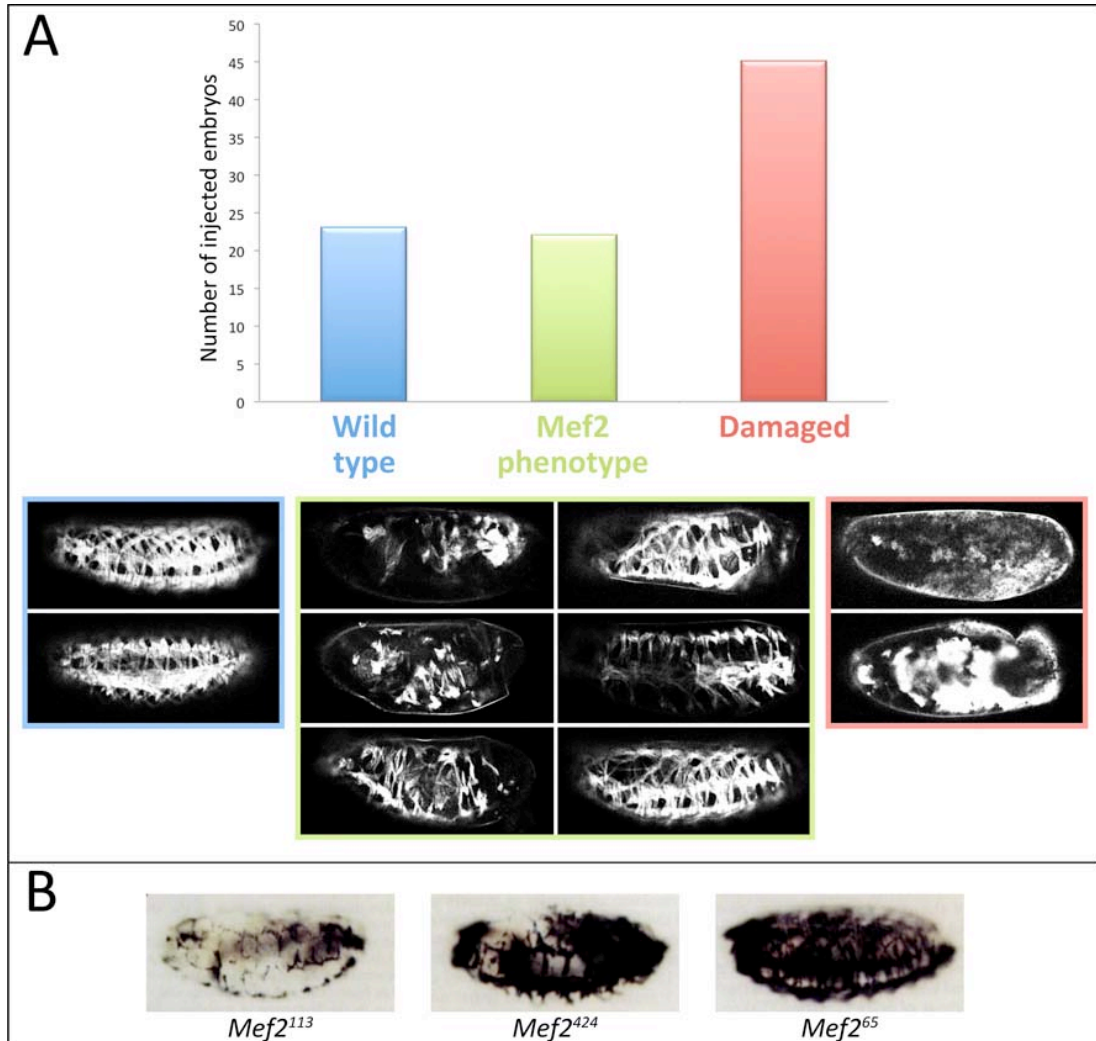
Figure 5.1-1: Injection of *Mef2* dsRNA results in a hypomorphic phenotypic series

(A) Outcome of a typical experiment when 94 embryos were injected with dsRNA against *Mef2*. The injected embryos carried a Myosin heavy chain (MHC)-τGFP transgene which labelled all muscle types with a microtubule-associated form of GFP, enabling live imaging of the specimen. Injection of the dsRNA resulted in a range of phenotypes, from wild type (blue), through muscle defects of variable severity (green), to generally damaged embryos as a result of the injection (red).

(B) A series of embryos homozygous for different hypomorphic *Mef2* alleles stained with an anti-Myosin antibody. Images are from [253].

5.2 The Mediator complex in *Drosophila* myogenesis

While the Mediator complex and other members of the general transcriptional apparatus have been implicated in vertebrate myogenesis (see Section 1.1.3.4), their role in *Drosophila* embryonic muscle development has not been directly established. The molecular screen described here provides indications that distinct subunits of the Mediator complex are specifically required for *Drosophila* myogenesis as well (summarized in Table 5.2-1), similar to the situation in vertebrates.

The observation that the injection of dsRNA against *MED4*, *MED11*, and *MED24* led to embryonic muscle defects [52] prompted us to investigate the necessity of these and other subunits for muscle development. Knockdown of each of the three and additional four Mediator subunits specifically in mesoderm leads to lethality (Table 5.2-1, Appendix 7.1). As the lethality is at rather late stages (discussed in Section 5.1.2.1), no abnormalities in embryonic musculature could be observed (not shown). However, mesoderm-specific knockdown of the key regulators Mef2, Tin, or Lmd similarly results in larval or pupal lethality, as opposed to the embryonic defects observed with null alleles ([252], unpublished). Consequently, the RNAi lines do suggest, but do not definitively confirm, that the mesodermal cells require expression of at least 7 different Mediator subunits for their survival and/or development.

Tissue-specific usage of Mediator subunits may require specific expression during mesoderm development. I therefore examined the mRNA expression of 25 subunits in an *in situ* hybridization assay. Most genes show ubiquitous expression throughout embryogenesis, but mRNA levels of several genes seem to be higher in the presumptive mesoderm compared to other tissues (Appendix 7.2). Nevertheless, the ubiquitously expressed subunits can contribute to spatio-temporal gene regulation by interacting with sequence-specific transcription factors.

Finally, two subunits, MED24 and MED14, were analysed in more detail where new specific deletion lines were generated. While deletion of *MED14* did not lead to any observable defects in embryonic musculature, *MED24* seems to be involved in the control of cell number as in its absence supernumerary myoblasts are present. Regulation of cell number might be a general role of MED24 as it was isolated in a screen for cell death regulators in salivary gland [47].

	dsRNA injection phenotype [52]	twi, 24B > RNAi lethality	Expression in the mesoderm
MED4	light green	light green	red
MED6	white	light green	dark green
MED7	white	light green	white
MED11	light green	light green	light green
MED14	white	light green	dark green
MED16	white	red	dark green
MED23	white	red	white
MED24	light green	light green	dark green
MED26	white	red	light green
MED27	white	red	light green
MED30	white	light green	light green

Table 5.2-1: Association between the Mediator complex and muscle development

Overview of three types of experimental evidence (injection of dsRNA, knockdown with UAS-RNAi line, and mRNA expression in mesoderm) for 11 different Mediator subunits. Light green indicates a positive result (presence of phenotype or expression in mesoderm), dark green indicates ubiquitous expression, red stands for a negative result (no lethality or no mesodermal expression) and white for no information available.

5.3 Tramtrack is a novel regulator of *Drosophila* myogenesis

Negative regulation of gene expression mediated by Ttk69 is essential for the proper specification of cell fates in diverse developmental contexts and an absence of Ttk69 is detrimental to several embryonic processes [155, 174, 176]. In contrast, the contribution of repression to the specification of muscle cell identity remains poorly understood and the current myogenic regulatory network is based primarily on positive regulatory inputs [54]. Using a combination of genetic as well as genomic approaches, I analysed the possibility that Tramtrack could be one of the missing links between transcriptional repression and *Drosophila* muscle development.

5.3.1 Functional analysis links Ttk69 to myoblast fusion

The *ttk* loss-of-function phenotype in the somatic muscle differs from previously described phenotypes. Whereas at earlier stages it shows signs of a classic myoblast fusion defect, at later stages only very few myoblasts remain unfused. Instead, they aggregate into large cell clusters which fail to extend and attach to the ectodermal tendon cells. This is in contrast to the typical myoblast fusion phenotype of mini-muscles where the unfused founder cells extend and form correct attachments to the epidermis.

Analysis of the markers specific for the two fusing populations of cells, founder cells (FCs) and fusion competent myoblasts (FCMs), revealed a disproportion in their number. While the number of FCMs decreases in *ttk* loss-of-function mutants, there are ectopic, FC-like cells formed. Consequently, there are at least three populations of somatic myoblasts in *ttk* mutants: original, presumably unaffected FCs, FCMs, and a population of ectopic FC-like cells. The number of cells undergoing mitosis is, however, comparable to wild type embryos, therefore it is unlikely that the imbalance between FCs and FCMs is caused by a cell cycle defect. Instead, some FCMs seem to have acquired molecular properties characteristic of FCs and have lost some of their FCM-like characteristics.

In wild type embryos, the differential expression of *sns* and *duf* (rP298-lacZ reporter) in FCMs and FCs, respectively, is essential for successful myoblast fusion.

When a source of *duf* is introduced outside of the mesoderm, it has the ability to attract FCMs which then aggregate around the site of its expression [221]. Similarly, the ectopic, *duf*-expressing FCs in *ttk* mutants seem to attract FCMs which then arrange in aggregates. As these aggregates are rather large, it is possible that the ectopic FC-like cells have an ability to fuse with fibres generated from the original FCs (Figure 5.3-1). Clustering between FC-like cells themselves may also occur as *duf* expressed in S2 cells can mediate their homotypic fusion [254]. In wild type embryos, a defined number of FCMs fuse with a single FC, and fusion between a FC and an FCM that is fusing with another FC has not been observed. The exact mechanisms controlling the number and direction of fusion events are however not known. It is possible that in *ttk* mutants these mechanisms are perturbed and the presence of ectopic FC-like cells interferes with proper fusion of the unaffected FCs as well.

The possibility that ectopic FC-like cells affect the development of "normal" FCs is supported by their inability to migrate and attach to their ectodermal sites of attachment. Expression of the founder identity factors Kr and Eve, which are assumed to dictate the specificity of guidance of particular muscles, does not seem to be affected in *ttk* mutants (not shown). It is however possible that the large, compact aggregates physically block directed migration. Alternatively, development of the tendon cells might be affected and the secreted guidance cues may be absent.

Based on the genetic analysis of *ttk69* mutant embryos, Ttk69 seems to antagonize the FC specification program in FCMs. Consistent with this, its ectopic expression in FCs results in embryonic lethality and their inability to fuse with the surrounding FCMs. This proposed model predicts that expression of Ttk69 in FCMs should not interfere with their specification as FC-specific genes are already repressed in this context. Compared to the ectopic expression in FCs, Ttk69 expression in FCMs indeed results in a milder phenotype, with some embryos surviving to pupal stages. However, mild defects in myoblast fusion can still be observed, possibly due to prolonged expression of Ttk69 after stage 14 while in the wild type embryos its expression in somatic mesoderm ceases at stage 12.

Finally, the premature expression of Ttk69 in the early mesoderm interferes with the specification of all three muscle types, implying that early Ttk69 expression is incompatible with myogenic cell fate acquisition. The timing of the onset of its expression in mesoderm must therefore be subjected to very tight regulatory control.
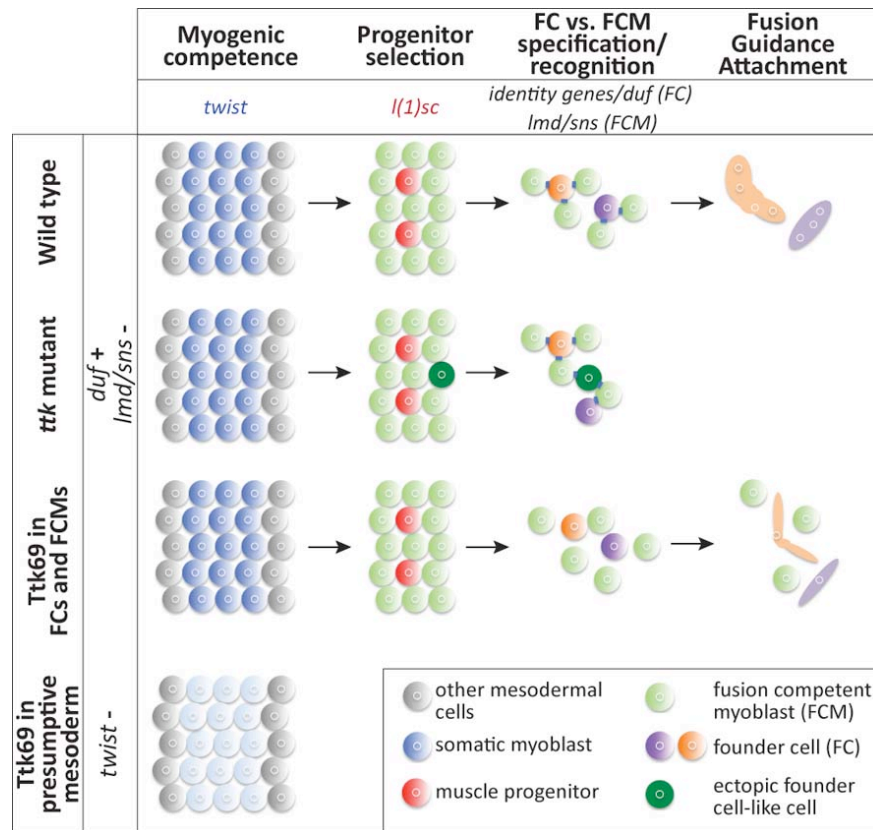
Figure 5.3-1: Schematic summary of *ttk* phenotypes

Overview of *ttk* loss- and gain-of-function phenotypes and gene misregulation based on *in situ* hybridization experiments (minus indicates downregulation, plus indicates upregulation).

5.3.2 Regulation of Ttk69 expression in the somatic mesoderm

According to the proposed model, Ttk69 represses FC-specific genes within the FCMs (Figure 5.3-3). An important prerequisite for this model is the expression of Ttk69 specifically in FCMs, but not in FCs. In spite of a large number of attempts, I failed to detect differential expression of *ttk69* mRNA in the two populations of myoblasts. However, there are several reasons why this might be challenging to capture.

First, Ttk69 expression is very dynamic and occurs within the somatic mesoderm only very shortly when the mRNA, but not the protein, can be detected during stage 11. The fact that the expression of the protein is even harder to detect is not surprising, considering the multiple mechanisms that suppress *ttk69* mRNA translation and promote protein degradation (see Section 1.4.3). In addition, small differences in the levels of Ttk69 protein in two types of tracheal cells are essential

for their subsequent fusion [162]. Similarly, the difference between *ttk69* mRNA levels between the muscle FCs and FCMs may be too subtle to be uncovered. Additional mechanisms would then prevent accumulation of functional Ttk69 protein in FCs. One such possible mechanism is proteasome-dependent degradation which reduces Ttk69 protein levels in the eye imaginal disc and requires the presence of Phyllopod, an adaptor protein linking E3 ubiquitin ligases to their targets [187, 193]. Interestingly, Phyllopod is activated by RAS/MAPK signalling not only in the eye disc [255], but also in muscle founder cells where it is necessary for the specification of distinct muscle identities [256]. Proteolysis of Ttk69 exclusively in the FCs through Phyllopod might therefore be one mode of achieving its differential expression.

5.3.2.1 Lame duck as a possible regulator of Ttk69 in FCMs

Analysis of the occupancy of Ttk69 and Lmd suggests that they may both function in specification of FCM fate, separately and jointly. However, the regulatory relationship between Ttk69 and Lmd themselves has not been explored. In *ttk^{D2-50}* homozygous embryos, the levels of *lmd* mRNA are decreased at stage 13, but this is probably a secondary effect due to decreased number of correctly specified FCMs rather than a direct regulatory input of Ttk69.

Several observations however suggest that Lmd could contribute to Ttk69 activation directly:

o Lmd and Mef2 co-bind to multiple CRMs in the vicinity of *ttk* locus (Figure 5.3-2).

o One of the CRMs associated with *ttk* was previously tested by P. Cunha in our lab, using luciferase reporter assay [214], where *Drosophila* S2 cells were transiently transfected with Lmd and Mef2 expression vectors. The assay revealed that either Lmd or Mef2 alone could weakly activate the CRM, while their combination led to a dramatic enhancement of its activity. Lmd and Mef2 might therefore cooperatively activate *ttk69* expression also *in vivo*.

o Expression profiling of *lmd* mutant embryos revealed a slight decrease of *ttk69* mRNA levels when compared to their wild type counterparts [257].

Twist may also activate *ttk69* expression by binding to multiple regions within the *ttk69* locus (Figure 5.3-2). However, the input from Lame duck (possibly in co-

operation with Mef2) may be another mechanism by which the presumed FCM-specific expression of Ttk69 is reinforced.
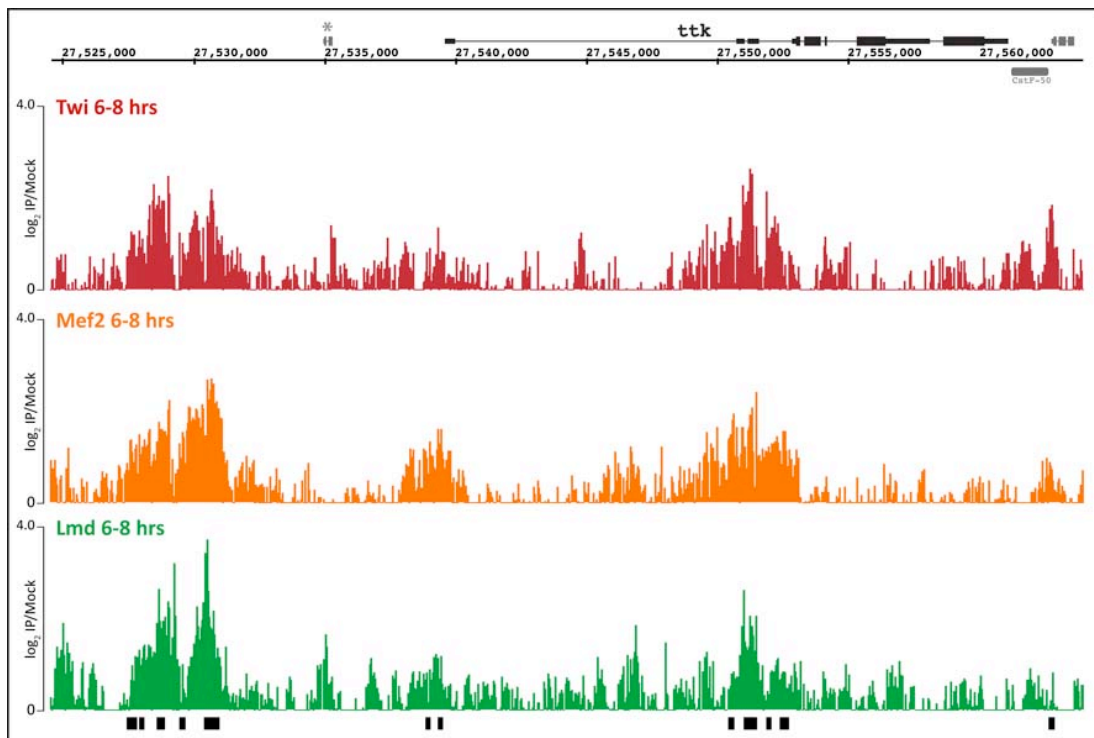
Figure 5.3-2: Twist, Mef2, and Lmd co-occupy CRMs within the *ttk* locus

ChIP-on-chip binding profiles of Twi, Mef2 (from [196]), Lmd, and mesodermal CRMs bound by at least one core myogenic TF at one assayed time point (black, from [196]). An asterisk marks a gene annotated only with the Release 5 of *Drosophila melanogaster* genome whose existence is considered dubious [258].

5.3.2.2 Notch signalling may contribute to Ttk69 regulation in the mesoderm

The Notch pathway is essential at multiple stages of somatic mesoderm development (see Section 1.3.4.1). Ttk69, in turn, genetically interacts with Notch in several distinct cell types. It is therefore tempting to speculate that Notch signalling and Ttk69 are associated in their function in muscle development.

Selection of a single progenitor from the proneural and myogenic competence fields is controlled by a remarkably similar set of regulators and principles. In the nervous system, Notch signalling suppresses neuronal fate and in the mesoderm it antagonizes myogenic fate. Ttk69 acts downstream of Notch signalling and represses neuronal genes in non-neuronal cells [163, 182]. Similarly, Ttk69 may be activated by Notch signalling in the muscle non-progenitor, FCM cells, where it could repress

the myogenic, or FC-specific, genes. Notch activity is necessary for the activation of *lmd* in the somatic mesoderm [132], hence it could in parallel contribute to *ttk69* activation. The proposed model of the role of Ttk69 role in myogenesis would therefore be consistent with its activation by Notch signalling and it could be a novel Notch effector reinforcing FCM fate. However, experimental evidence for a genetic interaction between Notch and Ttk69 would be required to confirm their association in the somatic mesoderm.
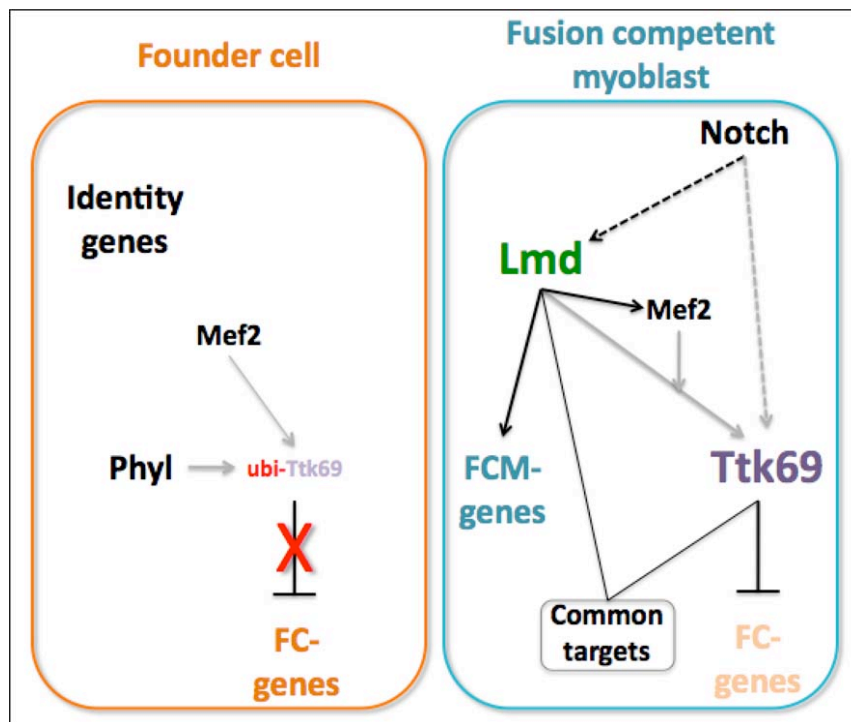
Figure 5.3-3: Model of a Ttk69 expression and function in FCM specification

A simplified model of experimentally confirmed and predicted regulatory connections upstream and downstream of Ttk69. Dashed lines indicate indirect regulation, solid direct regulation, and the connections in grey are hypothetical and would require experimental evidence. Ubi- ubiquitinated.

5.3.3 *In vivo* activity of Ttk69 CRMs suggests that Ttk69 might contribute to transcriptional activation

5.3.3.1 Ttk69 CRMs have diverse spatio-temporal activities

Analysis of Ttk69 CRMs identified in the ChIP-on-chip experiment provided several unexpected insights into the regulatory potential of Ttk69:

o All three tested CRMs drive a highly dynamic *lacZ* reporter expression in complex spatio-temporal patterns throughout embryogenesis.

o The CRM activity peaks with the onset of segmentation, coinciding with the robust initiation of Ttk69 expression in the mesoderm, nervous system, and ectoderm (Table 1.4-1). The CRMs drive expression in the tissues where Ttk69 itself is expressed, such as visceral mesoderm, tracheal placodes or nervous system. Ttk69 is however not present in all the cells within these tissues, for example in the nervous system it is only expressed in glia and not neurons [155]. Therefore a double *in situ* hybridization assay against *ttk69* and *lacZ* would be necessary to reveal their overlapping or complementary expression.

o In the case of one CRM, *jumuA*, mutagenesis of the Ttk69 binding sites resulted in ectopic *lacZ* expression in seven transverse stripes at stage 5. Maternal Ttk69 has a well documented role in the repression of pair rule genes [160]. Interestingly, Ttk69 itself is expressed ubiquitously so derepression of the CRM in vertical stripes cannot be simply explained by the inability of Ttk69 to repress the mutated CRM in these domains. However, Ttk69 may act as a short-range repressor and counteract activation of the CRM by a transcriptional activator that is expressed in a pair rule-like fashion. This would suggest that in the wild type situation, the CRM is inactive not because of a lack of activation but rather because of active repression by Ttk69. A mechanism of active repression is used during maternal-zygotic transition when a progressive dilution of Ttk69 protein levels was hypothesized to establish the onset of zygotic transcription [169]. The wild type *jumuA* enhancer however remains inactive until the end of stage 10, which is several hours after depletion of maternal Ttk69 protein, arguing against its control at the maternal-zygotic transition.

o Most CRMs show decreased activity upon mutagenesis of Ttk69 binding sites. The reduction is not global, but rather specific as it differs between the CRMs and the sites of activity. *jumuA-Ttkmut* not only shows derepression at stage 5, but also

reduced tracheal expression at stage 11. The wild type CRM *jumuB* is expressed in trachea as well, but its tracheal activity is unaffected when Ttk69 sites are mutated. Instead, *jumuB-Ttkmut* is active in fewer cells of the nervous system compared to the wild type enhancer. The loss of expression upon mutagenesis of Ttk69 sites implies that Ttk69 is needed for the activation of these CRMs rather than repression.

o Finally, all three wild type CRMs drive strong expression in the caudal visceral mesoderm, the primordium of the longitudinal visceral muscle. In mutant versions of all CRMs, this very specific and strong expression is completely lost. This result is even more surprising considering that I could not detect Ttk69 expression in these cells. A possible reason for the lost expression could be that the mutagenesis of Ttk69 sites interfered with the binding sites of another, positive regulator, which is not able to bind and hence activate the mutant CRMs in the cells of caudal visceral mesoderm.

Importantly, in the CRMs *jumuA* and *jumuB* only some of the predicted Ttk69 binding sites were mutated (see Appendix section 4.4 for sequences). Therefore it cannot be excluded that the complex changes in their CRM activity reflect residual binding of Ttk69. The Ttk69 motifs might be used differentially across space and time so the generated mutations could affect expression only in some tissues or at some time points. Consequently, a complete mutagenesis of all putative Ttk69 might result in different *cis*-activities.

5.3.3.2 Ttk69 as a potential transcriptional activator

As the analysis of the activity of Ttk69-bound CRMs *in vivo* activities suggests, Ttk69 may not be a dedicated repressor as it has been traditionally thought. Instead, it could act as a bimodal transcriptional regulator that is acting in a context-dependent manner. To validate that the loss of enhancer activity upon site mutagenesis is indeed due to absence of Ttk69 binding, the activity of the wild type CRMs should be investigated in *ttk69* loss-of-function background. Such an experiment would also clarify whether the activity of wild type CRMs in the caudal visceral mesoderm is indeed genetically dependent on Ttk69.

Another piece of evidence supporting a positive regulatory role of Ttk69 comes from studies during tracheal development [162]. In *ttk69* mutant embryos, expression of *bnl* and *pyd* is lost in the trunk visceral mesoderm and tracheal placodes, respectively [162]. Although this loss of expression was presumed to be an indirect

effect, our ChIP-on-chip data identified Ttk69 binding in the vicinity of these genes, suggesting that it may be direct. Additional experimental evidence would therefore be required to confirm the direct positive regulatory input from Ttk69.

5.3.4 The many aspects of Ttk69 regulatory roles

Dissection of the myogenic role played by Ttk69 has revealed a remarkable variety of processes in which Ttk69 is likely to be involved:

o Primarily, by repression of founder cell-specific genes, Ttk69 participates in the specification of fusion competent myoblasts and its presence is vital for the proper balance between the number of founder cells and fusion competent myoblasts.

o Ttk69 and Lmd regulate distinct sets of target genes, yet they co-occupy a significant number of mesodermal CRMs. As there are indications that both transcription factors might be able to activate as well as repress transcription [257], the *cis*-regulatory input to their combinatorial binding might be highly diverse (Figure 5.3-3). However, the possibility that the co-binding does not occur within a single cell cannot be excluded as the CRMs might be activated by Lmd in the FCMs but repressed by Ttk69 at other sites of its expression, such as visceral mesoderm, ectoderm, or endoderm.

o In addition to the developmental cell fate choice between two closely related cell identities, Ttk69 might also act across tissues to reinforce tissue-specific gene expression. The experimental evidence comes from the mesodermal expression of a neuronal marker and the ectodermal and endodermal expression of a founder cell marker in *ttk69* mutant embryos.

o Computational analysis, validated by genome-wide *in vivo* occupancy studies, implicates Ttk69 in incoherent feed-forward regulatory loops. Through combinatorially bound mesodermal CRMs, Ttk69 may act to control temporal expression of early mesodermal genes.

o Ttk69 may not act as a dedicated repressor but instead may contribute to transcriptional activation. Mutations in a number of Ttk69 binding motifs result in specific loss or downregulation of CRM activity. A number of genes are also genetically dependent on Ttk69 for their active expression [162], further supporting its possible function in positive gene regulation.

5.4 Snail is a bimodal factor that can both activate and repress transcription

Based on multiple experimental observations in *Drosophila* as well as other species, the possibility that Snail plays a dual regulatory function as a transcriptional activator and repressor was investigated. Luciferase reporter assays in *Drosophila* cells showed that Snail can indeed stimulate activity of three mesodermal CRMs. Interestingly, its positive function seems to be dependent on other factors, as Snail transfection alone does not affect reporter levels. Snail enhanced Twist-mediated activation of two mesodermal enhancers approximately 2-fold, while in combination with the TF Dorsal it could stimulate the CRM activity to more than 4-fold. Mutagenesis of the putative Snail binding motifs in one enhancer, *Mef2 I-D[L]*, showed that direct DNA binding is necessary for Snail to elicit this transcriptional activation both *in vitro* and *in vivo*.

At these same stages of development and in the same cells, Snail acts to repress non-mesodermal genes. How can Snail discriminate between its positive and negative targets to elicit two opposite transcriptional responses? Two possible models, for which there is also precedence in other developmental contexts, may explain this.

According to the first model, the DNA sequence recognized by the Snail protein plays an active role in dictating the direction of its regulatory input. Several bimodal transcription factors undergo allosteric changes upon binding to DNA (reviewed in [259]) which in turn might lead to direct protein-protein interaction with specific protein factors. In the simplest scenario, there would be two versions of Snail binding motifs, leading to different conformational changes of Snail upon DNA binding and therefore recruitment of distinct co-regulators. For example, the repressive motif would expose the protein sequences mediating interaction with the co-repressors CtBP [13] and Ebi [9], whereas the activatory binding site would lead to interaction with transcriptional co-activators. A similar mechanism was suggested for the DNA-binding protein dTCF, the effector of *Drosophila* Wingless (Wg) signalling. dTCF was traditionally viewed as a transcriptional activator induced by Wg signalling, but recently, genes repressed by dTCF were identified in several developmental processes [260, 261]. Analysis of the *cis*-regulatory region of one of these genes revealed a binding motif distinct from the typical dTCF binding site necessary for activation

[262]. In addition, the two different modes of activity depend on different domains of the TCF partner protein Armadillo, suggesting that they may be responsible for recruiting distinct co-factors [262]. If Snail functions in a similar manner, it should be possible to identify different Snail binding sites within the mesodermal enhancers compared to non-mesodermal enhancers.

In an alternative model, Snail would use one type of binding motif, but the response to Snail binding would be controlled by additional factors bound in its vicinity. Interaction with such factors would influence, possibly again via conformational changes, whether a co-activator or a co-repressor would be recruited. The morphogen TF Dorsal functions in a context-dependent manner. Although Dorsal is inherently a weak transcriptional activator, it can also repress specific genes. The Dorsal binding motifs that are necessary for the two modes of regulation are very similar. When the Dorsal binding sites in an activated enhancer are replaced by the sites from a repressed enhancer, Dorsal can still activate the enhancer [263], suggesting that it is not the binding sites themselves, but rather the context of the enhancer that influences the direction of regulation. If Snail requires an additional DNA-binding factor to determine the output of its regulation, the binding motifs of this factor should be found in the vicinity of the Snail-bound sequences.

To elucidate which mechanism is used by Snail, Snail-bound enhancers were classified as activated or repressed and scanned for differences in their sequence composition (A. Stark, unpublished data). Unfortunately, up to now no alternative Snail binding motif or a binding site of another regulator was found to be enriched in either of the two enhancer classes. It cannot be excluded that Snail operates by another, novel mechanism. An informative experiment would be to purify Snail protein complexes bound to mesodermal and non-mesodermal enhancers and subsequently compare the proteins interacting with Snail in each context.

The importance of Snail contribution to transcriptional activation is exemplified by the gastrulation defects and loss of mesodermal gene expression in *snail* mutants. Being one of the earliest zygotically expressed genes [264], in wild type embryos Snail might be necessary to establish fast and robust activation of the mesodermal specification program in combination with Twist and Dorsal. What is the exact mode of co-regulation between these three TFs and whether they interact physically remains to be addressed.

6   CONCLUSIONS AND FUTURE DIRECTIONS

As we are only beginning to decode the regulatory events leading to muscle specification and differentiation, the currently known regulators cannot generate the diversity and complexity of observed *cis*-regulatory responses. Other, as of yet unidentified, transcription factors clearly must play a role. I therefore initiated my PhD. research by conducting a molecular screen with the main objective to identify transcriptional regulators that are required for muscle development and hence are likely to contribute to the control of mesodermal gene expression. The screen identified 20 genetic loci, deletion of which led to a spectrum of somatic muscle phenotypes. Although the extent and strength of the mutant phenotypes varied, myoblast fusion, determination of distinct muscle fibre identity, and myotube guidance were the most frequently affected processes.

Since the used deficiency lines typically deleted multiple gene loci, an additional dissection of the genomic region is required to identify the individual gene responsible for the observed muscle phenotype. The regions whose deletion led to the most severe defects, could be systematically dissected using smaller deletion lines generated with available transposon-based recombination systems. Alternatively, the deleted genes could be knocked down by RNA interference once its technical limitations are resolved. When the identity of causative gene is known, its role in muscle development could be further characterized by genetic as well as genomic approaches. I used a combination of these approaches to perform a detailed analysis of the role of members of the Mediator complex and the transcriptional repressor Tramtrack in muscle development.

A more thorough analysis of two genes encoding subunits of the general co-activator Mediator complex was conducted. Although MED24 seems to play a role in mesodermal cell proliferation rather than specification, the results of RNA interference experiments point to a broader function of the Mediator complex in muscle development. Given the accumulating evidence for specific developmental roles of distinct Mediator subunits in *Drosophila* as well as other species, it probably remains only a matter of time until a link between the general transcriptional apparatus and embryonic muscle development will be established.

One of the more surprising outcomes of the molecular screen was the identification of Tramtrack (Ttk) as a potential new regulator of muscle development. The role of Ttk has been extensively studied in the context of neuronal and sensory organ precursor development. As there are many genetic tools already available, I performed a detailed genetic and molecular characterisation of the role of this gene in mesoderm development. Ttk is a transcriptional repressor, which may account for its rather unusual myoblast fusion phenotype in loss-of-function embryos. The myoblast fusion defect can be attributed to an imbalance between the number of the two types of fusing cells. Similar to the nervous system, Ttk69 contributes to repression of the default, founder cell fate in the fusion competent myoblast (FCM) cells. Establishment of an FCM identity therefore requires active repression of the alternative, founder cell fate, concomitantly with activation of the FCM program. Importantly, the role of Ttk69 in repression of myogenic genes was confirmed by a global ChIP-on-chip analysis, which revealed an extensive overlap with the binding profiles of other mesodermal transcription factors. One of them, Lame duck, the key regulator of FCM specification, co-occupies a substantial number of Ttk69-bound CRMs, which are associated with mesodermal genes. The functional significance of this co-binding for mesodermal gene expression and myogenesis however needs to be further dissected.

Analysis of Ttk69-bound CRMs in transgenic reporter assays revealed that the regulatory inputs of Ttk69 are highly complex and diverse. Surprisingly, it also indicated that Ttk69 might act as a bimodal factor, regulating expression of its target genes both positively and negatively. The ability of sequence-specific developmental regulators to influence transcription in both directions may be a more widespread phenomenon than generally anticipated. During my Ph.D. studies, I showed that Snail, another well-characterized repressor, can enhance activity of several mesodermal enhancers both *in vitro* and *in vivo*. Elucidating the mechanisms by which Snail, and possibly Ttk69, discriminate between the positively and negatively regulated target genes and elicit their differential transcriptional response is of foremost interest, especially if the ability to stimulate transcription is a more general property of transcriptional repressors.

APPENDICES

and

REFERENCES

7 APPENDICES

7.1 Analysed RNAi lines and their phenotypes

Gene symbol	CG number	Line ID	This study — twi, 24B-Gal4: Viability	Lethality stage	Schnorrer et al. [252] — Mef2-Gal4: Viability	Lethality stage/ adult phenotype
CG7839	CG7839	12691	lethal	unknown*	lethal	early pupal
dlg1	CG1725	41134	lethal	early larval	lethal	early pupal
		41136	lethal	unknown*	Not available	
MED6	CG9473	15337	lethal	unknown*	lethal	early pupal
		15339	lethal	unknown*	lethal	early pupal
MED7	CG31390	11504	lethal	pupal	lethal	early pupal
CG8173	CG8173	35845	lethal	unknown*	lethal	late larval
		35846	lethal	unknown*	lethal	late larval
mRpL17	CG13880	24809	lethal	pupal	lethal	late larval
		24810	lethal	pupal	viable	
CG7372	CG7372	35214	lethal	pupal	lethal	late pupal
		35213	viable		viable	
CG7987	CG7987	22649	lethal	pupal	lethal	late pupal
		22647	viable		viable	flightless
msk	CG7935	38963	lethal	pupal	lethal	late pupal
CG4217	CG4217	37819	lethal	pupal	lethal	pharate/adult
CG30426	CG30426	21173	lethal	pupal	viable	
		21172	viable		viable	
MED24	CG7999	15878	lethal	pupal	viable	
MED4	CG8609	14032	lethal	pupal	viable	
sbb	CG5580	41845	lethal	pupal	viable	
MED11	CG6884	27749	lethal	early larval	Not available	
		27750	lethal	unknown*	Not available	
MED14	CG12031	15877	lethal	unknown*	Not available	
MED30	CG17183	32459	lethal	early larval	Not available	
tai	CG13109	15709	lethal	early larval	Not available	
CG12391	CG12391	43264	viable		viable	flightless
		43265	viable		viable	flightless
zf30c	CG3998	15754	viable		viable	flightless
		15753	viable		Not available	
CG10462	CG10462	31246	viable		viable	weak flyer
		31247	viable		viable	
CG10565	CG10565	38393	viable		viable	weak flyer
Adf1	CG15845	4278	viable		viable	
CG11902	CG11902	16361	viable		viable	
CG12744	CG12744	38672	viable		viable	
		38673	viable		viable	
CG1572	CG1572	23220	viable		viable	
		23221	viable		viable	
CG2199	CG2199	20839	viable		viable	
CG31365	CG31365	19236	viable		viable	
		19237	viable		viable	
CG33235	CG33235	48441	viable		viable	
		48440	viable		viable	
CG33936	CG16751	19068	viable		viable	
		19069	viable		viable	
CG4903	CG4903	21974	viable		viable	
CG7188	CG7188	37108	viable		viable	
		3235	viable		Not available	

Gene symbol	CG number	Line ID	This study		Schnorrer et al. [252]	
			twi, 24B-Gal4		Mef2-Gal4	
			Viability	Lethality stage	Viability	Lethality stage/ adult phenotype
CG8092	CG8092	28196	viable		viable	
CG9650	CG9650	23170	viable		viable	
Dip3	CG12767	31828	viable		viable	
gem	CG30011	48745	viable		viable	
		43071	viable		viable	
		43070	viable		viable	
MED16	CG5465	14916	viable		viable	
MED23	CG3695	28361	viable		viable	
		28363	viable		viable	
MED27	CG1245	13697	viable		viable	
miple	CG1221	45241	viable		viable	
		45239	viable		viable	
		45242	viable		viable	
		15396	viable		Not available	
miple2	CG18321	39026	viable		viable	
		39025	viable		viable	
mri	CG1216	17043	viable		viable	
ttk	CG1856	10855	viable		viable	
CG11247	CG11247	52116	viable		Not available	
		52117	viable		Not available	
CG14122	CG14122	51782	viable		Not available	
dmrt11E	CG15749	16744	viable		Not available	
egg	CG32105	51267	viable		Not available	
		51269	viable		Not available	
MED26	CG1793	51476	viable		Not available	
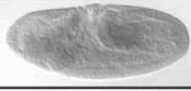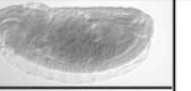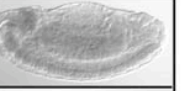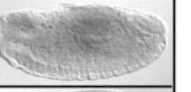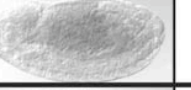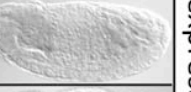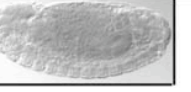
Appendix 7.1: Analysed RNAi lines and their phenotypes

Expression of mesodermal genes whose deletion by deficiency lines shows muscle phenotype, or encode subunits of the Mediator complex, was knocked down specifically in mesoderm using the panmesodermal twi, 24B-Gal4 driver and one or more gene-specific UAS-RNAi lines. Phenotype was scored as "viable" (in light grey) or "lethal" (in red) and the time point of lethality was recorded. The results were compared with the genome-wide RNAi screen performed with Mef2-Gal4 driver [252].
* the stage of lethality was not determined as the RNAi line was homozygous lethal and so adult, heterozygous offspring was always born (lethality was identified as absence of adult flies lacking the balancer chromosome)

7.2 Embryonic expression of genes encoding the Mediator complex subunits

	stage 5-7	stage 8-9	stage 10	stage 11	stage 12-16	
MED10						Ubiquitous
MED27						
MED9						
MED17						
MED22						
MED26						
MED4						Not expressed
CycC						
MED1						
lacZ						

Appendix 7.2: Embryonic expression of genes encoding the Mediator complex subunits

Expression of 25 out of 33 Mediator genes was investigated by colorimetric *in situ* hybridization in wild type embryos and imaged during five different stages of embryogenesis. The genes were grouped by their overall expression and presence in mesoderm. Probe against *lacZ* was used as a control for no expression.

7.3 Binding of Tramtrack69 in the vicinity of known downstream targets
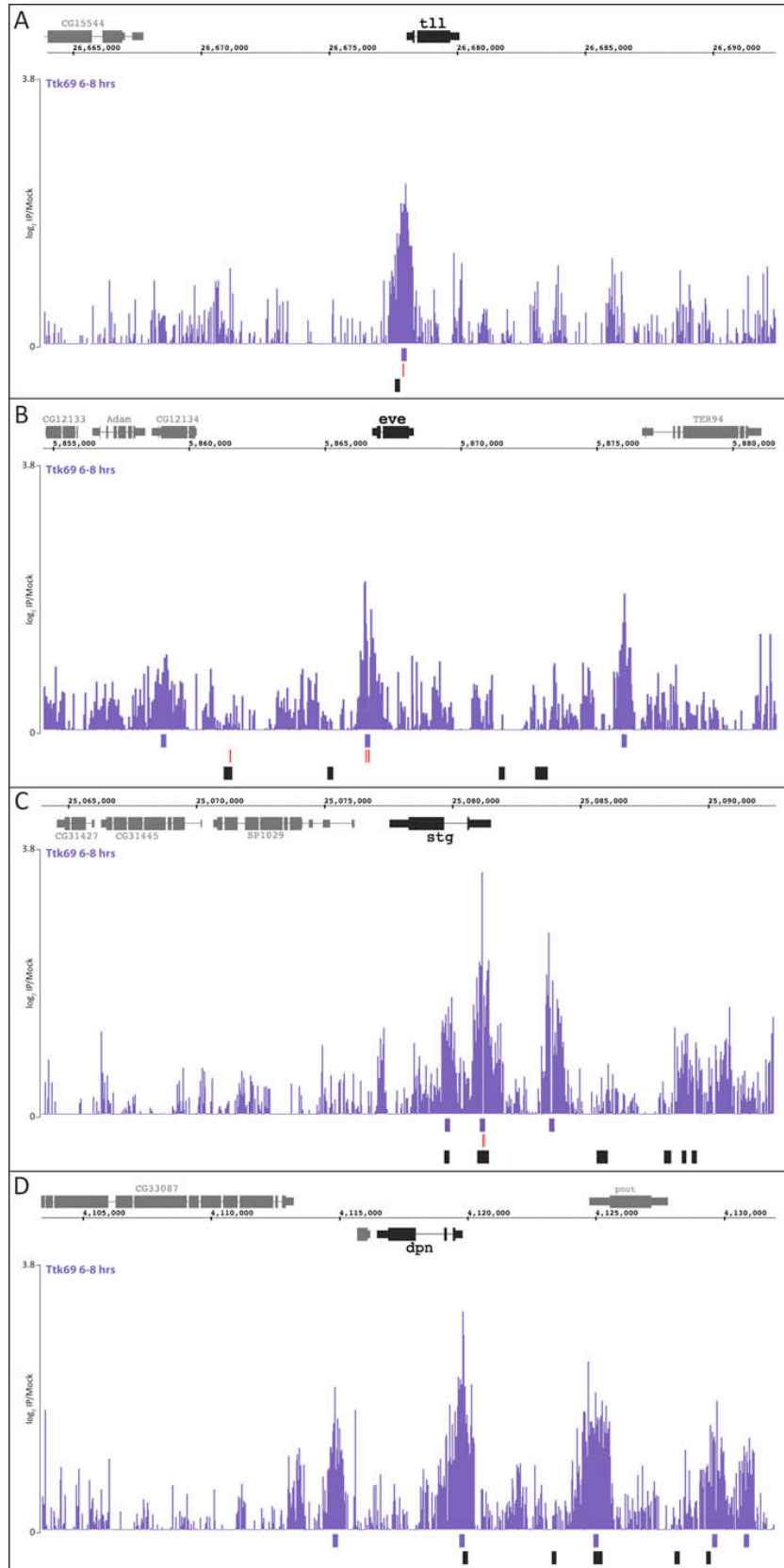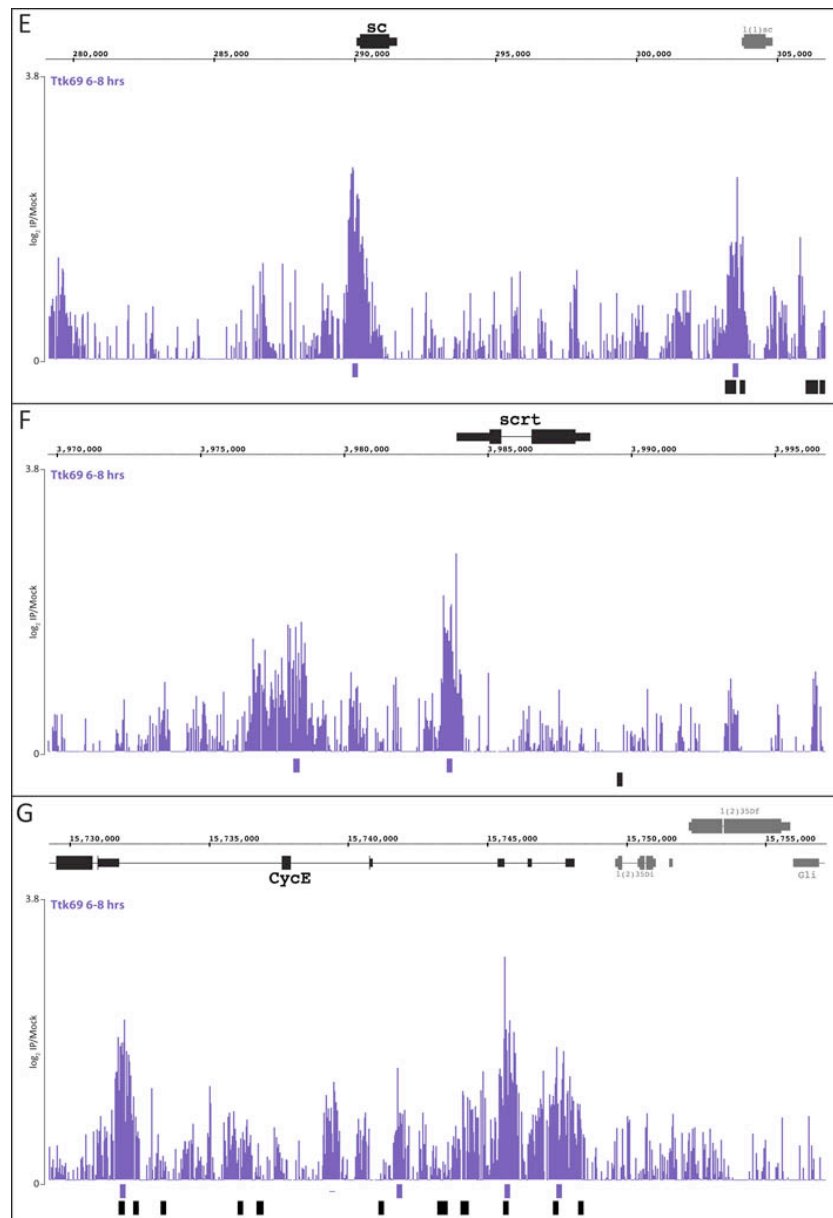
Appendix 7.3: Overview of Ttk69 CRMs in the vicinity of its known downstream targets

Normalized ChIP signal for Ttk69 is plotted in purple colour with significantly bound CRMs in the track below. In black are mesodermal CRMs derived from ChIP-on-chip experiments with 5 key mesodermal TFs over 5 time points [196]. Ttk69 has been previously shown to directly regulate *tll* (A), *eve* (B), and *stg* (C) and either directly or indirectly *dpn* (D), *sc* (E), *scrt* (F), and *CycE* (G) in black. Red lines indicate previously experimentally verified Ttk69 binding sites in the promoter regions of *tll* (A), *eve* (B), and *stg* (C).

7.4 Sequences of cloned enhancers

7.4.1 Sequences of Tramtrack69-bound enhancers

Underlined are all predicted Ttk69 binding sites and in red are highlighted the base pairs that were mutated.

JumuA (671 bp)

GGTTGTGTTTATTTTGCTAGCCACCACTGTATGTACTCTCAGAAAATGTTT
ACATGTTGCTGATACTTTTTT<u>TGTTGTACTGGTT</u>CTGGTGTTGCAACATGA
GTTATTTTCGTTGTTTTGGTTGTTGTTGGGCGTGCT<u>ACAGGGAACTCAG</u>AA
<u>ATATCCAAAC</u>CCCCGACTTTTTCCCTTTTCCTATCCTTTGGCCGTCTGCTTC
GGTTTTGTTTATTTTTTTTCCTCTTACAGCCTGCCTGCCTTGGCGCGCAAAA
CTGGTTTTGAACTCGGCTCTGTTGCTCCTGTGGTCCCTCATTTGTTGTTGTT
GCTATTTTGTGTATTGT<u>TGTTGTCCGTGC</u>CTCATGTAGAAAATGAAAACTT
ATACGATCCCCACAAATATCAGC<u>ACAG</u>GGA<u>TATCG</u>GATTCGTATTCGGAC
CCAGATTCGAGCGACTTGTGTTACAGATTTAATCATTATAATATATGTACA
TAT<u>GTATGTTCTTAT</u>ATACAT<u>ACATATACTTGT</u>ATGTATGCGTGCCTATAT
GTGTGCAAAGGAATTGTAACCTATGGAATTCTTAGGAAAATAAATG<u>CGCG</u>
<u>GGAAAACG</u>ATCAGTACGTGTTTCTGGATTTGTTTATGTTTTGATTGTGCTA
AGCGTAGCGGATAAAATGAAATA<u>GTATATT</u>CTTGCATAGCGGGCAAATAC
TTGATTCC

JumuA-Ttkmut (671 bp)

GGTTGTGTTTATTTTGCTAGCCACCACTGTATGTACTCTCAGAAAATGTTT
ACATGTTGCTGATACTTTTTT<u>TGTTGTACTGGTT</u>AAAGTGTTGCAACATGA
GTTATTTTCGTTGTTTTGGTTGTTGTTGGGCGTGCT<u>ACAGGGAACTCAG</u>AA
<u>ATATCCAAAC</u>CCCCGACTTTTTCCCTTTTCCTATCCTTTGGCCGTCTGCTTC
GGTTTTGTTTATTTTTTTTCCTCTTACAGCCTGCCTGCCTTGGCGCGCAAAA
CTGGTTTTGAACTCGGCTCTGTTGCTCCTGTGGTCCCTCATTTGTTGTTGTT
GCTATTTTGTGTATTGT<u>TGTTGTCCGTGC</u>CTCATGTAGAAAATGAAAACTT
ATACGATCCCCACAAATATCAGC<u>ACAG</u>AAC<u>TATCG</u>GATTCGTATTCGGAC
CCAGATTCGAGCGACTTGTGTTACAGATTTAATCATTATAATATATGTACA
TAT<u>GTATGTTCTTAT</u>ATACAT<u>ACATATACTTGT</u>ATGTATGCGTGCCTATAT
GTGTGCAAAGGAATTGTAACCTATGGAATTCTTAGGAAAATAAATG<u>CGCG</u>
<u>GGAAAACG</u>ATCAGTACGTGTTTCTGGATTTGTTTATGTTTTGATTGTGCTA
AGCGTAGCGGATAAAATGAAATA<u>GTATATT</u>AAAGCATAGCGGGCAAATA
CTTGATTCC

JumuB (857 bp)

AACCAACTAATTGCCCCAAAAAACCCTCACAGTCTCAGCAAATGGACTTC
CAGGCTAAAAATAAACATTTGGCAAATCTGATTTCTTGAAACACATGCAG
TTAATGAATGAAATAACCCGGAGAATCTGCCGATTTACTTCCCGCTAAAC
CGAGAGATTTGAAAACGATTTACGACCCTGGATCATCCTGTGTGTTTAAG
GATGTTACTTTGCGTTGGGAATCGGTCGGCAATTCAGGCAAGCGACTTTT
ATTGCACCATAAAGTAATTTACAAATGGGCGCTCATGACCTCAGGGAAGA
TTCGTCCAGAAAGGTCCACGGAAATCGTTGGTTTTACCTGAGTGAGGGAT
TCCATTGCATTCCATTCCTTACCATCCTCCTCCCATCATTATTTACCATCTC
TTCTTGTGTTTTCTCTGTGTGTCAGCAGAGTGATTTCCTTCGGAGGTGAAA
GGCGGAGAAAGATGGAAAATATGTGCAACAAAAGAAACGAAGCAGATGC
AGGAAATGCAGGAGAACACTTACAACCATCGCCAGTGACGAATCCTGCCG
AGAAAGGACGAAAATACCCTCGTCGGCAGCTCCTTGGTTCGTTGCACAAC
TTGAAAAGGCCAACTGTAGGCAGCTGCAACAGAGCAGCAAAAACTTTTTA
TGGCCTACGGTGGGTTGTATGGGAGGGGGTGCTGAAATACGGGAGTGGTG
GCGGTGTCTGTAAGCCTCGTCGAAGTTTTACTGACAACTTTTGCATTTTCC
TTTAATCCCCCCGATAACCAAAAGCAGCTTGCTGTGAACTTTGCTCGCCTG
TTTGATGCATTTTTATTTCGGCACATTTTTTCGAGAATTTTTTGGGGGGGCA
G

JumuB-Ttkmut (857 bp)

AACCAACTAATTGCCCCAAAAAACCCTCACAGTCTCAGCAAATGGACTTC
CAGGCTAAAAATAAACATTTGGCAAATCTGATTTCTTGAAACACATGCAG
TTAATGAATGAAATAACCCGGAGAATCTACCGAATTAAAACCCGCTAAAC
CGAGAGATTTGAAAACGATTTACGACCCTCAATCATCCTGTGTGTTTAAG
GATGTTACTTTGCGTTGGGAATCGGTCGGCAATTCAGGCAAGCGACTTTT
ATTGCACCATAAAGTAATTTACAAATGGGCGCTCATGACCTCAGGGAAGA
TTCGTCCAGAACGAACAACAGCAATCGTTGGTTTTACCTGAGTGAGGGAT
TCCATTGCATTCCATTCCTTACCATCCTCCTCCCATCATTATTTACCATCTC
TTCTTGTGTTTTCTCTGTGTGTCAGCAGAGTGATTTCCTTCGGAGGTGAAA
GGCGGAGAAAGATGGAAAATATGTGCAACAAAAGAAACGAAGCAGATGC
AGGAAATGCAGGAGAACACTTACAACCATCGCCAGTGACGAAAAATGCC
GAGAAAGGACGAAAATACCCTCGTCGGTCACAAATTGGTTCGTTGCACAA
CTTGAAAAGGCCAACTGTAGGCAGCTGCAACAGAGCAGCAAAAACTTTTT
ATGGCCTACGGTGGGTTGTATGGGAGGGGGTGCTGAAATACGGGAGTGGT
GGCGGTGTCTGTAAGCCTCGTCGAAGTTTTACTGACAACTTTTGCATTTTC
CTTTAATCCCCCCGATAACCAAAAGCAGCTTGCTGTGAACTTTGCTCGCCT
GTTTGATGCATTTTTATTTCGGCACATTTTTTCGAGAATTTTTTGGGGGGGC
AG

rib (426 bp)

AGACTGGGACACAGCTATGCCCCTCTCCCTCGCACGCACTGCATCCAATG
GCAGCAACCATATATGCAGCTCGCACACACATGCACGCACGCACACGCAC
ACAAGCAGTCGCAATCGCAATCTCGCTCACGCACAAAACTGCAACTGCGC
GAAAGAGAGGGCGCGAGGGTGGATTCGTAGCACAGTTTTCCTGTTTCCCC
TTCGCATTCGAACAGTTTTCTCGCCGCTCTCTCGCTCTCACCCTCTGTGCAT
CGCTCCATCCCGCTCTGGCAGCGCTCGGGAAATTTGGTCTTCAATGACTTG
CTGCACACGAAAACCGTGGCAGTTTTATTCAAGTTTGAGCGCTCGAATCT
GCAACAACGTGCACGCAAAACTCTCAAAATTTTCTCAAAAACATCAAAAT
CACAAGGCGAGCGGAAAACTCAAG

rib-Ttkmut (426 bp)

AGACTGGGACACAGCTATGCCCCTCTCCCTCGCACGCACTGCATCCAATG
GCAGCAACCATATATGCAGCTCGCACACACATGCACGCACGCACACGCAC
ACAAGCAGTCGCAATCGCAATCTCGCTCACGCACAAAACTGCAACTGCGC
GAAAGAGAGGGCGCGAGGGTGGATTCGTAGCACAGTTTAAATGTTTCCCC
TTCGCATTCGAACAGTTTTCTCGCCGCTCTCTCGCTCTCACCCTCTGTGCAT
CGCTCCATCCCGCTCTGGCAGCGCTCGAACAATTTGGTCTTCACACACTTG
CTGCACACGAAAACCGTGGCAGTTTTATTCAAGTTTGAGCGCTCGAATCT
GCAACAACGTGCACGCAAAACTCTCAAAATTTTCTCAAAAACATCAAAAT
CACAAGGCGAGCGGAAAACTCAAG

7.4.2 Sequences of Snail-bound enhancers

Underlined are putative Snail binding sites and in red are highlighted the base pairs that were mutated (shown only for the enhancers that had a mutant version generated).

Mdr49 early mesoderm (382 bp) [56]

GCAACAAAGTCGATCGTATAACTAATCGGGATTCCCGCCTGGCAACCCGC
CTGTTCCTCGCACTCGTCCAGGCGGCGGGAAATACTGGTTCCTAGTCGCTG
GTCCTGTAACCAAAACAGAGCCATAAAGACGATGCAGATTAGATTGATTA
GACAGTTATGTTGTACGTACATACCATCGATTCGGCGTAAGTTTGTTATCA
GTCTCTCGTAACCGGTATTTGTTTACTGACCCAGCGTGTCCATAGATGCAG
TGGAGTCGGGTCTTTTTTGGCCAAAACAAAAACGGACTAAGCGTAATCCA
TTTGCGCGCATGGTTTTCCCCGCGCAAAAGATTTTAAGCTGAATTCG
CCAGATCGGGGGGACTAACCGAGAGGAAATCTG

WntD-lacZ (523 bp) [56]

ATAGCCTGCAAATCCCAAGCCAGGGCGCCCTCCTGGGGCCGGCCCGTGGG
AATTTCGGGCCTGCTCAAAAAACCGGAAATTTGCCGTTTTCCACTTGGAA
ATTTTGCATGGGCAGGGGGTAGGAACTCCCGGCAATGGACGGGTACAAA
AACCCACTGGCAGCCCGAGACGCAATTGCGGAGCAGCCCAGTTTCCTGGT
TGACTACCTGCTCTCGTCCTGCGCCGGCGGAGGTGAAGGATCCGCCTTGCT
GCGAGCAAGTTTCCCACGCTTAGGCAGGTAGAGCCGTAAACGGCACCCGA
CGTGCTCATGAATGAAGCCCAGTCGAGTCCATTCAATACGGCCGGATTTT
CCCGGACTCACACTGCACAACATCAATGCCCGATACGGGGACGGGTTTGT
TTGGGTTTTGGACTGGTCAAGCCAATTATATAACAAACATATGACCAAC
AGTATATACACGTATAATCTGGGAAATTAATTGTCCTTGTGGTGAGCCGG
CGGGAACTGCCAGTATTGTTGC

Mef2 I-D[L] (480 bp) [236]

CTGTAAAAATCACGCATAACCGATACCAAGTGCAGCGCGACCAAGTTACC
CGTTTTATGGCCCTCGCCTCTCGGCGGCGCCACAATGTTGCATTTTGAGAA
CCGAACCCAGCAGCAAACGATCCATGTGTGTATATCCGAAGATCCCTTTC
TCACTTCCACTCCCATTACCATTCCCATTTCCATTTCCATTTCCATTCTGAT
TCCCGTTTGCAGTGTCCTTGTGACTTTTAACGCTTCCACTGGGTGCGACAA
CTGCCAACTGCAAACAGTCGTCGACAGTCTTAATTCAATAAACGCCGCCC
GGCGATTTGCGCATGACCATGTACCCCGATGCTGTGCGCCGTACGGTTGA
TGCTG<u>CATGTTGCATGCACTCAACACATGTGCAACATGC</u>GGCATCTGCGG
CAGTAGCAGCAGCAGCAGCATCGGCCTCAATGTGCGTTGAGTCCGA
CTTTAACTGCTTGGCAGGGGTTTCTTCAGG

Mef2 I-D[L] Sna123 (480 bp)

CTGTAAAAATCACGCATAACCGATACCAAGTGCAGCGCGACCAAGTTACC
CGTTTTATGGCCCTCGCCTCTCGGCGGCGCCACAATGTTGCATTTTGAGAA
CCGAACCCAGCAGCAAACGATCCATGTGTGTATATCCGAAGATCCCTTTC
TCACTTCCACTCCCATTACCATTCCCATTTCCATTTCCATTTCCATTCTGAT
TCCCGTTTGCAGTGTCCTTGTGACTTTTAACGCTTCCACTGGGTGCGACAA
CTGCCAACTGCAAACAGTCGTCGACAGTCTTAATTCAATAAACGCCGCCC
GGCGATTTGCGCATGACCATGTACCCCGATGCTGTGCGCCGTACGGTTGA
TGCTG<u>AGTATTGCATGCACTCATCACATGTGCAATACT</u>CGGCATCTGCGGC
AGTAGCAGCAGCAGCAGCATCGGCCTCAATGTGCGTTGAGTCCGA
CTTTAACTGCTTGGCAGGGGTTTCTTCAGG

Tin B-374 (403 bp) [60]

TCAAGCGTTGAGCGTTGAGCTCGAGGCTTTGACAAATCATCGTTATTTGTA
CAAGAGGCCGAAGACAAAGACGAGGAGAGACAGTGTTAACAATAAGAAC
AACAATGTCCGAGGCAGGCAGTCGGGAGTCACGGCGATCGCTGGTCCGCC
CGATCCCTTCTGGGCTGGTCAACATGTGTGATTCGCATGTGTGGACCGCCG
CACAGGGGCGTCCTTAATTGCCTGATGAGCCATGAAATGATGTCACCATG
GATCCTGTCGCGCCAGCAGGAAGTGGGCAAAAAGTCCTCGTCCCAGCTCC
CCGGGCTGTGTCCTCCGTGATGCAACATATGGCGGCCATATACGAGACTT
TGTATTGCCTTGTTTCTAACCTTGCGCTTTTCCATTTTCCTGCTTTCCCGCA

Tin B-374 Sna4 (403 bp)

TCAAGCGTTGAGCGTTGAGCTCGAGGCTTTGACAAATCATCGTTATTTGTA
CAAGAGGCCGAAGACAAAGACGAGGAGAGACAGTGTTAACAATAAGAAC
AACAATGTCCGAGGCAGGCAGTCGGGAGTCACGGCGATCGCTGGTCCGCC
CGATCCCTTCTGGGCTGGTCAACATGTGTGATTCGCATGTGTGGACCGCCG
CACAGGGGCGTCCTTAATTGCCTGATGAGCCATGAAATGATGTCACCATG
GATCCTGTCGCGCCAGCAGGAAGTGGGCAAAAAGTCCTCGTCCCAGCTCC
CCGGGCTGTGTCCTCCGTGATGCAACATATGGCGGCCATATACGAGACTT
TGTATTGCCTTGTTTCTAACTTCTCGCTTTTCCATTTTCCTGCTTTCCCGCA

REFERENCES

1.   Orphanides, G. and Reinberg, D. *A unified theory of gene expression.* Cell, 2002. **108**(4): p. 439-51.

2.   Wray, G.A. et al. *The evolution of transcriptional regulation in eukaryotes.* Mol Biol Evol, 2003. **20**(9): p. 1377-419.

3.   Archambault, J. and Friesen, J.D. *Genetics of eukaryotic RNA polymerases I, II, and III.* Microbiol Rev, 1993. **57**(3): p. 703-24.

4.   Atchison, M.L. *Enhancers: mechanisms of action and cell specificity.* Annu Rev Cell Biol, 1988. **4**: p. 127-53.

5.   Firulli, A.B. and Olson, E.N. *Modular regulation of muscle gene transcription: a mechanism for muscle cell diversity.* Trends Genet, 1997. **13**(9): p. 364-9.

6.   Hartenstein, V. *Atlas of Drosophila development.* 1993, Plainview, N.Y.: Cold Spring Harbor Laboratory Press. v, 57 p.

7.   Gray, S. and Levine, M. *Transcriptional repression in development.* Curr Opin Cell Biol, 1996. **8**(3): p. 358-64.

8.   Furlong, E.E. *Integrating transcriptional and signalling networks during muscle development.* Curr Opin Genet Dev, 2004. **14**(4): p. 343-50.

9.   Qi, D. et al. *Drosophila Ebi mediates Snail-dependent transcriptional repression through HDAC3-induced histone deacetylation.* EMBO J, 2008. **27**(6): p. 898-909.

10.   Stanojevic, D., Small, S. and Levine, M. *Regulation of a segmentation stripe by overlapping activators and repressors in the Drosophila embryo.* Science, 1991. **254**(5036): p. 1385-7.

11.   Arnosti, D.N. et al. *The eve stripe 2 enhancer employs multiple modes of transcriptional synergy.* Development, 1996. **122**(1): p. 205-14.

12.   Gray, S., Szymanski, P. and Levine, M. *Short-range repression permits multiple enhancers to function autonomously within a complex promoter.* Genes Dev, 1994. **8**(15): p. 1829-38.

13.   Nibu, Y., Zhang, H. and Levine, M. *Interaction of short-range repressors with Drosophila CtBP in the embryo.* Science, 1998. **280**(5360): p. 101-4.

14. Cai, H.N., Arnosti, D.N. and Levine, M. *Long-range repression in the Drosophila embryo.* Proc Natl Acad Sci U S A, 1996. **93**(18): p. 9309-14.

15. Kelleher, R.J., 3rd, Flanagan, P.M. and Kornberg, R.D. *A novel mediator between activator proteins and the RNA polymerase II transcription apparatus.* Cell, 1990. **61**(7): p. 1209-15.

16. Boube, M. et al. *Evidence for a mediator of RNA polymerase II transcriptional regulation conserved from yeast to man.* Cell, 2002. **110**(2): p. 143-51.

17. Gu, J.Y. et al. *Novel Mediator proteins of the small Mediator complex in Drosophila SL2 cells.* J Biol Chem, 2002. **277**(30): p. 27154-61.

18. Loncle, N. et al. *Distinct roles for Mediator Cdk8 module subunits in Drosophila development.* EMBO J, 2007. **26**(4): p. 1045-54.

19. Bosveld, F., van Hoek, S. and Sibon, O.C. *Establishment of cell fate during early Drosophila embryogenesis requires transcriptional Mediator subunit dMED31.* Dev Biol, 2008. **313**(2): p. 802-13.

20. Ashburner, M. et al. *Gene ontology: tool for the unification of biology. The Gene Ontology Consortium.* Nat Genet, 2000. **25**(1): p. 25-9.

21. Guglielmi, B. et al. *A high resolution protein interaction map of the yeast Mediator complex.* Nucleic Acids Res, 2004. **32**(18): p. 5379-91.

22. Dotson, M.R. et al. *Structural organization of yeast and mammalian mediator complexes.* Proc Natl Acad Sci U S A, 2000. **97**(26): p. 14307-10.

23. Asturias, F.J. et al. *Conserved structures of mediator and RNA polymerase II holoenzyme.* Science, 1999. **283**(5404): p. 985-7.

24. Bjorklund, S. and Gustafsson, C.M. *The yeast Mediator complex and its regulation.* Trends Biochem Sci, 2005. **30**(5): p. 240-4.

25. Collins, S.R. et al. *Functional dissection of protein complexes involved in yeast chromosome biology using a genetic interaction map.* Nature, 2007. **446**(7137): p. 806-10.

26. Rickert, P. et al. *Cyclin C/CDK8 is a novel CTD kinase associated with RNA polymerase II.* Oncogene, 1996. **12**(12): p. 2631-40.

27. Balciunas, D. and Ronne, H. *Three subunits of the RNA polymerase II mediator complex are involved in glucose repression.* Nucleic Acids Res, 1995. **23**(21): p. 4421-5.

28. Elmlund, H. et al. *The cyclin-dependent kinase 8 module sterically blocks Mediator interactions with RNA polymerase II.* Proc Natl Acad Sci U S A, 2006. **103**(43): p. 15788-93.

29. Gaytan de Ayala Alonso, A. et al. *A genetic screen identifies novel polycomb group genes in Drosophila.* Genetics, 2007. **176**(4): p. 2099-108.

30. Park, J.M. et al. *Drosophila Mediator complex is broadly utilized by diverse gene-specific transcription factors at different types of core promoters.* Mol Cell Biol, 2001. **21**(7): p. 2312-23.

31. Phatnani, H.P. and Greenleaf, A.L. *Phosphorylation and functions of the RNA polymerase II CTD.* Genes Dev, 2006. **20**(21): p. 2922-36.

32. Gobert, V. et al. *A genome-wide RNA interference screen identifies a differential role of the Mediator CDK8 module subunits for GATA/RUNX-activated transcription in Drosophila.* Mol Cell Biol.

33. Park, J.M. et al. *Mediator, not holoenzyme, is directly recruited to the heat shock promoter by HSF upon heat shock.* Mol Cell, 2001. **8**(1): p. 9-19.

34. Kuras, L., Borggrefe, T. and Kornberg, R.D. *Association of the Mediator complex with enhancers of active genes.* Proc Natl Acad Sci U S A, 2003. **100**(24): p. 13887-91.

35. Zhu, X. et al. *Genome-wide occupancy profile of mediator and the Srb8-11 module reveals interactions with coding regions.* Mol Cell, 2006. **22**(2): p. 169-78.

36. Fan, X., Chou, D.M. and Struhl, K. *Activator-specific recruitment of Mediator in vivo.* Nat Struct Mol Biol, 2006. **13**(2): p. 117-20.

37. Yudkovsky, N., Ranish, J.A. and Hahn, S. *A transcription reinitiation intermediate that is stabilized by activator.* Nature, 2000. **408**(6809): p. 225-9.

38. Lim, J. et al. *Drosophila TRAP230/240 are essential coactivators for Atonal in retinal neurogenesis.* Dev Biol, 2007. **308**(2): p. 322-30.

39. Gim, B.S. et al. *Drosophila Med6 is required for elevated expression of a large but distinct set of developmentally regulated genes.* Mol Cell Biol, 2001. **21**(15): p. 5242-55.

40. Janody, F. et al. *Two subunits of the Drosophila mediator complex act together to control cell affinity.* Development, 2003. **130**(16): p. 3691-701.

41. Carrera, I. et al. *Pygopus activates Wingless target gene transcription through the mediator complex subunits Med12 and Med13.* Proc Natl Acad Sci U S A, 2008. **105**(18): p. 6644-9.

42. Treisman, J. *Drosophila homologues of the transcriptional coactivation complex subunits TRAP240 and TRAP230 are required for identical processes in eye-antennal disc development.* Development, 2001. **128**(4): p. 603-15.

43. Boube, M. et al. *Drosophila homologs of transcriptional mediator complex subunits are required for adult cell and segment identity specification.* Genes Dev, 2000. **14**(22): p. 2906-17.

44. Terriente-Felix, A., Lopez-Varea, A. and de Celis, J.F. *Identification of Genes Affecting Wing Patterning Through a Loss-of-Function Mutagenesis Screen and Characterization of med15 Function During Wing Development.* Genetics.

45. Kim, T.W. et al. *MED16 and MED23 of Mediator are coactivators of lipopolysaccharide- and heat-shock-induced transcriptional activators.* Proc Natl Acad Sci U S A, 2004. **101**(33): p. 12153-8.

46. Park, J.M. et al. *Signal-induced transcriptional activation by Dif requires the dTRAP80 mediator module.* Mol Cell Biol, 2003. **23**(4): p. 1358-67.

47. Wang, L. et al. *A genetic screen identifies new regulators of steroid-triggered programmed cell death in Drosophila.* Genetics, 2008. **180**(1): p. 269-81.

48. Deato, M.D. and Tjian, R. *Switching of the core transcription machinery during myogenesis.* Genes Dev, 2007. **21**(17): p. 2137-49.

49. Deato, M.D. and Tjian, R. *An unexpected role of TAFs and TRFs in skeletal muscle differentiation: switching core promoter complexes.* Cold Spring Harb Symp Quant Biol, 2008. **73**: p. 217-25.

50. Pham, A.D., Muller, S. and Sauer, F. *Mesoderm-determining transcription in Drosophila is alleviated by mutations in TAF(II)60 and TAF(II)110.* Mech Dev, 1999. **84**(1-2): p. 3-16.

51. Beyer, K.S. et al. *Mediator subunit MED28 (Magicin) is a repressor of smooth muscle cell differentiation.* J Biol Chem, 2007. **282**(44): p. 32152-7.

52. Parrish, J.Z. et al. *Genome-wide analyses identify transcription factors required for proper morphogenesis of Drosophila sensory neuron dendrites.* Genes Dev, 2006. **20**(7): p. 820-35.

53. Wei, C.L. et al. *A global map of p53 transcription-factor binding sites in the human genome.* Cell, 2006. **124**(1): p. 207-19.

54. Bonn, S. and Furlong, E.E. *cis-Regulatory networks during development: a view of Drosophila.* Curr Opin Genet Dev, 2008. **18**(6): p. 513-20.

55. Sandmann, T. et al. *A core transcriptional network for early mesoderm development in Drosophila melanogaster.* Genes Dev, 2007. **21**(4): p. 436-49.

56. Zeitlinger, J. et al. *Whole-genome ChIP-chip analysis of Dorsal, Twist, and Snail suggests integration of diverse patterning processes in the Drosophila embryo.* Genes Dev, 2007. **21**(4): p. 385-90.

57. Alon, U. *Network motifs: theory and experimental approaches.* Nat Rev Genet, 2007. **8**(6): p. 450-61.

58. Davidson, E.H. *The regulatory genome : gene regulatory networks in development and evolution.* 2006, Burlington, MA ; San Diego: Academic. xi, 289 p.

59. Cripps, R.M. et al. *The myogenic regulatory gene Mef2 is a direct target for transcriptional activation by Twist during Drosophila myogenesis.* Genes Dev, 1998. **12**(3): p. 422-34.

60. Yin, Z., Xu, X.L. and Frasch, M. *Regulation of the twist target gene tinman by modular cis-regulatory elements during early mesoderm development.* Development, 1997. **124**(24): p. 4971-82.

61. Zaffran, S. et al. *biniou (FoxF), a central component in a regulatory network controlling visceral mesoderm development and midgut morphogenesis in Drosophila.* Genes Dev, 2001. **15**(21): p. 2900-15.

62. Sandmann, T. et al. *A temporal map of transcription factor activity: mef2 directly regulates target genes at all stages of muscle development.* Dev Cell, 2006. **10**(6): p. 797-807.

63. Roth, S. et al. *cactus, a maternal gene required for proper formation of the dorsoventral morphogen gradient in Drosophila embryos.* Development, 1991. **112**(2): p. 371-88.

64. Kidd, S. *Characterization of the Drosophila cactus locus and analysis of interactions between cactus and dorsal proteins.* Cell, 1992. **71**(4): p. 623-35.

65. Morisato, D. and Anderson, K.V. *The spatzle gene encodes a component of the extracellular signaling pathway establishing the dorsal-ventral pattern of the Drosophila embryo.* Cell, 1994. **76**(4): p. 677-688.

66. Moussian, B. and Roth, S. *Dorsoventral axis formation in the Drosophila embryo--shaping and transducing a morphogen gradient.* Curr Biol, 2005. **15**(21): p. R887-99.

67. Ip, Y.T. et al. *The dorsal gradient morphogen regulates stripes of rhomboid expression in the presumptive neuroectoderm of the Drosophila embryo.* Genes Dev, 1992. **6**(9): p. 1728-39.

68. Ip, Y.T. et al. *dorsal-twist interactions establish snail expression in the presumptive mesoderm of the Drosophila embryo.* Genes & Development, 1992. **6**(19): p. 1518-1530.

69. Chopra, V.S. and Levine, M. *Combinatorial patterning mechanisms in the Drosophila embryo.* Brief Funct Genomic Proteomic, 2009. **8**(4): p. 243-9.

70. Thisse, C. et al. *Sequence-specific transactivation of the Drosophila twist gene by the dorsal gene product.* Cell, 1991. **65**: p. 1191-1201.

71. Francois, V. et al. *Dorsal-ventral patterning of the Drosophila embryo depends on a putative negative growth factor encoded by the short gastrulation gene.* Genes Dev, 1994. **8**(21): p. 2602-16.

72. Stathopoulos, A. et al. *pyramus and thisbe: FGF genes that pattern the mesoderm of Drosophila embryos.* Genes Dev, 2004. **18**(6): p. 687-99.

73. Mellerick, D.M. and Nirenberg, M. *Dorsal-ventral patterning genes restrict NK-2 homeobox gene expression to the ventral half of the central nervous system of Drosophila embryos.* Dev Biol, 1995. **171**(2): p. 306-16.

74. Morel, V. and Schweisguth, F. *Repression by suppressor of hairless and activation by Notch are required to define a single row of single-minded expressing cells in the Drosophila embryo.* Genes Dev, 2000. **14**(3): p. 377-88.

75. Jiang, J. et al. *Individual dorsal morphogen binding sites mediate activation and repression in the Drosophila embryo.* EMBO J, 1992. **11**(8): p. 3147-54.

76. Dubnicoff, T. et al. *Conversion of dorsal from an activator to a repressor by the global corepressor Groucho.* Genes Dev, 1997. **11**(22): p. 2952-7.

77. Huang, J.D. et al. *The interplay between multiple enhancer and silencer elements defines the pattern of decapentaplegic expression.* Genes Dev, 1993. **7**(4): p. 694-704.

78. Ip, Y.T. et al. *The dorsal morphogen is a sequence-specific DNA-binding protein that interacts with a long-range repression element in Drosophila.* Cell, 1991. **64**(2): p. 439-46.

79. Campos-Ortega, J.A. and Hartenstein, V. *The embryonic development of Drosophila melanogaster*. 1985: Springer-Verlag.

80. Thisse, B., el Messal, M. and Perrin-Schmitt, F. *The twist gene: isolation of a Drosophila zygotic gene necessary for the establishment of dorsoventral pattern.* Nucleic Acids Res, 1987. **15**(8): p. 3439-53.

81. Baylies, M.K. and Bate, M. *twist: a myogenic switch in Drosophila.* Science, 1996. **272**(5267): p. 1481-4.

82. Leptin, M. and Grunewald, B. *Cell shape changes during gastrulation in Drosophila.* Development, 1990. **110**(1): p. 73-84.

83. Leptin, M. *twist and snail as positive and negative regulators during Drosophila mesoderm development.* Genes Dev, 1991. **5**(9): p. 1568-76.

84. Alberga, A. et al. *The snail gene required for mesoderm formation in Drosophila is expressed dynamically in derivatives of all three germ layers.* Development, 1991. **111**: p. 983-992.

85. Ray, R.P. et al. *The control of cell fate along the dorsal-ventral axis of the Drosophila embryo.* Development, 1991. **113**(1): p. 35-54.

86. Lilly, B. et al. *D-MEF2: a MADS box transcription factor expressed in differentiating mesoderm and muscle cell lineages during Drosophila embryogenesis.* Proc Natl Acad Sci U S A, 1994. **91**(12): p. 5662-6.

87. Bodmer, R., Jan, L.Y. and Jan, Y.N. *A new homeobox-containing gene, msh-2, is transiently expressed early during mesoderm formation of Drosophila.* Development, 1990. **110**(3): p. 661-9.

88. Shishido, E. et al. *Two FGF-receptor homologues of Drosophila: one is expressed in mesodermal primordium in early embryos.* Development, 1993. **117**(2): p. 751-61.

89. Costa, M., Wilson, E.T. and Wieschaus, E. *A putative cell signal encoded by the folded gastrulation gene coordinates cell shape changes during Drosophila gastrulation.* Cell, 1994. **76**(6): p. 1075-89.

90. Hemavathy, K., Meng, X. and Ip, Y.T. *Differential regulation of gastrulation and neuroectodermal gene expression by Snail in the Drosophila embryo.* Development, 1997. **124**(19): p. 3683-91.

91. Casal, J. and Leptin, M. *Identification of novel genes in Drosophila reveals the complex regulation of early gene activity in the mesoderm.* Proc Natl Acad Sci U S A, 1996. **93**(19): p. 10327-32.

92. Ip, Y.T., Maggert, K. and Levine, M. *Uncoupling gastrulation and mesoderm differentiation in the Drosophila embryo.* EMBO J, 1994. **13**(24): p. 5826-34.

93. Mauhin, V. et al. *Definition of the DNA-binding site repertoire for the Drosophila transcription factor SNAIL.* Nucleic Acids Res, 1993. **21**(17): p. 3951-7.

94. Reece-Hoyes, J.S. et al. *The C. elegans Snail homolog CES-1 can activate gene expression in vivo and share targets with bHLH transcription factors.* Nucleic Acids Res, 2009. **37**(11): p. 3689-98.

95. Hu, C.T. et al. *Snail associates with EGR-1 and SP-1 to upregulate transcriptional activation of p15INK4b.* FEBS J. **277**(5): p. 1202-18.

96. Technau, G.M. *A single cell approach to problems of cell lineage and commitment during embryogenesis of Drosophila melanogaster.* Development, 1987. **100**(1): p. 1-12.

97. Maggert, K., Levine, M. and Frasch, M. *The somatic-visceral subdivision of the embryonic mesoderm is initiated by dorsal gradient thresholds in Drosophila.* Development, 1995. **121**(7): p. 2107-16.

98. Bodmer, R. *The gene tinman is required for specification of the heart and visceral muscles in Drosophila.* Development, 1993. **118**(3): p. 719-29.

99. Azpiazu, N. et al. *Segmentation and specification of the Drosophila mesoderm.* Genes Dev, 1996. **10**(24): p. 3183-94.

100. Lee, H.H. and Frasch, M. *Wingless effects mesoderm patterning and ectoderm segmentation events via induction of its downstream target sloppy paired.* Development, 2000. **127**(24): p. 5497-508.

101. Lee, H.H. and Frasch, M. *Nuclear integration of positive Dpp signals, antagonistic Wg inputs and mesodermal competence factors during Drosophila visceral mesoderm induction.* Development, 2005. **132**(6): p. 1429-42.

102. Park, M. et al. *The wingless signaling pathway is directly involved in Drosophila heart development.* Dev Biol, 1996. **177**(1): p. 104-16.

103. Sam, S., Leise, W. and Hoshizaki, D.K. *The serpent gene is necessary for progression through the early stages of fat-body development.* Mech Dev, 1996. **60**(2): p. 197-205.

104. Colombani, J. et al. *A nutrient sensor mechanism controls Drosophila growth.* Cell, 2003. **114**(6): p. 739-49.

105. Gutierrez, E. et al. *Specialized hepatocyte-like cells regulate Drosophila lipid metabolism.* Nature, 2007. **445**(7125): p. 275-80.

106. Sink, H. *Muscle development in drosophila.* Molecular biology intelligence unit. 2006, Georgetown, Tex. New York, N.Y.: Landes Bioscience/Eurekah.com; Springer Science+Business Media. 207 p.

107. Azpiazu, N. and Frasch, M. *tinman and bagpipe: two homeo box genes that determine cell fates in the dorsal mesoderm of Drosophila.* Genes Dev, 1993. **7**(7B): p. 1325-40.

108. Jakobsen, J.S. et al. *Temporal ChIP-on-chip reveals Biniou as a universal regulator of the visceral muscle transcriptional network.* Genes Dev, 2007. **21**(19): p. 2448-60.

109. Martin, B.S. et al. *A distinct set of founders and fusion-competent myoblasts make visceral muscles in the Drosophila embryo.* Development, 2001. **128**(17): p. 3331-8.

110. Hosono, C. et al. *Functional subdivision of trunk visceral mesoderm parasegments in Drosophila is required for gut and trachea development.* Development, 2003. **130**(3): p. 439-49.

111. Immergluck, K., Lawrence, P.A. and Bienz, M. *Induction across germ layers in Drosophila mediated by a genetic cascade.* Cell, 1990. **62**(2): p. 261-8.

112. Ward, E.J. and Skeath, J.B. *Characterization of a novel subset of cardiac cells and their progenitors in the Drosophila embryo.* Development, 2000. **127**(22): p. 4959-69.

113. Reim, I. and Frasch, M. *The Dorsocross T-box genes are key components of the regulatory network controlling early cardiogenesis in Drosophila.* Development, 2005. **132**(22): p. 4911-25.

114. Gajewski, K. et al. *Pannier is a transcriptional target and partner of Tinman during Drosophila cardiogenesis.* Dev Biol, 2001. **233**(2): p. 425-36.

115. Jagla, K. et al. *ladybird, a new component of the cardiogenic pathway in Drosophila required for diversification of heart precursors.* Development, 1997. **124**(18): p. 3471-9.

116. Gajewski, K. et al. *Genetically distinct cardial cells within the Drosophila heart.* Genesis, 2000. **28**(1): p. 36-43.

117. Molina, M.R. and Cripps, R.M. *Ostia, the inflow tracts of the Drosophila heart, develop from a genetically distinct subset of cardial cells.* Mech Dev, 2001. **109**(1): p. 51-9.

118. Cox, V.T. and Baylies, M.K. *Specification of individual Slouch muscle progenitors in Drosophila requires sequential Wingless signaling.* Development, 2005. **132**(4): p. 713-24.

119. Riechmann, V. et al. *Control of cell fates and segmentation in the Drosophila mesoderm.* Development, 1997. **124**(15): p. 2915-22.

120. Tapanes-Castillo, A. and Baylies, M.K. *Notch signaling patterns Drosophila mesodermal segments by regulating the bHLH transcription factor twist.* Development, 2004. **131**(10): p. 2359-72.

121. Carmena, A., Bate, M. and Jimenez, F. *Lethal of scute, a proneural gene, participates in the specification of muscle progenitors during Drosophila embryogenesis.* Genes Dev, 1995. **9**(19): p. 2373-83.

122. Corbin, V. et al. *A role for the Drosophila neurogenic genes in mesoderm differentiation.* Cell, 1991. **67**(2): p. 311-23.

123. Carmena, A. et al. *Inscuteable and numb mediate asymmetric muscle progenitor cell divisions during Drosophila myogenesis.* Genes & Development, 1998. **12**(3): p. 304-315.

124. Bate, M., Rushton, E. and Frasch, M. *A dual requirement for neurogenic genes in Drosophila myogenesis.* Dev Suppl, 1993: p. 149-61.

125. Chen, E.H. and Olson, E.N. *Towards a molecular pathway for myoblast fusion in Drosophila.* Trends Cell Biol, 2004. **14**(8): p. 452-60.

126. Lai, E.C. *Notch signaling: control of cell communication and cell fate.* Development, 2004. **131**(5): p. 965-73.

127. Giebel, B. *The notch signaling pathway is required to specify muscle progenitor cells in Drosophila.* Mech Dev, 1999. **86**(1-2): p. 137-45.

128. Ruiz-Gomez, M. et al. *Specific muscle identities are regulated by Kruppel during Drosophila embryogenesis.* Development, 1997. **124**(17): p. 3407-14.

129. Knirr, S., Azpiazu, N. and Frasch, M. *The role of the NK-homeobox gene slouch (S59) in somatic muscle patterning.* Development, 1999. **126**(20): p. 4525-35.

130. Jagla, T. et al. *ladybird determines cell fate decisions during diversification of Drosophila somatic muscles.* Development, 1998. **125**(18): p. 3699-708.

131. Duan, H., Skeath, J.B. and Nguyen, H.T. *Drosophila Lame duck, a novel member of the Gli superfamily, acts as a key regulator of myogenesis by controlling fusion-competent myoblast development.* Development, 2001. **128**(22): p. 4489-500.

132. Ruiz-Gomez, M. et al. *myoblasts incompetent encodes a zinc finger transcription factor required to specify fusion-competent myoblasts in Drosophila.* Development, 2002. **129**(1): p. 133-41.

133. Furlong, E.E. et al. *Patterns of gene expression during Drosophila mesoderm development.* Science, 2001. **293**(5535): p. 1629-33.

134. Koebernick, K. and Pieler, T. *Gli-type zinc finger proteins as bipotential transducers of Hedgehog signaling.* Differentiation, 2002. **70**(2-3): p. 69-76.

135. Strunkelnberg, M. et al. *rst and its paralogue kirre act redundantly during embryonic muscle development in Drosophila.* Development, 2001. **128**(21): p. 4229-39.

136. Shelton, C. et al. *The immunoglobulin superfamily member Hbs functions redundantly with Sns in interactions between founder and fusion-competent myoblasts.* Development, 2009. **136**(7): p. 1159-68.

137. Beckett, K. and Baylies, M.K. *3D analysis of founder cell and fusion competent myoblast arrangements outlines a new model of myoblast fusion.* Dev Biol, 2007. **309**(1): p. 113-25.

138. Bate, M. *The embryonic development of larval muscles in Drosophila.* Development, 1990. **110**(3): p. 791-804.

139. Berger, S. et al. *WASP and SCAR have distinct roles in activating the Arp2/3 complex during myoblast fusion.* J Cell Sci, 2008. **121**(Pt 8): p. 1303-13.

140. Richardson, B.E. et al. *SCAR/WAVE and Arp2/3 are crucial for cytoskeletal remodeling at the site of myoblast fusion.* Development, 2007. **134**(24): p. 4357-67.

141. Kim, S. et al. *A critical function for the actin cytoskeleton in targeted exocytosis of prefusion vesicles during myoblast fusion.* Dev Cell, 2007. **12**(4): p. 571-86.

142. Schnorrer, F. and Dickson, B.J. *Muscle building; mechanisms of myotube guidance and attachment site selection.* Dev Cell, 2004. **7**(1): p. 9-20.

143. Volk, T. *Singling out Drosophila tendon cells: a dialogue between two distinct cell types.* Trends Genet, 1999. **15**(11): p. 448-53.

144. Vorbruggen, G. and Jackle, H. *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 U S A, 1997. **94**(16): p. 8606-11.

145. Bate, M. and Rushton, E. *Myogenesis and muscle patterning in Drosophila.* C R Acad Sci III, 1993. **316**(9): p. 1047-61.

146. Crossley, A. *The morphology and development of the Drosophila muscular system*, in *The Genetics and Biology of Drosophila*, M. Ashburner and T. Wright, Editors. 1978, New York: Academic Press. p. 2b:499-560.

147. Horiuchi, T. and Aigaki, T. *Alternative trans-splicing: a novel mode of pre-mRNA processing.* Biol Cell, 2006. **98**(2): p. 135-40.

148. Zollman, S. et al. *The BTB domain, found primarily in zinc finger proteins, defines an evolutionarily conserved family that includes several developmentally regulated genes in Drosophila.* Proc Natl Acad Sci U S A, 1994. **91**(22): p. 10717-21.

149. Fairall, L. et al. *The crystal structure of a two zinc-finger peptide reveals an extension to the rules for zinc-finger/DNA recognition.* Nature, 1993. **366**(6454): p. 483-7.

150. Read, D. and Manley, J.L. *Alternatively spliced transcripts of the Drosophila tramtrack gene encode zinc finger proteins with distinct DNA binding specificities.* EMBO J, 1992. **11**(3): p. 1035-44.

151. Brown, J.L. et al. *Repression of the Drosophila fushi tarazu (ftz) segmentation gene.* EMBO J, 1991. **10**(3): p. 665-74.

152. Harrison, S.D. and Travers, A.A. *The tramtrack gene encodes a Drosophila finger protein that interacts with the ftz transcriptional regulatory region and shows a novel embryonic expression pattern.* EMBO J, 1990. **9**(1): p. 207-16.

153. Chen, Y.J. et al. *Tramtrack69 is required for the early repression of tailless expression.* Mech Dev, 2002. **116**(1-2): p. 75-83.

154. Stein, L.D. et al. *The generic genome browser: a building block for a model organism system database.* Genome Res, 2002. **12**(10): p. 1599-610.

155. Giesen, K. et al. *Glial development in the Drosophila CNS requires concomitant activation of glial and repression of neuronal differentiation genes.* Development, 1997. **124**(12): p. 2307-16.

156. Murawsky, C.M. et al. *Tramtrack69 interacts with the dMi-2 subunit of the Drosophila NuRD chromatin remodelling complex.* EMBO Rep, 2001. **2**(12): p. 1089-94.

157. Schaeper, U. et al. *Molecular cloning and characterization of a cellular phosphoprotein that interacts with a conserved C-terminal domain of adenovirus E1A involved in negative modulation of oncogenic transformation.* Proc Natl Acad Sci U S A, 1995. **92**(23): p. 10467-71.

158. Wen, Y. et al. *The N-terminal BTB/POZ domain and C-terminal sequences are essential for Tramtrack69 to specify cell fate in the developing Drosophila eye.* Genetics, 2000. **156**(1): p. 195-203.

159. Dhordain, P. et al. *The LAZ3(BCL-6) oncoprotein recruits a SMRT/mSIN3A/histone deacetylase containing complex to mediate transcriptional repression.* Nucleic Acids Res, 1998. **26**(20): p. 4645-51.

160. Brown, J.L. and Wu, C. *Repression of Drosophila pair-rule segmentation genes by ectopic expression of tramtrack.* Development, 1993. **117**(1): p. 45-58.

161. French, R.L., Cosand, K.A. and Berg, C.A. *The Drosophila female sterile mutation twin peaks is a novel allele of tramtrack and reveals a requirement for Ttk69 in epithelial morphogenesis.* Dev Biol, 2003. **253**(1): p. 18-35.

162. Araujo, S.J., Cela, C. and Llimargas, M. *Tramtrack regulates different morphogenetic events during Drosophila tracheal development.* Development, 2007. **134**(20): p. 3665-76.

163. Guo, M. et al. *tramtrack acts downstream of numb to specify distinct daughter cell fates during asymmetric cell divisions in the Drosophila PNS.* Neuron, 1995. **14**(5): p. 913-25.

164. Harvie, P.D., Filippova, M. and Bryant, P.J. *Genes expressed in the ring gland, the major endocrine organ of Drosophila melanogaster.* Genetics, 1998. **149**(1): p. 217-31.

165. Xiong, W.C. and Montell, C. *tramtrack is a transcriptional repressor required for cell fate determination in the Drosophila eye.* Genes Dev, 1993. **7**(6): p. 1085-96.

166. Jordan, K.C. et al. *Notch signaling through tramtrack bypasses the mitosis promoting activity of the JNK pathway in the mitotic-to-endocycle transition of Drosophila follicle cells.* BMC Dev Biol, 2006. **6**: p. 16.

167. Sun, J. and Deng, W.M. *Hindsight mediates the role of notch in suppressing hedgehog signaling and cell proliferation.* Dev Cell, 2007. **12**(3): p. 431-42.

168. Althauser, C. et al. *Fringe-dependent notch activation and tramtrack function are required for specification of the polar cells in Drosophila oogenesis.* Dev Dyn, 2005. **232**(4): p. 1013-20.

169. Pritchard, D.K. and Schubiger, G. *Activation of transcription in Drosophila embryos is a gradual process mediated by the nucleocytoplasmic ratio.* Genes Dev, 1996. **10**(9): p. 1131-42.

170. Wheeler, J.C. et al. *Distinct in vivo requirements for establishment versus maintenance of transcriptional repression.* Nat Genet, 2002. **32**(1): p. 206-10.

171. Badenhorst, P. *Tramtrack controls glial number and identity in the Drosophila embryonic CNS.* Development, 2001. **128**(20): p. 4093-101.

172. Vervoort, M. et al. *Genetic basis of the formation and identity of type I and type II neurons in Drosophila embryos.* Development, 1997. **124**(14): p. 2819-28.

173. Audibert, A., Simon, F. and Gho, M. *Cell cycle diversity involves differential regulation of Cyclin E activity in the Drosophila bristle cell lineage.* Development, 2005. **132**(10): p. 2287-97.

174. Badenhorst, P., Finch, J.T. and Travers, A.A. *Tramtrack co-operates to prevent inappropriate neural development in Drosophila.* Mech Dev, 2002. **117**(1-2): p. 87-101.

175. Siddall, N.A. et al. *Ttk69-dependent repression of lozenge prevents the ectopic development of R7 cells in the Drosophila larval eye disc.* BMC Dev Biol, 2009. **9**: p. 64.

176. Lai, Z.C. et al. *Loss of tramtrack gene activity results in ectopic R7 cell formation, even in a sina mutant background.* Proc Natl Acad Sci U S A, 1996. **93**(10): p. 5025-30.

177. Baonza, A. et al. *Pointed and Tramtrack69 establish an EGFR-dependent transcriptional switch to regulate mitosis.* Nat Cell Biol, 2002. **4**(12): p. 976-80.

178. Lai, Z.C. and Li, Y. *Tramtrack69 is positively and autonomously required for Drosophila photoreceptor development.* Genetics, 1999. **152**(1): p. 299-305.

179. Sutherland, D., Samakovlis, C. and Krasnow, M.A. *branchless encodes a Drosophila FGF homolog that controls tracheal cell migration and the pattern of branching.* Cell, 1996. **87**(6): p. 1091-101.

180. Jung, A.C. et al. *Polychaetoid/ZO-1 is required for cell specification and rearrangement during Drosophila tracheal morphogenesis.* Curr Biol, 2006. **16**(12): p. 1224-31.

181. Read, D., Levine, M. and Manley, J.L. *Ectopic expression of the Drosophila tramtrack gene results in multiple embryonic defects, including repression of even-skipped and fushi tarazu.* Mech Dev, 1992. **38**(3): p. 183-95.

182. Guo, M., Jan, L.Y. and Jan, Y.N. *Control of daughter cell fates during asymmetric division: interaction of Numb and Notch.* Neuron, 1996. **17**(1): p. 27-41.

183. Okabe, M. et al. *Translational repression determines a neuronal potential in Drosophila asymmetric cell division.* Nature, 2001. **411**(6833): p. 94-8.

184. Crew, J.R., Batterham, P. and Pollock, J.A. *Developing compound eye in lozenge mutants of Drosophila : lozenge expression in the R7 equivalence group.* Development Genes and Evolution, 1997. **206**(8): p. 481-493.

185. Burgler, C. and Macdonald, P.M. *Prediction and verification of microRNA targets by MovingTargets, a highly adaptable prediction method.* BMC Genomics, 2005. **6**(1): p. 88.

186. Iovino, N., Pane, A. and Gaul, U. *miR-184 has multiple roles in Drosophila female germline development.* Dev Cell, 2009. **17**(1): p. 123-33.

187. Li, S. et al. *Photoreceptor cell differentiation requires regulated proteolysis of the transcriptional repressor Tramtrack.* Cell, 1997. **90**(3): p. 469-78.

188. Spencer, M.L., Theodosiou, M. and Noonan, D.J. *NPDC-1, a novel regulator of neuronal proliferation, is degraded by the ubiquitin/proteasome system through a PEST degradation motif.* J Biol Chem, 2004. **279**(35): p. 37069-78.

189. Rogers, S., Wells, R. and Rechsteiner, M. *Amino acid sequences common to rapidly degraded proteins: the PEST hypothesis.* Science, 1986. **234**(4774): p. 364-8.

190. Li, S., Xu, C. and Carthew, R.W. *Phyllopod acts as an adaptor protein to link the sina ubiquitin ligase to the substrate protein tramtrack.* Mol Cell Biol, 2002. **22**(19): p. 6854-65.

191. Tang, A.H. et al. *PHYL acts to down-regulate TTK88, a transcriptional repressor of neuronal cell fates, by a SINA-dependent mechanism.* Cell, 1997. **90**(3): p. 459-67.

192. Hirota, Y. et al. *Musashi and seven in absentia downregulate Tramtrack through distinct mechanisms in Drosophila eye development.* Mech Dev, 1999. **87**(1-2): p. 93-101.

193. Cooper, S.E. et al. *Two modes of degradation of the tramtrack transcription factors by Siah homologues.* J Biol Chem, 2008. **283**(2): p. 1076-83.

194. Lehembre, F. et al. *Covalent modification of the transcriptional repressor tramtrack by the ubiquitin-related protein Smt3 in Drosophila flies.* Mol Cell Biol, 2000. **20**(3): p. 1072-82.

195. Barolo, S., Carver, L.A. and Posakony, J.W. *GFP and beta-galactosidase transformation vectors for promoter/enhancer analysis in Drosophila.* BioTechniques, 2000. **29**(4): p. 726-732.

196. Zinzen, R.P. et al. *Combinatorial binding predicts spatio-temporal cis-regulatory activity.* Nature, 2009. **462**(7269): p. 65-70.

197. Bischof, J. et al. *An optimized transgenesis system for Drosophila using germ-line-specific phiC31 integrases.* Proc Natl Acad Sci U S A, 2007. **104**(9): p. 3312-7.

198. Sandmann, T., Jakobsen, J.S. and Furlong, E.E. *ChIP-on-chip protocol for genome-wide analysis of transcription factor binding in Drosophila melanogaster embryos.* Nat Protoc, 2006. **1**(6): p. 2839-55.

199. Goldstein, L., E. Fyrberg, ed. *Drosophila melanogaster: Practical Uses in Cell and Molecular Biology.* Methods in Cell Biology. 1994.

200. Parks, A.L. et al. *Systematic generation of high-resolution deletion coverage of the Drosophila melanogaster genome.* Nat Genet, 2004. **36**(3): p. 288-92.

201. Adryan, B. and Teichmann, S.A. *FlyTF: a systematic review of site-specific transcription factors in the fruit fly Drosophila melanogaster.* Bioinformatics, 2006. **22**(12): p. 1532-3.

202. Tomancak, P. et al. *Systematic determination of patterns of gene expression during Drosophila embryogenesis.* Genome Biol, 2002. **3**(12): p. RESEARCH0088.

203. Wen, S. et al. *Analysis of three doublesex related genes suggests that they play roles in sexual differentiation.* Program and Abstracts. 42nd Annual Drosophila Research Conference, Washington DC, 2001, 2001: p. 656B.

204. Ragab, A., Thompson, E.C. and Travers, A.A. *High mobility group proteins HMGD and HMGZ interact genetically with the Brahma chromatin remodeling complex in Drosophila.* Genetics, 2006. **172**(2): p. 1069-1078.

205. De Graeve, F. et al. *Identification of the Drosophila progenitor of mammalian Kruppel-like factors 6 and 7 and a determinant of fly development.* Gene, 2003. **314**: p. 55-62.

206. Ryder, E. et al. *The DrosDel collection: a set of P-element insertions for generating custom chromosomal aberrations in Drosophila melanogaster.* Genetics, 2004. **167**(2): p. 797-813.

207. Leiss, D. et al. *Beta 3 tubulin expression characterizes the differentiating mesodermal germ layer during Drosophila embryogenesis.* Development, 1988. **104**(4): p. 525-31.

208. Kremser, T. et al. *Tinman regulates the transcription of the beta3 tubulin gene (betaTub60D) in the dorsal vessel of Drosophila.* Dev Biol, 1999. **216**(1): p. 327-39.

209. Ashburner, M. et al. *Report of New Mutants.* Drosophila Information Service, 1981. **56**: p. 186-191.

210. Reuter, R., Grunewald, B. and Leptin, M. *A role for the mesoderm in endodermal migration and morphogenesis in Drosophila.* Development, 1993. **119**(4): p. 1135-45.

211. Wolfstetter, G. et al. *Fusion of circular and longitudinal muscles in Drosophila is independent of the endoderm but further visceral muscle differentiation requires a close contact between mesoderm and endoderm.* Mech Dev, 2009. **126**(8-9): p. 721-36.

212. Thibault, S.T. et al. *A complementary transposon tool kit for Drosophila melanogaster using P and piggyBac.* Nat Genet, 2004. **36**(3): p. 283-7.

213. Wang, L., Lam, G. and Thummel, C.S. *Med24 and Mdh2 are required for Drosophila larval salivary gland cell death.* Dev Dyn. **239**(3): p. 954-64.

214. Cunha, P.e.a. *Combinatorial transcription factor binding leads to diverse regulatory responses: Lmd is a tissue-specific modulator of Mef2 activity.* PLoS Genet, 2009**, in press**.

215. Hummel, T., Schimmelpfeng, K. and Klambt, C. *Commissure formation in the embryonic CNS of Drosophila.* Dev Biol, 1999. **209**(2): p. 381-98.

216. Hendzel, M.J. et al. *Mitosis-specific phosphorylation of histone H3 initiates primarily within pericentromeric heterochromatin during G2 and spreads in an ordered fashion coincident with mitotic chromosome condensation.* Chromosoma, 1997. **106**(6): p. 348-60.

217. Roos, J. et al. *Drosophila Futsch regulates synaptic microtubule organization and is necessary for synaptic growth.* Neuron, 2000. **26**(2): p. 371-82.

218. Hummel, T. et al. *Drosophila Futsch/22C10 is a MAP1B-like protein required for dendritic and axonal development.* Neuron, 2000. **26**(2): p. 357-70.

219. Klambt, C., Jacobs, J.R. and Goodman, C.S. *The midline of the Drosophila central nervous system: a model for the genetic analysis of cell fate, cell migration, and growth cone guidance.* Cell, 1991. **64**(4): p. 801-15.

220. Nose, A., Mahajan, V.B. and Goodman, C.S. *Connectin: a homophilic cell adhesion molecule expressed on a subset of muscles and the motoneurons that innervate them in Drosophila.* Cell, 1992. **70**(4): p. 553-67.

221. Ruiz-Gomez, M. et al. *Drosophila dumbfounded: a myoblast attractant essential for fusion.* Cell, 2000. **102**(2): p. 189-98.

222. Nose, A., Isshiki, T. and Takeichi, M. *Regional specification of muscle progenitors in Drosophila: the role of the msh homeobox gene.* Development, 1998. **125**(2): p. 215-23.

223. Brand, A.H. and Perrimon, N. *Targeted gene expression as a means of altering cell fates and generating dominant phenotypes.* Development, 1993. **118**(2): p. 401-15.

224. Kocherlakota, K.S. et al. *Analysis of the cell adhesion molecule sticks-and-stones reveals multiple redundant functional domains, protein-interaction motifs and phosphorylated tyrosines that direct myoblast fusion in Drosophila melanogaster.* Genetics, 2008. **178**(3): p. 1371-83.

225. Stute, C. *Establishment of cell type specific Gal4-driver lines for the mesoderm of Drosophila.* Dros. Inf. Serv., 2006. **89**: p. 111-115.

226. Menon, S.D. and Chia, W. *Drosophila rolling pebbles: a multidomain protein required for myoblast fusion that recruits D-Titin in response to the myoblast attractant Dumbfounded.* Dev Cell, 2001. **1**(5): p. 691-703.

227. Sepp, K.J. and Auld, V.J. *Conversion of lacZ enhancer trap lines to GAL4 lines using targeted transposition in Drosophila melanogaster.* Genetics, 1999. **151**(3): p. 1093-101.

228. Pagans, S. et al. *The Drosophila transcription factor tramtrack (TTK) interacts with Trithorax-like (GAGA) and represses GAGA-mediated activation.* Nucleic Acids Res, 2002. **30**(20): p. 4406-13.

229. Ji, H. and Wong, W.H. *TileMap: create chromosomal map of tiling array hybridizations.* Bioinformatics, 2005. **21**(18): p. 3629-36.

230. Read, D., Nishigaki, T. and Manley, J.L. *The Drosophila even-skipped promoter is transcribed in a stage-specific manner in vitro and contains multiple, overlapping factor-binding sites.* Mol Cell Biol, 1990. **10**(8): p. 4334-44.

231. Bradley, P.L. and Andrew, D.J. *ribbon encodes a novel BTB/POZ protein required for directed cell migration in Drosophila melanogaster.* Development, 2001. **128**(15): p. 3001-3015.

232. Cheah, P.Y., Chia, W. and Yang, X. *Jumeaux, a novel Drosophila winged-helix family protein, is required for generating asymmetric sibling neuronal cell fates.* Development, 2000. **127**(15): p. 3325-3335.

233. Tomancak, P. et al. *Global analysis of patterns of gene expression during Drosophila embryogenesis.* Genome Biol, 2007. **8**(7): p. R145.

234. Estrada, B. et al. *An integrated strategy for analyzing the unique developmental programs of different myoblast subtypes.* PLoS Genet, 2006. **2**(2): p. e16.

235. Schneider, D.S. et al. *Dominant and recessive mutations define functional domains of Toll, a transmembrane protein required for dorsal-ventral polarity in the Drosophila embryo.* Genes Dev, 1991. **5**(5): p. 797-807.

236. Nguyen, H.T. and Xu, X. *Drosophila mef2 expression during mesoderm development is controlled by a complex array of cis-acting regulatory modules.* Dev Biol, 1998. **204**(2): p. 550-66.

237. Gisselbrecht, S. et al. *heartless encodes a fibroblast growth factor receptor (DFR1/DFGF-R2) involved in the directional migration of early mesodermal cells in the Drosophila embryo.* Genes Dev, 1996. **10**(23): p. 3003-17.

238. Lai, Z.C. et al. *Loss of function of the Drosophila zfh-1 gene results in abnormal development of mesodermally derived tissues.* Proc Natl Acad Sci U S A, 1993. **90**(9): p. 4122-6.

239. Nusslein-Volhard, C., Wieschaus, E. and Kluding, H. *Mutations affecting the pattern of the larval cuticle in Drosophila melanogaster.* Roux's Archives of Developmental Biology, 1984. **193**: p. 267-282.

240. Grau, Y., Carteret, G. and Simpson, P. *Mutation and chromosomal rearrangements affecting the expression of snail, a gene involved in embryonic patterning in Drosophila melanogaster.* Genetics, 1984. **108**: p. 347-360.

241. Bourouis, M. and Jarry, B. *Vectors containing a prokaryotic dihydrofolate reductase gene transform Drosophila cells to methotrexate-resistance.* EMBO J, 1983. **2**(7): p. 1099-104.

242. Sims, D. et al. *FLIGHT: database and tools for the integration and cross-correlation of large-scale RNAi phenotypic datasets.* Nucleic Acids Res, 2006. **34**(Database issue): p. D479-83.

243. Celniker, S.E. et al. *Unlocking the secrets of the genome.* Nature, 2009. **459**(7249): p. 927-30.

244. Ganguly, A., Jiang, J. and Ip, Y.T. *Drosophila WntD is a target and an inhibitor of the Dorsal/Twist/Snail network in the gastrulating embryo.* Development, 2005. **132**(15): p. 3419-29.

245. Alberga, A. et al. *The snail gene required for mesoderm formation in Drosophila is expressed dynamically in derivatives of all three germ layers.* Development, 1991. **111**(4): p. 983-92.

246. Nusslein-Volhard, C. and Wieschaus, E. *Mutations affecting segment number and polarity in Drosophila.* Nature, 1980. **287**(5785): p. 795-801.

247. Bour, B.A. et al. *Drosophila MEF2, a transcription factor that is essential for myogenesis.* Genes Dev, 1995. **9**(6): p. 730-41.

248. Ciglar, L. and Furlong, E.E. *Conservation and divergence in developmental networks: a view from Drosophila myogenesis.* Curr Opin Cell Biol, 2009. **21**(6): p. 754-60.

249. Fire, A. et al. *Potent and specific genetic interference by double-stranded RNA in Caenorhabditis elegans.* Nature, 1998. **391**(6669): p. 806-11.

250. Dietzl, G. et al. *A genome-wide transgenic RNAi library for conditional gene inactivation in Drosophila.* Nature, 2007. **448**(7150): p. 151-6.

251. Tavernarakis, N. et al. *Heritable and inducible genetic interference by double-stranded RNA encoded by transgenes.* Nat Genet, 2000. **24**(2): p. 180-3.

252. Schnorrer, F. et al. *Systematic genetic analysis of muscle morphogenesis and function in Drosophila.* Nature. **464**(7286): p. 287-91.

253. Ranganayakulu, G. et al. *A series of mutations in the D-MEF2 transcription factor reveal multiple functions in larval and adult myogenesis in Drosophila.* Dev Biol, 1995. **171**(1): p. 169-81.

254. Dworak, H.A. et al. *Characterization of Drosophila hibris, a gene related to human nephrin.* Development, 2001. **128**(21): p. 4265-76.

255. Chang, H.C. et al. *phyllopod functions in the fate determination of a subset of photoreceptors in Drosophila.* Cell, 1995. **80**(3): p. 463-72.

256. Artero, R. et al. *Notch and Ras signaling pathway effector genes expressed in fusion competent and founder cells during Drosophila myogenesis.* Development, 2003. **130**(25): p. 6257-72.

257. Cunha, P.e.a. *Combinatorial transcription factor binding leads to diverse regulatory responses: Lmd is a tissue-specific modulator of Mef2 activity.* PLoS Genet, 2010. **, in press**.

258. Lin, M. and Kellis, M., *Gene models flagged as uncertain or dubious.* 2006.

259. Lefstin, J.A. and Yamamoto, K.R. *Allosteric effects of DNA on transcriptional regulators.* Nature, 1998. **392**(6679): p. 885-8.

260. Piepenburg, O., Vorbruggen, G. and Jackle, H. *Drosophila segment borders result from unilateral repression of hedgehog activity by wingless signaling.* Mol Cell, 2000. **6**(1): p. 203-9.

261. Theisen, H. et al. *Wingless directly represses DPP morphogen expression via an armadillo/TCF/Brinker complex.* PLoS One, 2007. **2**(1): p. e142.

262. Blauwkamp, T.A., Chang, M.V. and Cadigan, K.M. *Novel TCF-binding sites specify transcriptional repression by Wnt signalling.* EMBO J, 2008. **27**(10): p. 1436-46.

263. Pan, D. and Courey, A.J. *The same dorsal binding site mediates both activation and repression in a context-dependent manner.* EMBO J, 1992. **11**(5): p. 1837-42.

264. De Renzis, S. et al. *Unmasking activation of the zygotic genome using chromosomal deletions in the Drosophila embryo.* PLoS Biol, 2007. **5**(5): p. e117.