Mef2 in muscle: A CRISPR-based approach to investigate the regulation of a master transcription factor

Sean Anthony Hubbert
December 2023

This thesis is submitted for the degree of Doctor of Philosophy

Acknowledgments

I would firstly like to thank the Biotechnology and Biological Sciences Research Council-funded SWBio DTP for funding, and supporting me during my PhD project.

There are many others, for whom this work would not have been possible. In particular, I would like to thank Dr Michael Taylor for his input and guidance over the past four years. I would also like all of those who have been a part of the Cardiff Fly community, which has provided valuable feedback at meetings, as well as generally being a wonderful community to be a part of.

In particular, I would like to thank Rob, Fi, Nicola, Rakhee, Toby, and Lainey, for whom I have been lucky to share time with throughout the entirety of my PhD and am very lucky to call my close friends.

Finally, I would like to thank my fiancé Candice, and my family for their unwavering support throughout my PhD.

Summary

Myocyte enhancer factor-2 (Mef2) is a highly conserved transcription factor required for muscle differentiation in *Drosophila*, and functions by activating the expression of hundreds of genes. By using CRISPR-Cas9 genome engineering, not previously used for the study Mef2 function *in vivo*, I have generated a novel suite of valuable tools which can be used to probe endogenous Mef2 function *in vivo*.

Interestingly, Mef2 function is also highly expressed in undifferentiated myoblast populations, many hours before these cells differentiate and express many Mef2 target genes. Therefore, Mef2 activity must be regulated in space and time to enable to proper coordination of muscle development *in vivo*. The Class IIa histone deacetylases (HDACs) have been shown to physically interact with, and negatively regulate Mef2 *in vitro*. However, before this project, there had been relatively little research as to their function in the context of muscle differentiation *in vivo*. In this project, I further our understanding of the role of Class IIa HDACs during muscle differentiation using *Drosophila melanogaster* as a model. Firstly, I use the Gal4-UAS system to probe the functional consequences of overexpressing *HDAC4 in vivo*, as well as to identify important residues for HDAC4 function *in vivo*. Secondly, I utilise CRISPR-Cas9 genome engineering to generate a novel, rescuable *HDAC4* null allele and show that *HDAC4* is essential for viability, while its loss also disrupts normal muscle patterning *in vivo*. Importantly, this allele will be a critical tool to further define the role of *HDAC4* during muscle differentiation *in vivo*, as well for investigating Class IIa HDAC function more broadly.

Finally, I also identify critical residues required for Mef2 transcriptional activity *in vivo*. I link these observations to, and investigate the function of the *Drosophila* orthologue of the p300/CBP Histone acetyltransferase (HAT), Nejire, during muscle differentiation *in vivo*.

In this work, I utilised CRISPR-Cas9 genome engineering to develop valuable tools to facilitate the understanding of Mef2 function and regulation, while I have also investigated the role of HDAC4 and Nejire in the regulation of muscle differentiation *in vivo*.

Contents

Table of Figures

Abbreviations

3xHA – hemagglutinin epitope tag

AMP – adult muscle progenitor

APF – after puparium formation

BDSC – Bloomington *Drosophila* stock center

CDS – coding sequence

CNS – central nervous system

CRISPR – clustered regularly interspersed short palindromic repeats

CT – C-terminal

DSB – double-strand break

DT – direct tag

FC – founder cells

FCM – fusion competent myoblast

GFP – green fluorescent protein

HA – homology arm

HAT – histone acetyltransferase

HDAC – histone deacetylase

HDR – homology directed repair

IRD – insertion-ready deletion

Mef2 – myocyte enhancer factor-2

Mhc – myosin heavy chain

Nej – Nejire

NES – nuclear export sequence

NHEJ – non-homologous end joining

NLS – nuclear localisation sequence

NT – N-terminal

TF – transcription factor

TSS – transcription start site

UAS – upstream activator sequence

UTR – untranslated region

VDRC – Vienna drosophila stock centre

WT – Wild-type

Chapter 1: Introduction

1.1 Gene expression regulation underpins development

Animals are complex multicellular eukaryotic organisms with a vast array of differentiated cell types which perform specialised functions to enable the correct functioning of the organism. Despite this complexity, all cells arise from a single totipotent zygote formed during fertilization and are genetically identical to one another. In addition, cell differentiation also proceeds throughout the life cycle in the context of tissue regeneration and repair, via the presence of distinct stem cell populations; while many of the concepts regarding differentiation are also relevant in the context of diseases such as cancer (Hanahan and Weinberg 2011; Nelson 2022). Therefore, understanding the underlying mechanisms of cell differentiation is important, and can ultimately be described by the fundamental question which underpins developmental biology: How are populations of genetically identically cells driven down different cell differentiation pathways?

Given the same genomic information, the underlying principle regulating cellular differentiation is not the genome sequence itself (Davidson 2010), but more how the expression of the same compliment of genetic information is tightly regulated in time and space to control the establishment of distinct cellular phenotypes. Ever since the first completed sequence of the human genome (Lander et al. 2001), significant progress has been made in understanding that the spatio-temporal control of gene expression is established through the formation of distinct transcriptional programmes, or regulatory states. In these states, different cell populations show a unique interplay between highly conserved cell signalling pathways and the regulated expression of transcription factors, which ultimately results in distinct cell populations that express some genes, and repress others (Levine and Davidson 2005; Reynolds et al. 2013).

Genes are complex structures in which their expression is tightly regulated by the presence of multiple cis-regulatory modules (CRM's), including the so-called 'core promoter' responsible for the recruitment and formation of the basal transcriptional machinery, and more distally placed enhancers (Spitz and Furlong 2012; Shlyueva et al. 2014). Although typically associated with being upstream of the core promoter and transcription start site (TSS), enhancers may be present both within, and downstream of the gene they regulate (Pennacchio et al. 2013; Panigrahi and O'Malley 2021). Transcription factors bind enhancers with a high level of sequence specificity, and can function to either activate, or repress the transcription of their target genes. Thus, the

expression of different complements of transcription factors in discrete cell populations is a fundamental way in which gene expression is regulated throughout development (Spitz and Furlong 2012). Because genes integrate signals from multiple CRM's to regulate their expression, CRM's and associated transcription factors can fine-tune the transcriptional output of their target genes (Wilczyski and Furlong 2010). Furthermore, given that most transcription factor DNA-binding motifs recognise relatively short DNA sequences, most transcription factors often have hundreds of potential gene targets, including tissue specific genes, but also other regulatory proteins, signalling pathway components, and transcription factors themselves (Wilczyski and Furlong 2010). This complexity of gene regulatory elements within higher eukaryotic genomes underpins the formation of gene-regulatory networks, whereby inductive signals coupled with the expression of specific transcription factors results in the formation of highly regulated transcriptional networks required for the proper differentiation of individual cell types (Stathopoulos and Levine 2004; Levine and Davidson 2005; Davidson and Erwin 2006; Wilczyski and Furlong 2010; Singh et al. 2018). Some of these factors, known as master transcriptional regulators, which include the *Drosophila* mesodermal specification gene *Twist* (Sandmann et al. 2007) and the mammalian pluripotency factors *Oct-4, Sox-2 and* Nanog (Whyte et al. 2013) often sit at the top of gene regulatory networks, modulate the expression of hundreds of targets, which indeed themselves also may regulate at lower levels in the network. Such master regulators can ultimately promote differentiation, or even trans-differentiation of existing cell types**.** Thus, the regulation of gene expression through transcription factor recruitment is crucial for coordinating gene expression patterns within a given cell through the establishment of complex gene-regulatory networks (Wilczyski and Furlong 2010).

1.2 *Drosophila melanogaster* as a model for gene expression regulation during development

Since Thomas Hunt Morgan's work understanding the basis of the inheritance of sex-linked genes (Morgan 1910), the humble fruit fly, *Drosophila melanogaster* has become a dominant model in the field of developmental genetics, while also having an impressive record of informing human biology. Despite the obvious phenotypic differences, and the fact that species' lineages diverged from a common ancestor in

excess of 750 million years ago (Shih et al. 2015), 60% of *Drosophila* genes have a human orthologue (Maqbool and Jagla 2007). Many of these conserved genes form components of conserved signalling pathways and transcriptional regulatory proteins which underpin developmental processes. A brilliant example of this comes from the from the forward genetic screen carried out by Nüsslein-Volhard and Wieschaus (1980). They identified 15 loci in which mutations caused patterning defects and disruption of the formation of the *Drosophila* larval body plan. Eleven of these fifteen genes encode transcription factors, while the other four are components of the Hedgehog or Wingless signalling pathways. Human orthologues of many of these genes are important during development, as well as in human disease: the transcription factors *paired* and *gooseberry* are homologous to the Paired-Box (*Pax*) genes in humans, which are involved in the development of many organs and tissues, and mutations in these genes are causes of congenital developmental defects such as Waardenburg's syndrome (Blake and Ziman 2014). Moreover, Wnt and Hedgehog signalling are two of the most conserved signalling pathways in multicellular eukaryotes, and are involved in the regulation of a vast array of cellular and developmental processes (Logan and Nusse 2004; Perrimon et al. 2012; Sasai et al. 2019).

The true power of *Drosophila* as a model comes in its accessibility for genetic and functional analysis. Despite the conservation, and the availability of fully sequenced genomes for both species (Adams et al. 2000; Lander et al. 2001; Nurk et al. 2022), the *Drosophila* genome is considerably smaller, has fewer chromosomes, and has less redundancy, thereby reducing the complexity associated with understanding gene function. Furthermore, *Drosophila's* short generation time and relatively cheap maintenance costs mean significant biological conclusions can be reached quicker, and cheaper than in other model organisms, in particular vertebrates such as mice (Tolwinski 2017). In addition, and as a result of its extensive use by the research community, development of novel genetic techniques and transgenesis methods have made addressing research questions easier (Venken and Bellen 2007), For example: Convenient genetic tools such as balancer chromosomes, which allow the maintenance of stocks possessing such deleterious or lethal mutations (Miller et al. 2019); tissue-specific gene expression manipulation *in vivo* using the Gal4/UAS system (Brand and Perrimon 1993; Del Valle Rodríguez et al. 2012) ; or more recently

the development of CRISPR-Cas9 genome engineering (Ran et al. 2013), all significant benefit the use of *Drosophila melanogaster* in genetics research.

1.3 *Drosophila* myogenesis - a model system for gene expression regulation during development

Skeletal (or somatic) muscle is the largest tissue by mass in motile animals, constituting anywhere between 35% and 75% of the total body mass (Bothe and Baylies 2016). It also displays a vast array of structural and functional heterogeneity adapted for specialised body movements (Nikonova et al. 2020). Fundamentally though, the structural and functional units of somatic musculature are highly conserved throughout evolution. Muscle fibers are a single, syncytial muscle cell whose contractile ability is generated through long, continuous chains of sarcomeres (myofibrils). (**Fig. 1.1**). These sarcomeres, consisting of thick (Myosin heavy chain) and thin (Actin) filaments, as well as a plethora of other structural proteins, utilise ATP hydrolysis to generate the force required for muscle contractions in a process called excitation-contraction coupling (Poovathumkadavil and Jagla 2020).

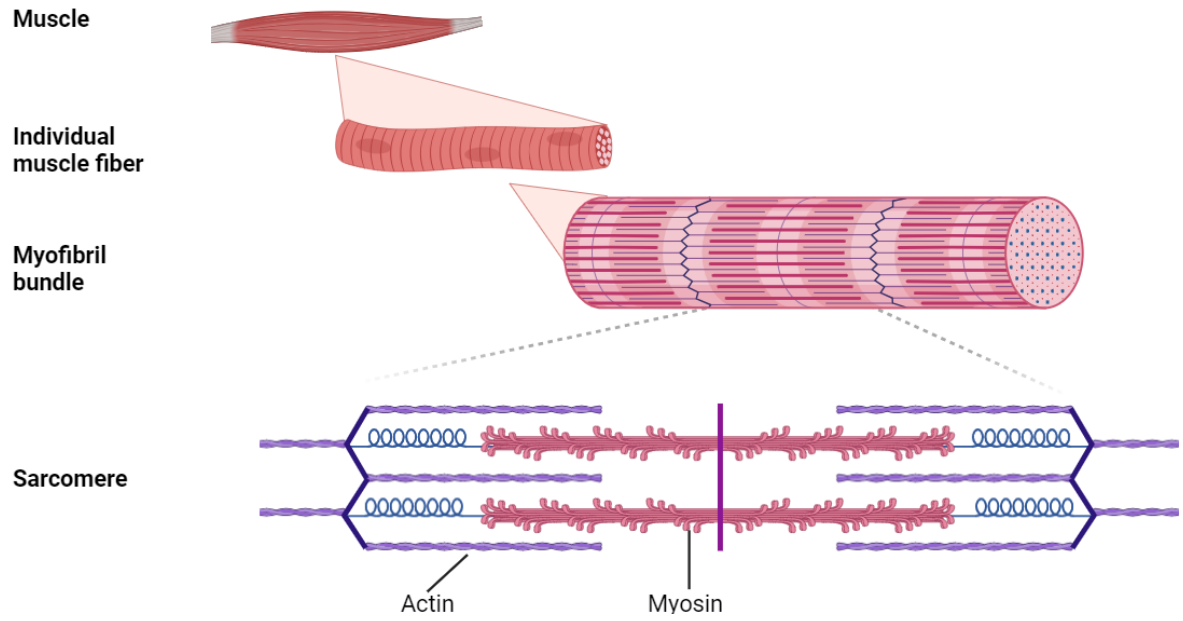


Figure 1. 1 Overview of somatic muscle structure
Somatic muscles consist of individual syncytial muscle cells, known as muscle fibers. These fibers contain bundles of myofibrils, which run the length of the muscle fiber and consist of repeating units of the functional unit of muscle, the sarcomere. Sarcomeres are highly organised structures consisting of Actin and Myosin, as well as many other structural proteins. Sarcomeres generate the force required for muscle contraction and stimulate muscle shortening.

Given their size, functional significance, and the existence of many myopathies affecting human skeletal muscle development and function, notably the most well-known group called muscular dystrophies (Mercuri et al. 2019), it is no surprise that understanding the mechanisms which regulate muscle development (myogenesis) has received significant attention among developmental geneticists. Indeed, it has become a paradigm for the study of gene expression regulation in the context of stem cell maintenance, proliferation, and differentiation (Bentzinger et al. 2012). Despite the evolutionary distance from humans, *Drosophila* has proved an excellent model for the study of such mechanisms (Weitkunat and Schnorrer 2014; Lemke and Schnorrer 2017). Many of the developmental processes that govern skeletal muscle formation are conserved throughout evolution: Undifferentiated pools of myoblasts proliferate, migrate, fuse with one another to form syncytial muscle fibres, which contain repeating units of the muscle functional unit - the sarcomere. Furthermore, many of the genes implicated in regulating these developmental processes, such as the transcription factor Myocyte-enhancer factor 2 (Mef2), which is thought to coordinate the gene expression program during myogenesis, are highly conserved between the two species (Poovathumkadavil and Jagla 2020).

Drosophila melanogaster is a holometabolic insect consisting of embryonic, larval, pupal, and adult life stages. To accompany this, there are two independent stages of myogenesis to facilitate the two motile stages of the *Drosophila* life cycle (larval and adult). The presence of two distinct stages is not surprising given the differing functional requirements of the somatic muscles: for example, the larval body-wall muscles facilitate crawling on a solid medium; whereas the adult somatic muscles carry out substantially different functions including walking, jumping, and flying.

The first wave of myogenesis occurs during embryogenesis and involves the generation of the larval body-wall muscles, which present as a pattern of 30 individual muscle fibres per abdominal hemisegment. The second occurs during metamorphosis, when most larval muscles histolyse and are replaced by the *de novo* formation of adult muscles. Along with most of the other adult tissues, the somatic

musculature develop from pools of tissue specific progenitor cells associated with one of nineteen imaginal discs found in the larva (Hales et al. 2015). The indirect flight muscles (IFM's) in the thorax are the most widely used in the study of adult myogenesis. Consisting of the dorso-longitudinal (DLM) and dorso-ventral (DVM), they constitute up to 65% of the fly's total body mass (Bothe and Baylies 2016), and generate the force required to power the wing stroke during flight. Given their size, accessibility, and the fact they are not required for viability, they are an ideal model for the study of adult myogenesis (Jawkar and Nongthomba 2020; Nikonova et al. 2020).

1.3.1 Overview of larval body wall muscle development during embryogenesis

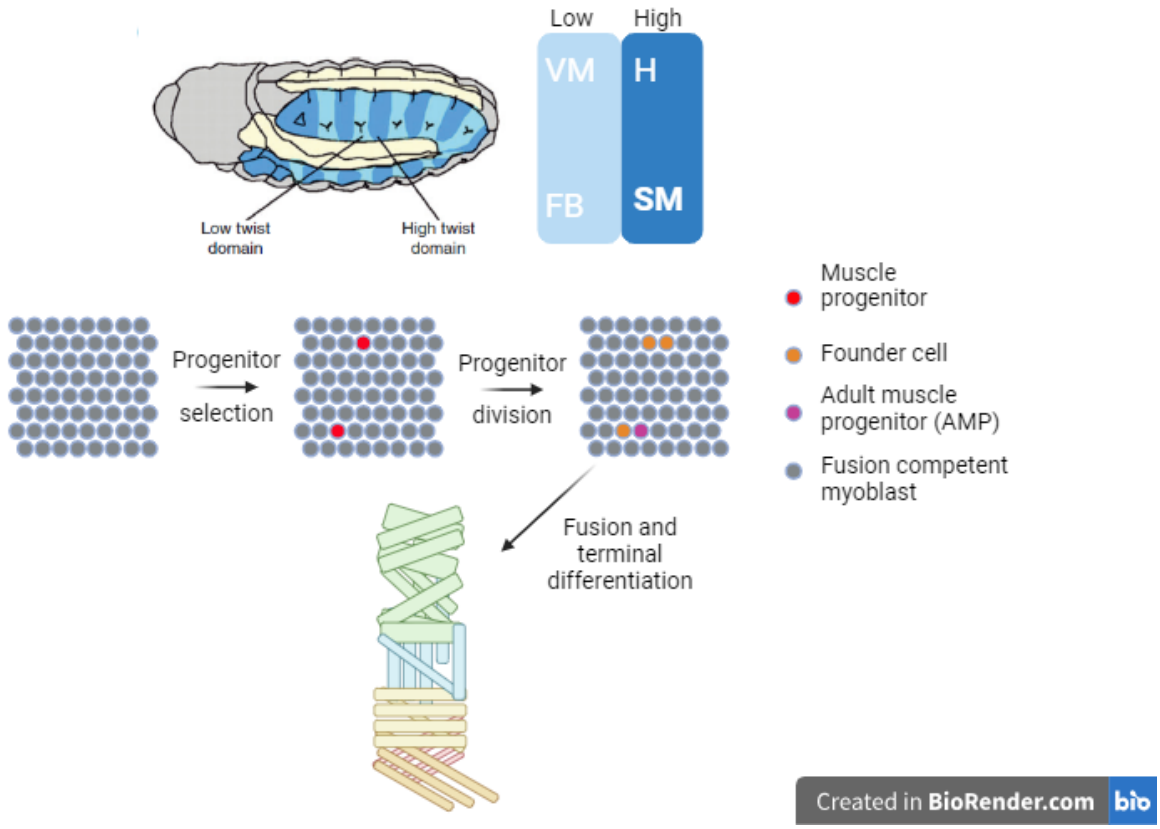


Figure 1. 2. Overview of larval somatic muscle differentiation during embryogenesis. Top: Schematic of a stage 10 *Drosophila* embryo showing mesodermal segments split into either into high (dark blue) and low (light blue) *Twist* domains. Somatic musculature develops from the high *Twist* domain. H, Heart; SM, somatic muscle; VM, visceral muscle; FB, fat body. Below: Schematic of SM domain to generate the larval somatic muscles. Progenitor cells are initially specified in the SM domain. These then undergo an asymmetric cell division to yield either two founder cells (FC), or one founder cell and one adult muscle progenitor (AMP). All other cells in the domain take on the fate of a fusion competent myoblast (FCM). Each FC seeds the formation of an individual muscle fibre. FCs fuse with FCMs, which ultimately gives rise to the fully differentiated larval muscle pattern. Image adapted from Dobi et al. (2015).

1.3.1.1 From mesoderm to muscle primordia

Somatic muscle is a derivative of the mesodermal germ layer, which is initially specified in the early syncytial blastoderm by the nuclear accumulation of the maternally deposited Dorsal protein. After cellularisation, Dorsal activates the expression of the transcription factors *Twist* and *Snail*, which is shortly followed by the invagination of these cells during gastrulation to form a layer underneath the overlying ectoderm. Both *Twist* and *Snail* are required for gastrulation and mesoderm formation, since both are disrupted by mutations in either gene, and double the mutant phenotype is more severe phenotype than either single mutant (Leptin and Grunewald 1990; Leptin 1991). *Twist* is required for establishing the transcriptional program required for mesoderm development by activating the expression of many mesodermal genes, including Mef2, a key transcription factor required in the muscle differentiation program, in the future somatic muscle primordia (Bour et al. 1995; Lilly et al. 1995; Taylor et al. 1995; Cripps et al. 1998; Sandmann et al. 2006; Sandmann et al. 2007). Conversely, *Snail* appears to recruit corepressor proteins upon DNA-binding and function as a transcriptional repressor of neur-ectodermal genes and help define the boundaries of the mesoderm domain (Leptin and Grunewald 1990; Leptin 1991). However, *snail* mutant embryos also display reduced expression of many mesodermal genes, suggesting that *snail* also contributes to the transcriptional activation of mesodermal genes (Rembold et al. 2014).

Multiple specialised tissues, including the somatic and visceral musculature, heart, and fat body, derive from the initially uniform population of cells within the mesoderm. Following gastrulation, the mesoderm is divided into segments and patterned along both the antero-posterior (AP) and dorso-ventral (DV) axes, which ultimately give rise to progenitors of these different cell types (Baylies et al. 1998) (**Fig. 1.2**). Importantly, a large component of mesodermal subdivision is reliant upon modulation of *Twist* expression into either high- or low-expressing domains within each para-segment of the post-gastrulation embryo. This subdivision influences the ultimate fate of the cells within each of those domains, with somatic and cardiac musculature forming within the high, and visceral musculature and fat body developing from the low-*Twist* domain.

Somatic muscle progenitors form in the anterior region of each segment, which maintains a high level of *Twist* expression (Lee and Frasch 2000; Tixier et al. 2010). The loss of *Twist* in post-gastrulation embryos inhibits somatic musculature development, while its overexpression throughout the entire mesoderm inhibits the formation of other mesodermal derivatives, such as the visceral musculature and fat body, which form from the low-*Twist* domain (Baylies and Bate 1996). Hence, although *Twist* is initially required to initiate the formation of the mesoderm prior to- and during gastrulation, its modulation post gastrulation is important for the proper mesodermal subdivision and somatic myogenesis.

1.3.1.2 Development of the somatic musculature from the high Twist domain

The somatic muscle domain is then patterned to give rise to the two functional 'building blocks' of muscle: the founder cells (FC's), and fusion competent myoblasts (FCM's). Both cell types are crucial for the formation of somatic muscles, however FC's appear to hold the genetic 'identity' for each individual muscle (Ruiz-Gómez et al. 1997; Ruiz-Gómez 1998). Each muscle fibre is a product of the fusion of many FCM's with a single FC, with the number of fusions varying depending on the muscle. Thus, each mesodermal segment must specify 30 founder cells, which can then fuse with surrounding FCM's.

FC and FCMs are specified initially through expression of the transcription factor *lethal of scute* (*l'sc*) in 18 clusters (termed equivalence groups) of 4-6 cells throughout the somatic muscle domain (Carmena et al. 1995). Initially, each cell within a cluster remains unspecified and retains the ability to either become a FC or FCM. However, through a subsequent process of Notch-signalling mediated lateral inhibition throughout these cells, the expression of *l'sc* is restricted to a single cell within each equivalence group. This single cell becomes a muscle progenitor, while the rest of the cells which lose *l'sc* expression become committed to an FCM cell fate. The progenitor can differentiate into two different cell types, each of which are important for different aspects of myogenesis. Following specification, the progenitor undergoes an asymmetric cell division, which yields either two FC's, which will 'seed' the formation of two independent muscle fibres, or one FC and one adult muscle progenitor (AMPs) (**Fig. 1.2**). Following the asymmetric division, the FC fuses with the surrounding cells which have acquired the FCM cell fate and differentiate into a characteristic syncytial

muscle fibre (**Fig. 1.2**). Conversely the AMPs, remain undifferentiated and instead are required for the formation of the adult musculature. The nature of progenitor cell asymmetric divisions appears to be dependent upon the unequal segregation of cytosolic Numb protein (Ruiz-Gómez and Bate 1997; Carmena et al. 1998). Numb inhibits the Notch signalling pathway, and in this context, active Notch signalling inhibits progenitor cell differentiation and promotes stem cell maintenance. Thus, in the context of asymmetric muscle progenitor division, daughter cells that inherit a large compliment of cytosolic Numb inhibit Notch and acquire a FC fate, while those which don't remain undifferentiated as AMPs.

Interestingly, upon their specification, FC's and FCM's lose, while AMPs maintain *T*wist expression, suggesting that, following specification of the somatic muscle lineage, *Twist* is not required throughout the differentiation process, indicating its role in establishing the muscle GRN occurs relatively early in myogenesis (Gunage et al. 2017). However, it does emphasise the importance of Twist in regulating the gene regulatory network that underpins muscle differentiation: It directly activates the expression of the transcription factor Mef2, which is first detected in the mesoderm shortly after gastrulation (Lilly et al. 1995; Taylor et al. 1995). Mef2 subsequently activates the expression of many target genes throughout muscle development, and is thought to be critical for coordinating the overall gene expression network required for normal muscle differentiation (Junion et al. 2005; Sandmann et al. 2006; Potthoff and Olson 2007; Taylor and Hughes 2017). Indeed, given Mef2, a pro-differentiation factor required for muscle differentiation is expressed so early in the myogenic pathway, this may point towards an important regulatory network required also to tightly control the normal differentiation pathway.

1.3.2 Formation of the indirect flight muscles (IFM's) during adult myogenesis

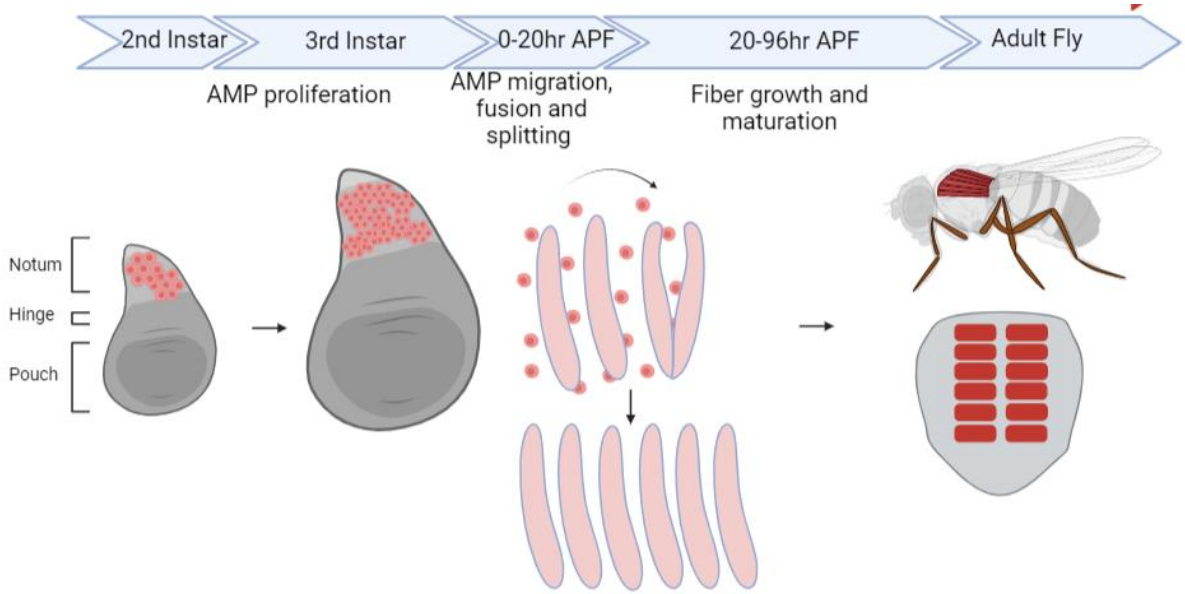


Figure 1. 3. Overview of dorso-longitudinal muscle (DLM) differentiation.
(from left to right). Adult muscle progenitors (AMPs) associated with the wing imaginal disc proliferate during the second and third larval stages. During metamorphosis, these AMPs migrate away from the disc and fuse to two pairs of 3 larval muscle templates. This is coupled with the splitting of each of these fibres, which gives rise to 6 nascent, immature muscles per hemithorax. The fibres grow and mature within the adult thorax to ultimately give rise to 6 fully differentiated DLM muscle fibres per hemithorax of the adult fly.

1.3.2.1 Adult somatic muscles derive from Adult Muscle Progenitors (AMPs)

The second wave of myogenesis generates the adult somatic musculature required for the many different functions within the adult fly. Of these, some of the most common muscles used as a model to study muscle differentiation *in vivo* are the indirect flight muscles (IFMs). These are located within the adult thorax and are comprised of two distinct sets of muscle fibres: the Dorso-longitudinal (DLMs), and the Dorso-ventral muscles (DVMs). There are 12 DLMs and 14 DVMs within the *Drosophila* adult thorax, which themselves are arranged symmetrically into a distribution of 6 DLM, and 7 DVM fibres in each hemithorax. As their names suggest, DLMs and DVMs are aligned in opposite directions along either the antero-posterior axis (DLMs) or the dorso-ventral axis (DVMs).

All adult somatic muscles arise from the AMPs, muscle stem cells that arise during embryogenesis and remain undifferentiated until they enter the differentiation program to form the adult somatic musculature throughout metamorphosis (**Fig. 1.3**) (Gunage et al. 2017; Laurichesse and Soler 2020). The AMPs that give rise to the IFM's

originate from the notum of the wing imaginal disc, a tissue of undifferentiated cells that contains precursors of the wing, and majority of the thoracic tissue within the adult fly (Tripathi and Irvine 2022). These AMPs give rise to both the IFMs, and the direct flight muscles (DFMs), which are smaller muscles in the thorax responsible for controlling finer wing movements during flight.

In the wing imaginal disc, an initial population of approximately 10 AMPs divide during larval stages and expands the AMP pool to approximately 2'500 cells (**Fig. 1.3**). The mechanisms underlying AMP proliferation during larval stages remains incompletely understood, Gunage et al (2014) suggested that AMP proliferation is dependent upon the plane of cell division with respect to the underlying wing disc epithelium: During early larval stages, AMPs divide symmetrically to amplify the AMP population, with all daughter cells maintaining their 'stemness' and ability to self-renew. However, during the third instar, AMPs divide asymmetrically, and post-mitotic daughter cells, which lose contact with the disc epithelium, are those which become committed to differentiation. The population of AMPs which continue to remain undifferentiated are thought to play a role in repair and regeneration of adult muscle tissue, analogous to that of vertebrate muscle satellite cells (Chaturvedi et al. 2017; Boukhatmi and Bray 2018). Interestingly however, a recent RNA-seq dataset from Zappia et al (2020) found five distinct clusters of AMPs that differ in their gene expression profile. Furthermore, spatial characterisation of these cell clusters within the disc did not correlate with that suggested by Gunage et al (2014). Given these clusters were initially characterised based upon differential gene expression, this does suggest some cell populations possess distinct characteristics that may reflect different cell populations being more or less differentiated. This suggests that it is likely there is a more complex regulatory network in these cells than the binary model proposed by Gunage et al (2014).

1.3.2.2 AMP-derived myoblasts fuse with larval muscle templates to form the mature DLM's

During metamorphosis, much of the pre-existing larval tissue histolyses, and the tissues of the adult fly begin to differentiate. Most adult somatic muscles form *de novo* through the fusion of differentiating AMPs. By contrast, the DLM's form through the fusion of AMPs with larval muscle templates: Three pairs of larval oblique muscles (DO1-3) in the second thoracic hemisegments of the larva resist histolysis and function

as templates, like that of embryonic muscle FCs, in 'seeding' the formation of the DLM fibres (Fernandes et al. 1991; Dobi et al. 2015; Schulman et al. 2015). In the same manner by which FCM's fuse with FC's during embryonic myogenesis, AMPs associated with the wing imaginal disc take on a function analogous to that of the embryonic FCM's: They first migrate from the disc to the DOM's which have resisted histolysis early during pupation (<8hr APF); before fusing to these DOM template 'founders' (Approx 12-24hr APF) (**Fig. 1.3**). This is coupled with the splitting of three DOM templates (per hemisegment) into six, which ultimately generates the immature myotube structures which will form the future six individual DLM fibres per hemithorax. While it remains relatively poorly understood the mechanisms which guide myoblast migration over the relatively long distance, compared with the short distances of a few cells in diameter during FC/FCM interactions during embryonic myogenesis, many of the proteins involved are required for myoblast fusion during both stages. This includes the critical transmembrane proteins Dumbfounded and Sticks-n-stones. This indicates a general level in conservation in the developmental pathways regulating fusion during embryonic and adult myogenesis (Schulman et al. 2015; Gunage et al. 2017).

Akin to its expression in early mesodermal tissue during larval somatic muscle differentiation during embryogenesis, not only is Mef2 required for the differentiation of the adult IFMs (Soler et al. 2012), but the pro-differentiation transcription factor is expressed throughout muscle differentiation. It is expressed in all of the AMPs associated with the wing imaginal disc of the third instar larva (Cripps et al. 2004; Soler and Taylor 2009). Moreover, the AMPs still remain undifferentiated and do not fuse with the larval muscle templates until after metamorphosis. Therefore, given Mef2's importance in activating muscle gene expression throughout development, how is this gene expression regulated in order to coordinate the development of such complex tissues that perform a wide variety of functions?

1.4 Myocyte-enhancer factor 2 (Mef2) transcription factors: critical regulators of coordinating gene expression during muscle development

1.4.1 Structure and conservation of Mef2 factors

Mef2 factors are members of the MADS (MCM1, Agamous, Deficiens, and Serum response factor)-box containing transcription factors, named after the first four factors

to be identified in this family (Shore and Sharrocks 1995; Black and Olson 1998). There are four *Mef2* paralogues in vertebrates (*A-D*), while *Drosophila, C. elegans* and *S. cerevisiae* possess only a single copy (**Fig. 1.4**). Mef2 factors possess three distinguishable structural domains. At the extreme N-terminus is a 57 amino acid long MADS-box, immediately followed by a 29 amino acid Mef2-domain not present in other MADS-box containing proteins (Black and Olson 1998). The MADS/Mef2 domain is followed by a significantly longer, C-terminal transactivation domain. The MADS/Mef2 domains are highly conserved both between human paralogues and *Drosophila* Mef2 (*dMef2*). In contrast, there is significant sequence divergence within the transactivation domain **(Fig. 1.4)** (Potthoff and Olson 2007).

The MADS/Mef2 domain is responsible for mediating multiple aspects of Mef2 function (Potthoff and Olson 2007; Taylor and Hughes 2017). Despite subtle protein sequence variations in the MADS/Mef2 domain, all Mef2 factors bind to the same CTA(A/T)$_4$TAG motif (Black and Olson 1998; Estrella et al. 2015). Some data suggests the variation in the MADS/Mef2 domain may have subtle effects on the transcriptional activity of different Mef2 orthologues in vertebrates: Mef2B possesses a glutamine (Q) at amino acid 14, whereas all other paralogues contain a glutamic acid (E), and a Q14E mutation in Mef2B increased DNA-binding by two-fold (Molkentin et al. 1996b). The MADS/Mef2 domain is also required for dimerization, as well as the recruitment and interaction with cofactors that can modulate Mef2 activity (Molkentin et al. 1996a; Han et al. 2005; He et al. 2011). In cell culture, the MADS/Mef2 domain of Mef2C is sensitive to inactivating mutations via the disruption of either DNA-binding, dimerization, or both, while some mutants retain these abilities but lack the ability to activate target gene expression (Molkentin et al. 1996a). The latter possibly implies that additional coactivator proteins are required to interact with bound Mef2 to activate target gene expression. Thus, given its multitude of roles in mediating Mef2 function, it is not surprising that the MADS/Mef2 domain is highly conserved from vertebrates to *Drosophila*, and highlights the strength of *Drosophila* not only as a model system to investigate myogenesis, but also the function of Mef2. Indeed, *in* vivo, mutations in the MADS box of *Drosophila* Mef2 results in loss of DNA binding, while also causing embryonic lethality accompanied with severe muscle phenotypes (Nguyen et al. 2002). Moreover, truncated mutants containing an in-tact MADS/Mef2 domain but lacking the transactivation domain lack the ability to activate a transcriptional reporter,

despite retaining the ability to bind DNA, highlighting the requirement of all three domains for proper Mef2 function (Molkentin et al. 1996a; Molkentin et al. 1996b).

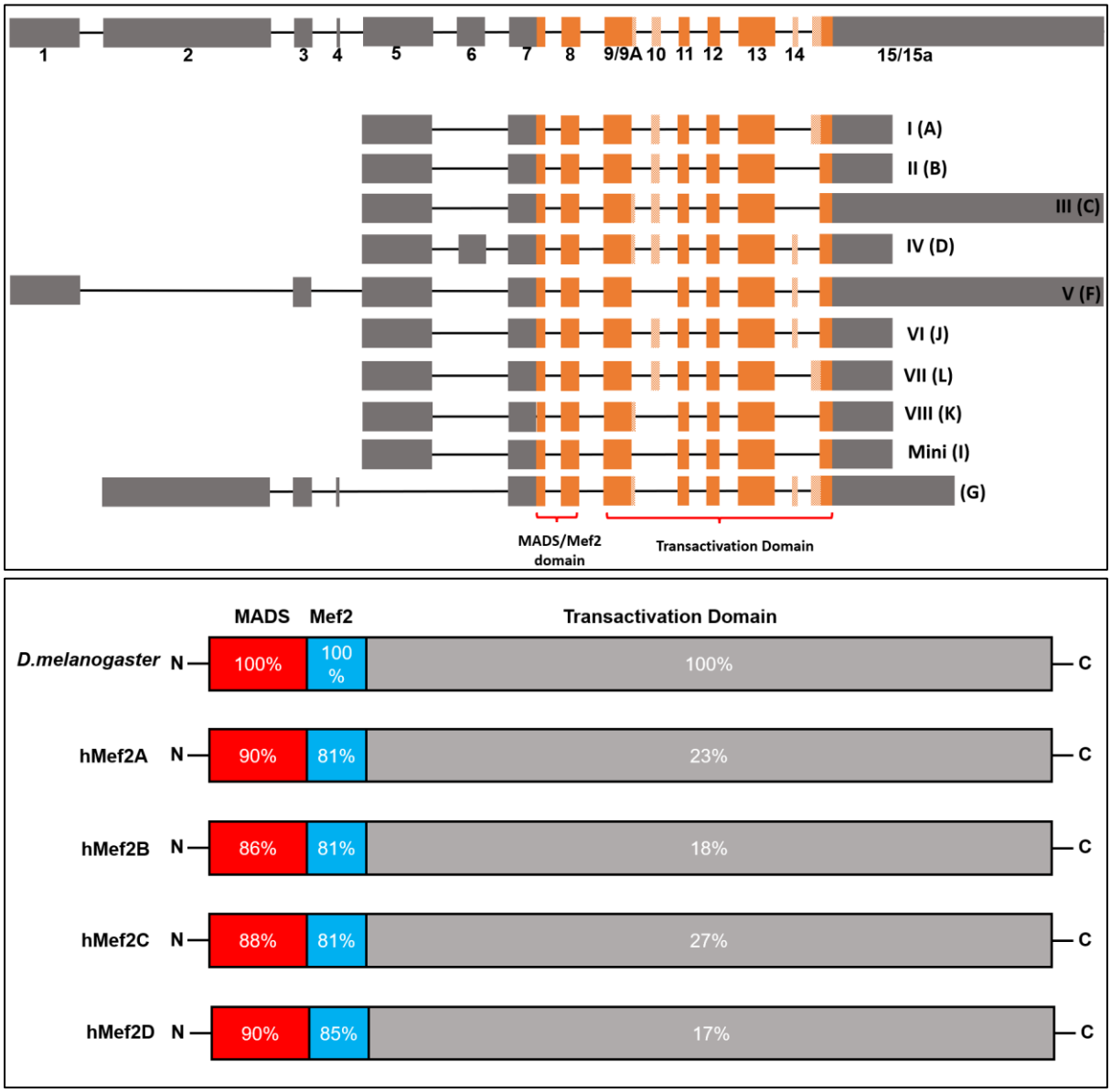


Figure 1. 4. *Mef2* gene structure and Mef2 protein organisation and conservation
(Top). Schematic of the intron-exon structure of the *Drosophila* Mef2 locus. There are many different isoforms encoded by the locus, which vary by transcription start site utilised, alternative splicing, and alternative poly-adenylation. The region encoding the MADS/Mef2 and transactivation domains are indicated. Non-coding regions are in grey, coding exons in orange. Alternatively spliced coding exons are patterned orange/white. (Bottom) Mef2 protein domain structure and conservation. Mef2 proteins have three domains: the MADS-Box (red), Mef2 (blue), and Transactivation (grey) domains. The MADS/Mef2 domains are highly conserved between *Drosophila* and human Mef2 genes, while the Transactivation domains are more divergent. Percentages indicate sequence identity compared to *Drosophila* Mef2.

1.4.2 *Mef2* is required for *Drosophila* muscle development

Mef2 function in vertebrates is complicated by the presence of four paralogues with overlapping patterns of expression, with all genes expressed in the developing skeletal muscle lineages of the mouse embryo (Edmondson et al. 1994; Molkentin et al. 1996b). *Mef2a* or *Mef2b* mutants have no distinguishable skeletal muscle phenotype (Potthoff et al. 2007a; Potthoff et al. 2007b), while global *Mef2c* knockouts die too early to analyse skeletal muscle phenotypes (Lin et al. 1997b). Conditional, skeletal muscle specific *Mef2* knockouts show either disorganised, fragmented skeletal muscle fibres (Potthoff et al. 2007a), or disrupted growth and glucose metabolism (Potthoff et al. 2007a; Anderson et al. 2015), although the mechanisms remain unclear. Interestingly, in cultured myoblasts that Mef2 paralogues regulate the expression of distinct, non-overlapping target genes (Estrella et al. 2015), possibly suggesting independent roles of Mef2 paralogues in vertebrate myogenesis may also exist.

The simpler *Drosophila* system has been fundamental in identifying the role of Mef2 during myogenesis. Mef2 null mutants die during embryogenesis due to the lack of differentiated somatic muscle, as well as visceral muscle and heart abnormalities (Bour et al. 1995; Lilly et al. 1995; Ranganayakulu et al. 1995). Furthermore, DNA-binding mutants generated through mutagenesis of the MADS/Mef2 domain are also embryonic lethal and die with severe muscle phenotypes similar to that of null mutants generated previously (Nguyen et al. 2002). This observation also emphasises the functional importance and sensitivity of the conserved MADS/Mef2 domain. In Mef2 null embryos, the myoblast population is specified normally, however fail to fuse and differentiate further into syncytial muscle fibres (Bour et al. 1995). Given the embryonic lethality, RNA-interference (RNAi) has been used to understand the requirement for Mef2 during adult myogenesis: Mef2 knockdown in the AMPs inhibits IFM formation, with phenotype severity varying with knockdown efficiency and the line used (Bryantsev et al. 2012; Soler et al. 2012). Thus, Mef2 is necessary for *Drosophila* somatic muscle development.

1.4.3 Mef2 activates the myogenic transcriptional program.

Mef2 factors play a central role in the activation of the gene expression network throughout myogenesis (Sandmann et al. 2006). Moreover, Mef2 expression is regulated by other pro-myogenic factors. In *Drosophila*¸ Mef2 expression is directly

activated by Twist (Nguyen et al. 1994; Lilly et al. 1995; Taylor et al. 1995; Cripps et al. 1998). The role of Twist is analogous to the myogenic regulatory factors (MRFs), transcription factors of which there are four in mammalian systems (MyoD, Myf5, myogenin, MRF4). Like the function of Twist, the MRFs function upstream of Mef2 in the myogenic gene expression network: They are expressed before Mef2 in the developing mouse embryo (Taylor and Hughes 2017), and have been shown to be able to directly activate *Mef2C* expression (Wang et al. 2001; Dodou et al. 2003).

Mef2 can also can ectopically induce muscle gene expression. In mammalian cell culture, Mef2 cooperates with the MRF's to enhance the myogenic conversion of 10T1/2 fibroblasts into *MHC*-expressing myoblasts (Molkentin et al. 1995; Molkentin and Olson 1996). Interestingly, the capacity for vertebrate MRF's alone to induce myogenic conversion of fibroblasts is not shared by Mef2C, which requires the MRF's to activate muscle gene expression (Molkentin et al. 1995). However, expression of a dominant-negative Mef2A lacking its transactivation domain blocks the differentiation of C2C12 myoblasts and the MyoD-induced myogenic conversion of 10T1/2 myoblasts (Ornatsky et al. 1997), suggesting Mef2 and vertebrate MRFs function cooperatively to activate muscle gene expression. In contrast to vertebrates, *Drosophila* Mef2 can activate the muscle gene expression program on its own: Ectopic expression of *Mef2* in the ectoderm of the *Drosophila* embryo induces expression of muscle genes including *Tropomyosin* (*Tm1*) and *Myosin heavy chain (MHC)* (Lin et al. 1997a; Gunthorpe et al. 1999), as well as other *Mef2* target genes such as *Holes in muscle (Him*) (Sandmann et al. 2006; Liotta et al. 2007). Moreover, *Mef2* overexpression in the L3 AMPs induces premature expression of a tau-GFP reporter under the control of MHC regulatory sequence (Soler and Taylor 2009).

High throughput analysis of Mef2 target gene expression has highlighted the important role of Mef2 in coordinating gene expression throughout myogenesis. Firstly, constant development of the *Drosophila* REDfly database continues to experimentally verify CRMs which modulate the expression of all *Drosophila* genes (Gallo et al. 2011; Rivera et al. 2019). This complements studies investigating Mef2 targets: In *Drosophila* embryos, *in silico* predictions (Junion et al. 2005) and Chromatin immunoprecipitation (ChIP) (Sandmann et al. 2006) have found over 200 Mef2-target genes, Interestingly, Sandmann et al. further categorised Mef2 targets into those expressed in three distinct temporal patterns during embryonic muscle differentiation.

Furthermore, Mef2 targets also respond differently to levels of Mef2, with some targets requiring more Mef2 to activate their expression (Elgar et al. 2008). Mef2 target genes also show enrichment for certain core promoter elements (CPEs) compared to other genes transcribed by RNA polymerase II. However, the lack of distinguishable effect upon mutation of these overrepresented sequences in Mef2-target genes may indicate a general level of redundancy in CPEs throughout the genome, rather than an additional level of regulation which controls Mef2 target gene activation (Bu and Cripps 2022). Analysis of Mef2 target gene expression has thus far been limited to embryogenesis, however it is likely similar characteristics of Mef2 target gene expression may apply during adult myogenesis, given that some known Mef2 targets, such as *βTub60D*, *Tm1* and *MHC* are known to be expressed at different times during *Drosophila* flight muscle development (Spletter et al. 2018; Zappia et al. 2020). Nevertheless, an important question resides in the expression patterns of Mef2 target genes. How is the expression of such a vast range of transcriptional targets' coordinated to allow the correct formation and patterning of muscle fibres? These data suggest some layer of spatio-temporal regulation of Mef2 activity.

1.4.4 Mef2 is regulated by multiple mechanisms throughout myogenesis

Many lines of evidence point to the requirement for spatiotemporal regulation of Mef2 activity for normal muscle differentiation. Firstly, *Mef2* is expressed in both the embryonic mesoderm post-gastrulation (Lilly et al. 1995; Taylor et al. 1995), and the AMPs residing in the L3 wing imaginal disc (Cripps et al. 2004; Soler and Taylor 2009; Spletter et al. 2018; Zappia et al. 2020). In both cases, this is many hours before the activation of Mef2-target genes and the associated differentiation of Mef2-expressing cells. Secondly, Mef2 target genes are expressed at different stages of muscle differentiation during embryonic muscle differentiation (Sandmann et al. 2006); and thirdly, overexpression of Mef2 disrupts normal muscle differentiation: In the embryo, this manifests in the presence of unfused, *MHC* expressing cells as well as patterning defects (Gunthorpe et al. 1999); while during adult muscle differentiation, Mef2 overexpression in L3 AMPs causes them to ectopically express a tau-GFP reporter. The expression of this reporter is controlled by regulatory sequence for *MHC*, a known Mef2 target gene (Soler and Taylor 2009).

Existing evidence, mainly using *in vitro* models, points towards multiple regulatory inputs which may control the expression, and activity of Mef2 during muscle differentiation. Importantly, these may either positively, or negatively regulate Mef2 at both the transcriptional, as well as the post-transcriptional level. At the transcriptional level, Mef2 can auto-activate its own expression through the presence of Mef2-binding sites in its CRM (Cripps et al. 2004). To overcome excessive, uncontrolled *Mef2* expression which may induce the aforementioned overexpression phenotypes, *Mef2* expression may be controlled by: the presence of transcriptional repressors such as *Zinc-finger homeodomain 1 (zfh1*), a transcription factor which represses *Mef2* transcription (Postigo et al. 1999; Boukhatmi and Bray 2018); as well as microRNA's such as mir-92B, which binds to the 3' UTR of *Mef2* transcripts and represses its translation (Chen et al. 2012). Interestingly, both genes are activated by Mef2, suggesting that a finely tuned balance of positive and negative regulatory inputs may help control *Mef2* transcription. At the co-transcriptional level, much of the aforementioned diversity observed within the TAD of Mef2 homologues is a result of multiple different isoforms expressed in both *Drosophila* and vertebrates (Potthoff and Olson 2007; Taylor and Hughes 2017). For example in *Drosophila*, at least 10 different isoforms have been identified (Gunthorpe et al. 1999; Lin and Baines 2019) with all of the coding sequence variation found outside of the MADS/Mef2 domain. However currently, there is very little understanding as to the expression, or function of different isoforms during myogenesis.

These transcriptional mechanisms however do not explain how Mef2 protein, present in undifferentiated myoblasts, does not precociously induce muscle gene expression and thus affect normal muscle formation. Thus far, *in vitro* evidence suggests a major regulator of Mef2 function may be its direct physical interaction with several co-regulator proteins. Although transcription factor-protein interactions are fundamental for gene expression, being required for the recruitment of RNA polymerase and the initiation of transcription (Shokri et al. 2019), the physical interaction of TFs with additional coactivators and corepressors can regulate TF activity in the context of target gene activation (Chen and Pugh 2021). In various contexts, Mef2 has been shown to directly interact with different proteins, including other transcription factors, kinases, signalling molecules, and chromatin modifiers (Dong et al. 2017; Shokri et al. 2019; Luck et al. 2020). These protein-protein interactions could potentially be

involved in multiple mechanisms that regulate Mef2 target gene activation, such as: direct modification of the nearby chromatin environment, facilitating the formation of coactivator and/or corepressor complexes, or modulating the post-translational modifications (PTMs) present directly on Mef2. For example, different PTMs have been shown to have different effects on Mef2 transcriptional activity *in vitro*, including either: enhancing transcriptional activation through acetylation(Ma et al. 2005; Angelelli et al. 2008), or phosphorylation (Ornatsky et al. 1999; Zhao et al. 1999; Kato et al. 2000; Wu et al. 2000); or the repression of Mef2 activity through phosphorylation-dependent sumoylation (Grégoire et al. 2006; Kang et al. 2006).

Of particular interest in the context of Mef2 regulation are the Class IIa Histone deacetylases (HDACs) and the histone acetyltransferase (HAT) p300. As will be introduced below, Class IIa HDACs can potential repress Mef2 activity *in vitro*, while p300, in the most general concept of gene expression regulation, is believed to predominantly function as a transcriptional coactivator. Significantly, both proteins have been shown to physically interact with the same region of Mef2. Thus, this may present the possibility of competition for binding between different regulator proteins, which may have contrasting roles in regulating Mef2 function *in vivo*. However, as will be described in the following sections, their role in normal muscle differentiation *in vivo*, as well as their possible functional relationships with Mef2, remain poorly understood.

1.5 Histone Acetyltransferases (HATs) and Class IIa Histone Deacetylases (HDACs): candidate regulators of Mef2 activity during *Drosophila* muscle differentiation *in vivo*

1.5.1 Histone Acetyltransferases (HATs) and Histone Deacetylases (HDACs): two superfamilies of diverse gene regulatory proteins

HATs and HDACs are two superfamilies of diverse proteins heavily implicated in the regulation of gene expression. There are many different classes/subfamilies of HATs and HDACs within their respective superfamilies, categorised based upon evolutionary origin, sequence similarity and substrate specificity. There are five subfamilies of HATs (HAT1,GCN5,MYST,Rtt109, and p300/CBP) (Marmorstein and Zhou 2014), and five classes of HDACs (I, IIa, IIb, III, IV) (Yang and Seto 2008).

Members of the HAT and HDAC superfamilies generally share a relatively well-conserved catalytic domain, which are then flanked by large regions with either poor sequence homology, or the presence or absence of additional domains that mediate substrate specificity and class-specific functions (Haberland et al. 2009; Marmorstein and Zhou 2014).

Functionally, as their name suggests, HATs and HDACs are most broadly associated with being key modifiers of the chromatin landscape and subsequent transcriptional regulation. The canonical mechanism by which HATs and HDACs function is by modulating the acetylation status on histone tails: lysine acetylation by HATs is associated with active gene transcription by promoting the formation of 'open chromatin'. Acetylation neutralises the lysine side chains' positive charge, thus weakening the interaction between histones and the associated DNA, while also providing binding sites for other proteins that contain bromodomains – domains which recognise and bind to acetylated lysine residues (Bannister and Kouzarides 2011; Narita et al. 2019). However, HATs and HDACs generally do not possess intrinsic DNA-binding ability, and thus must be recruited to DNA via directly interacting with other proteins, especially TFs (Yang and Seto 2007).

However, HATs and HDACs have also been shown to regulate gene expression in other ways. HATs and HDACs can target non-histone proteins, including transcription factors. Such post-translational modification (PTMs) subsequently can impact protein function by a variety of mechanisms (Narita et al. 2019). For example, in the context of transcription factor regulation, acetylation and deacetylation of the tumour suppressor p53 regulate its DNA-binding and subsequent coordination of the cellular response following DNA damage (Reed and Quelle 2014), while acetylation of Mef2 by the p300 HAT enhances its transcriptional activity (Ma et al. 2005; Angelelli et al. 2008). Interestingly, the Class IIa HDACs are unique within the HDAC superfamily because they appear to be catalytically inactive enzymes, displaying limited activity against acetyl-lysine substrates *in vitro* (Lahm et al. 2007; Bottomley et al. 2008; Jones et al. 2008). A single residue of vertebrate HDAC4 may be responsible for this interesting loss in enzymatic activity. Compared to class I HDACs, Class IIa's possess a histidine to tyrosine mutation which, when mutated back to a histidine, restores the catalytic activity of HDAC4 to 1000-fold higher than that of wild-type HDAC4 (Lahm et al. 2007). Despite this, Class IIa HDACs are well-established repressors of gene

expression and may operate through alternative, albeit functionally linked mechanisms. For example, they can interact with class I HDACs via the deacetylase domain, indicating they may function in repressor complexes (Fischle et al. 2002); or possibly by competing with other transcriptional coactivators for transcription factor binding, required for gene activation (Clocchiatti et al. 2011).

Two classes in particular, the Class IIa HDACs, and the p300/CBP family of HATs may hold significance in the regulation of Mef2 target gene expression due to their ability to physically interact with Mef2 in different contexts. This will be described in the following sections.

1.5.2 Class IIa HDACs and the p300/CBP HATs as candidates for regulating Mef2 during *Drosophila* myogenesis

In *Drosophila*, there is only one Class IIa HDAC (HDAC4), and one p300/CBP homologue (nejire), compared to four of the former in vertebrates (HDAC4,5,7,9), and two of the latter (p300 and CBP) (Yang and Seto 2007). Therefore, akin to only possessing a single Mef2, the less possible functional redundancy among these two protein families in *Drosophila* also makes it an ideal model system to analyse their function during muscle differentiation *in vivo*.

Based on their ability to physically interact with Mef2 in various contexts, Class IIa HDACs and the p300 HAT family are strong candidates for Mef2 regulatory proteins *in vivo*. In the case of their interaction with Mef2, Class IIa HDACs interact via a well-characterised Mef2-binding domain, which is not found in other HDAC classes (Han et al. 2005; Haberland et al. 2009). While there is no unique Mef2 interacting domain in p300, it does contain a C-terminal TAZ2 domain which is responsible for p300 interacting with many different proteins (He et al. 2011). Importantly, p300 has been shown to physically interact with Mef2 via this domain (He et al. 2011). Moreover in *Drosophila,* the p300 homologue, *nejire* has been shown to physically interact with Mef2 in a yeast-2-hybrid assay (Lin and Baines 2019).

Interestingly, both p300 and Class IIa HDACs interact with the same region of Mef2 in using Mef2A/p300 (He et al. 2011) and Mef2B/HDAC9 (Han et al. 2005) crystal structures. Although structures have not been determined for other Mef2/HDAC/HAT complexes, the interaction domains of Class IIa HDACs, p300, and Mef2 are well conserved between vertebrate and *Drosophila* proteins. This suggests both the

potential functional significance of these interactions, and the suitability to use *Drosophila* as a model to probe Mef2 protein interactions in the context of myogenesis. However, as will be discussed in the next two sections, there is evidence implicating potentially contrasting roles of these proteins in regulating Mef2 during muscle differentiation: In the broadest sense, p300/CBP HATs are associated with being transcriptional coactivators, facilitating the activation of target gene expression, while class IIa HDACs are associated with transcriptional corepression, shutting down transcriptional activity (Yang and Seto 2007). Thus, in this context, differential and/or competitive binding of Mef2 by either p300/CBP or Class IIa HDACs could be a mechanism that regulates Mef2 activity (**Fig. 1.5**). However, relatively little is currently understood as to the functions of these two protein classes during both muscle differentiation, and in Mef2 regulation *in vivo*.

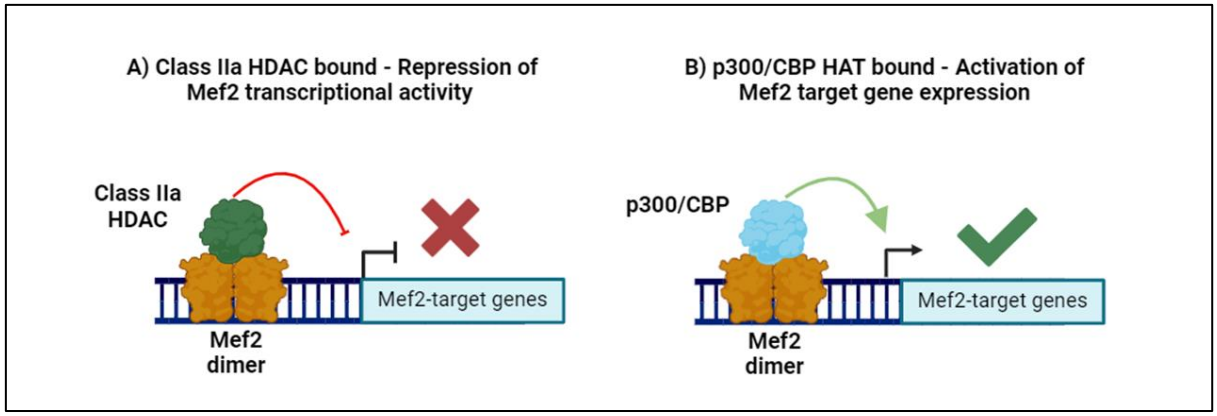


Figure 1. 5 Theoretical consequences of differential binding of p300 and Class IIa HDACs on Mef2 transcriptional activity

When physically interacting with a DNA-bound Mef2 dimer, Class IIa HDACs repress Mef2 transcriptional activation of target genes (**A**), while p300/CBP activates Mef2 transcriptional activity and facilitates the expression of Mef2 target genes (**B**).

.

1.5.3 Evidence for the role of p300/CBP in muscle differentiation and Mef2 regulation

Generally, p300 is associated with Mef2 transcriptional activation, which was first suggested following the observation that p300 can augment the activation of a Mef2-dependent reporter (Sartorelli et al. 1997). Later studies showed that p300 is responsible for both the acetylation of the Mef2C's TAD (Ma et al. 2005) and K4 in the MADS/Mef2 domain (Angelelli et al. 2008), both of which increase its DNA-binding and transcriptional activity in muscle cells lines. Interestingly this p300-mediated

acetylation of Mef2 appears to be correlated with the onset of differentiation, since acetylation of these residues occurs only in differentiating cells (Ma et al. 2005; Angelelli et al. 2008). Hence, this potentially suggests a mechanism whereby acetylation of Mef2 through p300 may stimulate Mef2 DNA binding and subsequent activation of target genes. Separate experiments have also shown that mutations within the p300-interacting domain renders Mef2 unable to activate a Mef2-induced reporter construct (Molkentin et al. 1996a). Therefore, it is possible these mutations inactivate Mef2 by rendering it unable to interact with p300, if it were a critical cofactor required for Mef2 function. However, an investigation as to the direct functional consequences of any such mutations have yet to be investigated *in vivo*.

However, given that p300 is a general HAT with the ability to interact with multiple transcription factors, it remains possible that p300 may also regulate myogenesis through alternative mechanisms (Holmqvist and Mannervik 2013). Thus far, over four hundred different interacting partners of p300 (and its homologue CBP) have been identified (Bedford et al. 2010; Dancy and Cole 2015). They have also been reported to directly acetylate over seventy different target proteins, in addition to the core histone proteins for which they earned their name (Wang et al. 2008). Currently, while no investigation into p300 function in the context of Mef2 during *Drosophila* muscle differentiation has been done, knockdown of *nej* in embryonic musculature is lethal (Schnorrer et al. 2010), indicating that it may play an important role during myogenesis *in vivo*. Interestingly however, another study found that nej inhibits Mef2-dependent transcription in cultured *Drosophila* cells, albeit not in the context of muscle differentiation (Lin and Baines 2019). Therefore, it remains an open question as to what the function of Nej is in the context of Mef2 regulation.

1.5.4 Class IIa HDACs regulate Mef2 during muscle differentiation *in-vitro.*

By contrast to p300/CBP, Class IIa HDACs have largely been implicated in the repression of Mef2 activity *in vitro,* making them ideal candidate proteins for down-regulating Mef2 activity during myogenesis *in* vivo. Co-expression of either HDAC4 (Miska et al. 1999; Wang et al. 1999; Lu et al. 2000a) or HDAC5 (Lemercier et al. 2000; Lu et al. 2000b) inhibits expression of a Mef2-dependent reporter in cell lines. Moreover, HDAC5 mutants lacking the Mef2-binding domain fail to inhibit similar

reporter gene activity, indicating that Class IIa HDACs inhibit Mef2-dependent transcription through directly interacting with Mef2 (Lemercier et al. 2000). However, this negative regulation does not appear to be through obstructing Mef2 binding to DNA, since HDAC4-Mef2 complexes can still bind to DNA (Lu et al. 2000b; Chan et al. 2003). Furthermore, the crystal structure of a Mef2B/HDAC9 complex is bound to DNA, further supporting that HDACs may inhibit Mef2 when bound to DNA (Han et al. 2005). Additional *in vitro* work further linked Class IIa HDAC's ability to repress Mef2 dependent transcription to models of muscle differentiation *in vitro*. HDAC4 and HDAC5 can both inhibit the MyoD-induced myogenic conversion of 10T1/2 fibroblasts (Lu et al. 2000b; McKinsey et al. 2000a; McKinsey et al. 2000b), and the differentiation of C2 myoblasts into myotubes (Lu et al. 2000b; Miska et al. 2001). This inhibition of myogenesis is also dependent upon an in-tact Mef2-binding domain, which is required for both the HDAC5-mediated inhibition of myogenic conversion of 10T1/2 fibroblasts (Lu et al. 2000b), and for the HDAC4-mediated inhibition of C2 myoblasts (Miska et al. 2001). Furthermore, as well as the Mef2-binding domain, Miska et al showed that the deacetylase domain of HDAC4 was also required for this inhibition, suggesting that, although lacking intrinsic deacetylase activity, the conserved HDAC domain of Class IIa HDACs may be required for the regulation of Mef2 activity.

By contrast to p300/CBP HATs, which are thought to be exclusively nuclear proteins (Narita et al. 2019), Class IIa HDAC function during muscle differentiation *in vitro* has also been linked to its subcellular localisation. Calcium-calmodulin dependent kinase (CamK)-mediated phosphorylation of HDAC5 induces 14-3-3 binding and subsequent nuclear export (McKinsey et al. 2000a). Mutation of either of these residues impedes 14-3-3 binding and inhibits nuclear export, whereas mutations in both residues almost completely abolishes nuclear (McKinsey et al. 2000a; McKinsey et al. 2000b). However, 14-3-3 binding alone is not sufficient to drive nuclear export, since HDAC4 and HDAC5 mutants possessing intact 14-3-3 binding sites, but lacking the NES, resist nuclear export and maintain a predominantly nuclear localization, even in the presence of active CamK (McKinsey et al. 2001; Wang and Yang 2001). Importantly, 14-3-3 binding is accompanied by a reduction in the capability of HDAC5 to interact with Mef2, even when nuclear export is blocked (McKinsey et al. 2000b), suggesting that, in addition to the spatial redistribution of HDAC5, the act of 14-3-3 binding itself inhibits the ability for HDAC5 to interact with Mef2. Importantly, activated CamK rescues the

inhibition of Mef2-reporter gene activity by HDAC4 and 5 (Lu et al. 2000a; Miska et al. 2001), as well as both the HDAC-mediated inhibition of both the myogenic conversion of fibroblasts (Lu et al. 2000b), and differentiation of C2C12 myoblasts (Miska et al. 2001). Mutant HDAC5 resistant to 14-3-3 binding retains the ability to inhibit myogenic conversion of fibroblasts, even in the presence of activated CamK (McKinsey et al. 2000b). Given also that endogenous HDAC5 localises to the nucleus in undifferentiated myoblasts, before being redistributed to the cytoplasm upon differentiation into syncytial myotubes (McKinsey et al. 2000a), these data ultimately led to a model for class IIa HDAC function in differentiating muscle (**Fig. 1.6**). In undifferentiated myoblasts, Class IIa HDACs localise to the nucleus where it physically interacts with Mef2 on DNA, negatively regulating its transcriptional activity. Then, upon differentiation, Class IIa HDACs are redistributed to the cytoplasm where they no longer can inihibit Mef2, therefore enabling Mef2 to thus coordinate the transcriptional program required for proper muscle differentiation. This model, however is complicated by the observation that HDAC4 shows the reciprocal localisation, where it is cytoplasmic in undifferentiated myoblasts and nuclear in myotubes (Miska et al. 2001). It is possible therefore that different Class IIa HDACs possess different functions throughout muscle development *in vivo*.

Finally, one additional possible mechanism by which Mef2 activity may be regulated is through a feedback mechanism between Mef2 and Class IIa HDACs: Mef2 has been shown to activate the expression of *HDAC9 in vitro* (Haberland et al. 2007) *.* This, along with the aforementioned mechanisms potentially regulating *Mef2* expression, may help to fine tune the transcriptional Mef2 and thus prevent precocious expression of many of its target genes throughout muscle differentiation *in vivo*

Nevertheless, despite this evidence supporting a role for Class IIa HDACs in regulating Mef2 *in vitro*, there has thus far, only been limited investigation into their role during muscle differentiation *in vivo*. *HDAC4* knockout mice display no obvious skeletal muscle phenotypes (Vega et al. 2004), although Class IIa HDACs have been shown to regulate exercise-induced fibre-type switching in mice following exercise (Potthoff et al. 2007b). Such analysis of mammalian Class IIa HDAC function is complicated by the presence of four family members, for which may have a level of functional redundancy. However, because of only a single Class IIa HDAC, *Drosophila*

represents an incredibly powerful model to investigate their function during muscle differentiation *in vivo*.

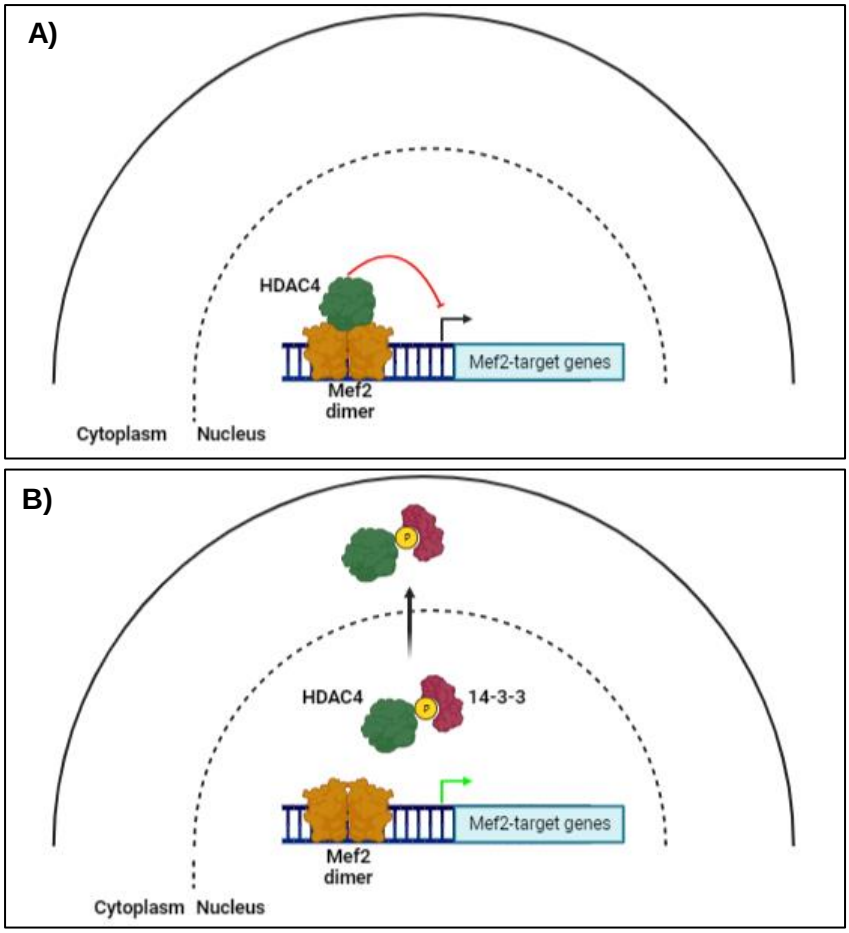


Figure 1. 6. An in-vitro model for Class IIa HDAC-mediated regulation of Mef2 during muscle differentiation

(A) In the nucleus, Class IIa HDACs can physically interact with Mef2, while it is bound to DNA, and repress expression of Mef2 target genes. (**B**) HDAC activity is regulated by phosphorylation-dependent changes in its subcellular localisation Upon phosphorylation, HDACs are bound by 14-3-3 proteins and actively exported out of the nucleus, where they can no longer inhibit Mef2 activity.

1.5.5 Class IIa HDACs can regulate Mef2 *in vivo*

Class IIa HDAC function has been investigated in the context of Mef2 regulation during development and other physiological processes. During skeletogenesis, *HDAC4*

overexpression inhibits normal bone development, a process dependent upon Mef2 function and for which the phenotype phenocopies that of Mef2 loss-of-function (Vega et al. 2004; Arnold et al. 2007). In the heart, HDAC5 and HDAC9 share partially redundant phenotypes in the regulation of stress-induced cardiac hypertrophy, a process associated with Mef2 hyperactivation (Zhang et al. 2002; Chang et al. 2004). Meanwhile, *HDAC7* mutants are embryonic lethal caused by loss of vascular integrity. and this phenotype is accompanied by upregulated levels of Mef2-target genes (Chang et al. 2006).These data not only implicate Class IIa HDAC function in a wide variety of tissues and context, but also emphasise their importance, and the subsequent potential consequences to the dysregulation of Mef2 during normal development and. However, although the role of the Class IIa HDACs in regulating Mef2 during skeletal muscle differentiation have been established *in vitro*, there has been no investigation as to the consequences of HDAC loss-of-function on Mef2 regulation during skeletal muscle development, either *in vitro* or *in vivo*.

However, Mef2-independent functions of *HDAC4* have also been described in *Drosophila*. Firstly, a rough-eye screen found multiple other genes which genetically interact with *HDAC4* (Schwartz et al. 2016). In the brain, both *HDAC4* overexpression disrupts long-term memory formation (LTM) (Fitzsimons et al. 2013). However, although overexpression causes a redistribution of Mef2 within neuronal nuclei, it did not cause any significant transcriptional changes (Schwartz et al. 2016); while mutant *HDAC4* which cannot interact with Mef2 similarly inhibits LTM, while Mef2 loss-of-function does not phenocopy that of *HDAC4* overexpression (Main et al. 2021). Moreover, *HDAC4* functions in the fat body to regulate lipolysis through interacting with another transcription factor, FOXO (Wang et al. 2011; Choi et al. 2015). Therefore, while roles for HDAC4 in regulating Mef2 activity have been established, it could likely have many other roles in different developmental and physiological contexts in *Drosophila in vivo*.

1.6 Techniques to study gene function and regulation: The CRISPR-Cas9 revolution

1.6.1 The desire to study gene function at an endogenous level

Largely because *Drosophila* has been at the forefront of genetic and developmental research for over a century, many technologies exist to study gene function in a variety of biological contexts. In particular, the development of the Gal4-UAS system, a binary

system based on the yeast transcription factor GAL4, has been fundamental in allowing tissue-specific genetic manipulation (Brand and Perrimon 1993; McGuire et al. 2004). GAL4-UAS is a binary system which allows the targeted of various constructs, namely protein coding sequences and RNA hairpin constructs for RNA interference (RNAi), in a tissue specific manner (**Fig. 1.7**). The system utilises the yeast transcription factor GAL4 whose expression is controlled by a tissue-specific promoter; and a transgene of interest placed downstream of GAL4-responsive, upstream activator sequences (UAS). In the absence of GAL4 protein, the UAS-transgene remains largely silent however, when GAL4 is expressed, it binds to UAS-sites and drives the expression of the transgene. Furthermore, the use of the bacteriophage phiC31-integrase has allowed researchers to precisely insert transgenes into exact genomic locations, termed 'landing sites' (Groth et al. 2004). This has allowed researchers to reliably compare different transgenes by controlling for 'position effects', the extent to which the location of the transgene within the genome affects its expression (Markstein et al. 2008). Multiple different fly lines currently exist with landing sites throughout the *Drosophila* genome, and are routinely used for the generation of UAS- driven lines (Venken and Bellen 2007; Markstein et al. 2008). Furthermore, insertion of mutant protein coding sequences into defined genomic locations allows the structure-function analysis of proteins *in vivo*, while RNAi analysis has also been aided by the collaborative nature of the *Drosophila* research community and development of large RNAi libraries (Dietzl et al. 2007; Ni et al. 2011; Green et al. 2014; Perkins et al. 2015). Thus, Gal4-UAS remains fundamental for analysing gene function in *Drosophila* (Del Valle Rodríguez et al. 2012).

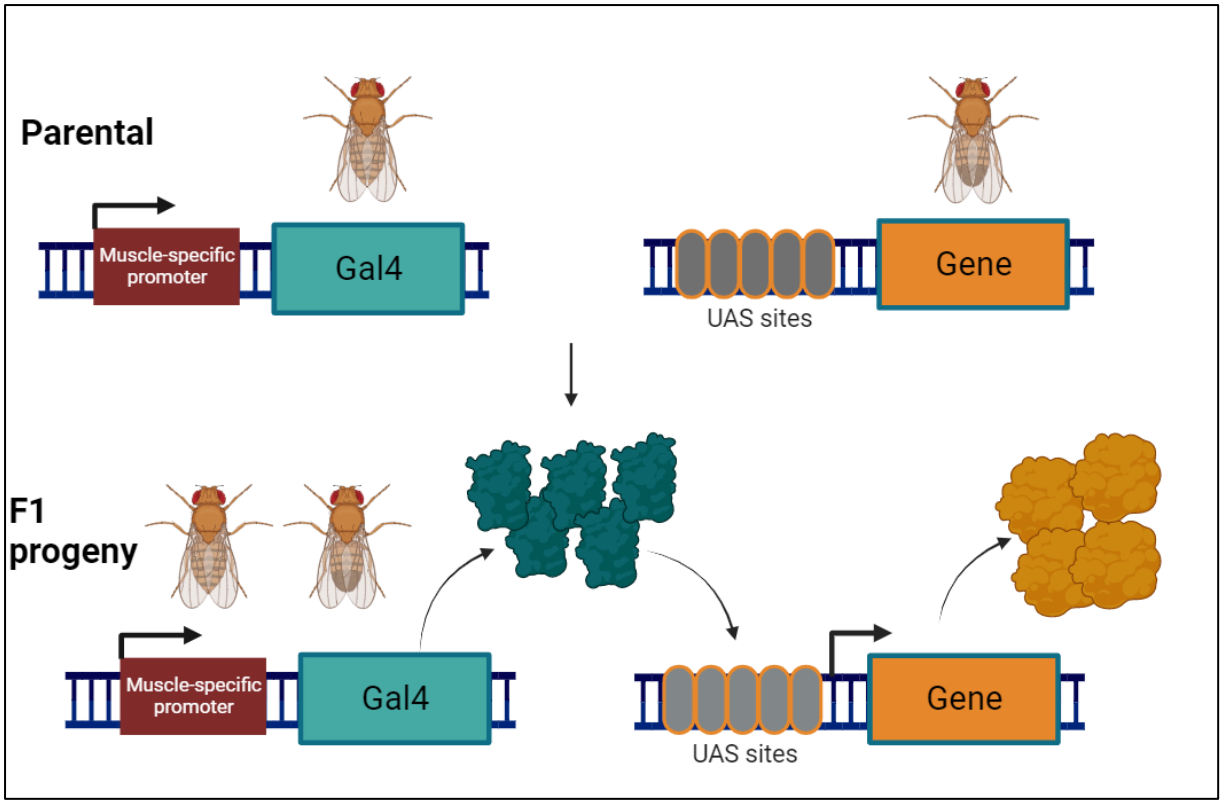


Figure 1. 7. Tissue specific manipulation of gene expression using the GAL4-UAS system.
Expression of the yeast transcription factor, Gal4, is under the control of a tissue-specific (for example, muscle) promoter. In another line, a transgene of choice is under the control of Gal4-responsive UAS-sites (either a cDNA for overexpression, or an RNAi construct for knockdown). When these lines are crossed together, Gal4 drives expression of the UAS-driven construct only in the tissues where Gal4 is expressed.

Nevertheless, while the tissue-specific overexpression and RNAi models have been incredibly powerful in the study of gene function, it has inherent limitations: Overexpression artificially increases expression far above endogenous levels, which may present biological phenotypes because of non-specific effects rather than elucidate the function of the native protein. RNAi on the other hand is typically limited by: the efficiency of the RNA knockdown and stability of the endogenous protein, as well as the potential for off-target effects. Therefore, one would ideally prefer to study gene function through modulation of the endogenous locus, thereby ensuring that any phenotype associated with a given change is a direct result of the change in function of the endogenous protein, with its expression under control of its endogenous regulatory sequences. The development of programmable nucleases, including Zinc-Finger nucleases (ZFNs), transcription activator-like proteins (TALENs) and, CRISPR-

Cas9, have become popular due to their ability to target specific DNA-sequences and introduce mutations effectively (Khalil 2020) .

1.6.2 CRISPR-Cas9: a revolution in genome engineering technology

CRISPR-Cas9 genome engineering is an application of the CRISPR-Cas adaptive immune response found in most archaea and bacteria, which utilises RNA-guided endonucleases to cut foreign DNA elements and protect against viral (phage) infection (Doudna and Charpentier 2014).The name itself originates from the components involved in the process: CRISPR-arrays containing Clustered Regularly Interspersed Short Palindromic Repeats (CRISPRs) separated by spacer sequences, which are unique and provide a template for a previously experienced foreign genetic element (i.e the memory); and Cas (CRISPR-associated) 'effector' proteins required for DNA cleavage (Makarova et al. 2011; Koonin and Makarova 2019). The exact mechanism by which CRISPR-Cas systems operate do vary, however key themes are shared. CRISPR-arrays are transcribed and processed into CRISPR-RNAs (crRNAs), which then form a complex with Cas protein(s). The crRNA provides the 'guide' to direct the Cas proteins to their intended target through DNA:RNA base pairing with the foreign genetic element. Such base-pairing induces cleavage of the foreign DNA through induction of a double-strand break (DSB) via the nuclease activity of the Cas protein, or protein complex. As with the diversity found among the bacterial and archaeal domains, there is huge diversity among the CRISPR-Cas systems found in different species (Makarova et al. 2011). For a more in-depth review of CRISPR-Cas system diversity and evolution, which lie outside the scope of this thesis, please see previously published comprehensive reviews (Makarova et al. 2011; Makarova et al. 2015; Koonin and Makarova 2019). The most notable differences that influenced the development of CRISPR-Cas for genome engineering is in the Cas effectors which induce DNA cleavage: Type I systems use a multi-Cas protein complex; whereas the Type II system from the bacterium *Streptococcus pyogenes* requires only one Cas protein, Cas9, induce (Jinek et al. 2012). Furthermore, only a single RNA molecule (termed single-guide RNA (sgRNA)) which has 20 nucleotides of homology required to provide target specificity, as well as a region required to interact with Cas9, is required for DNA cleavage (**Fig. 1.8a**) (Jinek et al. 2012; Ran et al. 2013). Cas9-induced cleavage is only limited by the requirement of an adjacent 5'NGG-3'

protospacer adjacent (PAM), a motif found on average every 16bp throughout any given genome, therefore meaning the CRISPR-Cas9 system can be utilised to efficiently induce DSB's throughout the genome. Therefore, the dual component nature and versatility in target site selection has made CRISPR-Cas9 an attractive as a genome engineering tool.

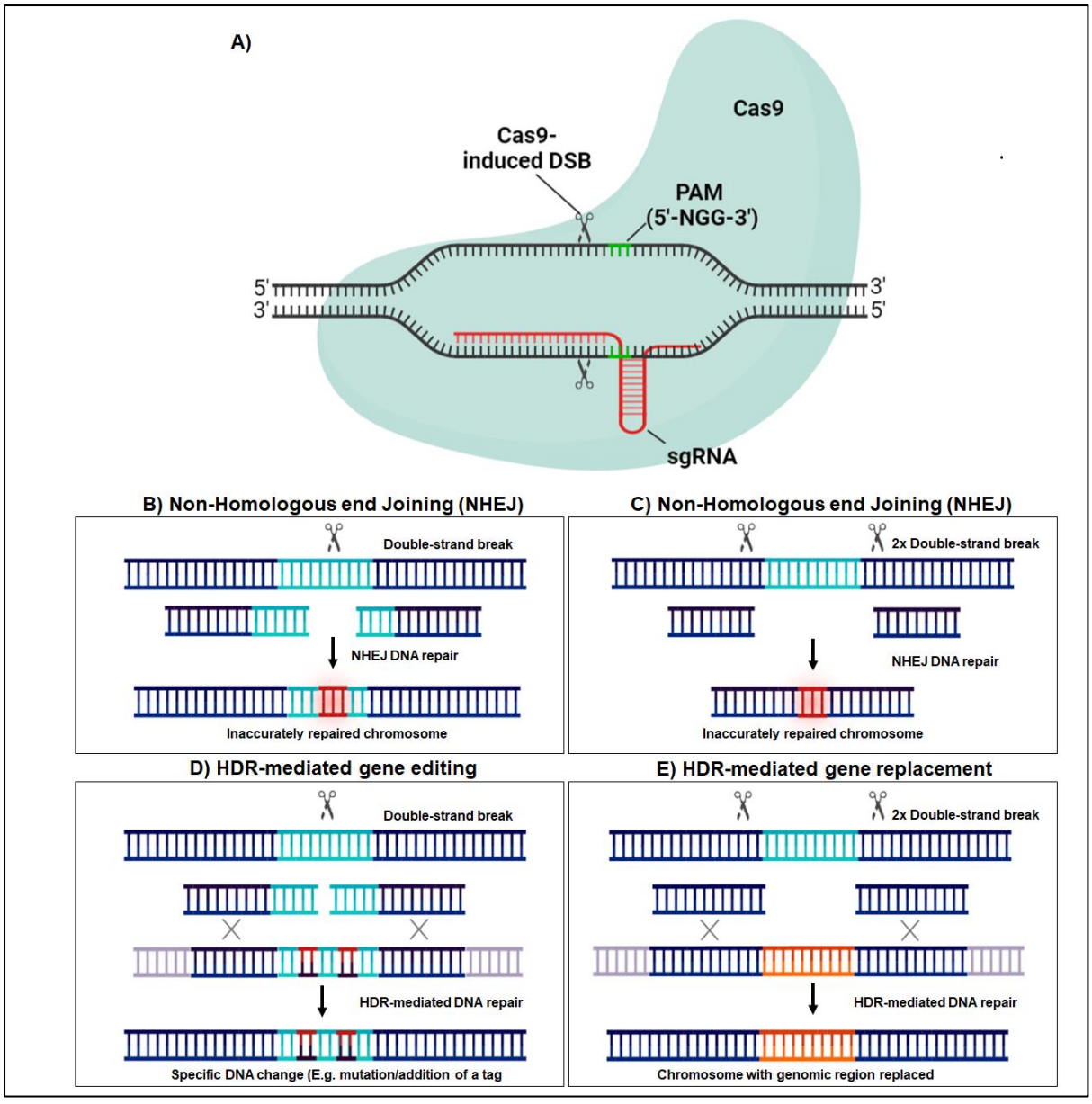


Figure 1. 8. CRISPR-Cas9 mediated genome engineering.
(**A**) The Cas9-endonuclease forms a complex with a sgRNA (single-guide), which guides the nuclease to its intended genomic target through DNA:RNA base pairing of 20 nucleotides between the genomic DNA, and the sequence-specific sgRNA. This induces cas9-cleavage of the genomic DNA between base-pairs 17 and 18 of the guide RNA sequence. (**B-D**) Different results of the DSB, based on cellular DNA-repair mechanisms. (**B-C**) Where there is

no homologous template, non-homologous end joining (NHEJ) inaccurately repairs the DSBs, causing mutagenesis. This can cause either small INDELs in the presence of only one cas9-induced DSB (**B**), entire genomic deletions where two cas9 DSBs are induced (**C**). (**D-E)** where there is a homologous DNA template available, the DSBs can be accurately repaired. By introducing desired changes into the HDR-repair construct, researchers can accurately introduce changes at single-bp resolution. This can either be utilised to generate specific changes at a desired location by using a single cas9 DSB (**D**), or generating entire gene-replacements by generating two cas9 DSBs (**E**).

Because of its sequence-specificity, CRISPR-Cas9 has been developed into a versatile tool for generating DSBs and making specific changes to DNA. Since the first seminal paper describing the CRISPR-Cas9 system (Jinek et al. 2012), it has been adapted for use in both fundamental and applied biological research, and in a wide variety of model organisms (Wang and Doudna 2023). By coupling Cas9-induced DSBs to the DNA-damage response (DDR), CRISPR-Cas9 can be used to introduce a wide variety of desired changes (**Fig. 1.8B-E**). In the absence of a homologous template, DSBs are repaired by error-prone non-homologous end joining (NHEJ), which typically introduces poorly-defined INDELS into the DNA sequence. By contrast, in the presence of a homologous repair template, DSB's can be accurately repaired by homologous recombination (HR). By introducing specific changes into the repair template, HDR-mediated repair can be used to introduce a variety of changes, such as nucleotide substitutions, defined deletions, insertion of additional genetic elements including protein tags, or even gene replacement (Wang and Doudna 2023).

CRISPR-Cas9 has been used to generate stable, heritable genomic modifications in *Drosophila* (Bassett et al. 2013; Gratz et al. 2013; Kondo and Ueda 2013; Ren et al. 2013; Gratz et al. 2014; Port et al. 2014). Since these initial studies, CRISPR-Cas9 has been widely adopted in *Drosophila* research (Bier et al. 2018; Zirin et al. 2022). However, prior to this project, there has been no application of CRISPR-Cas9 for the study of Mef2 function or regulation in *Drosophila*.

While making precise changes this way is valuable, its obvious limitation is the requirement to undergo new rounds of CRISPR for each desired change, which in turn involves relatively time-consuming vector design and synthesis protocols. Many genes, including Mef2, are regulated by multiple mechanisms involving different protein regions. Therefore, researchers may desire the ability to rapidly generate many different genetic changes for the study of gene function *in vivo*, while retaining the ability to do so by engineering the endogenous locus. By linking CRISPR-Cas9

genome engineering to site-specific recombination, the generation of 'Insertion-ready' deletion alleles can provide an excellent platform for the study of Mef2 function and regulation *in vivo*.

1.6.3 Insertion-ready deletion alleles and their potential for the study of *Drosophila* muscle biology

A more advanced, versatile application of CRISPR in the study of gene function is to generate an allele which would serve a dual purpose in the analysis of gene function: Firstly, in the generation of a null allele to analyse loss of function phenotypes; and secondly, the ability to rescue such a phenotype through the insertion of genetic elements, which could both restore the original gene's function and, in the case of insertion of a modified genetic element, generate mutated and/or tagged alleles for the analysis of gene function, while retaining the endogenous expression levels.

Insertion-ready deletion alleles in *Drosophila* have been identified as a technology that may enable such analysis of gene function in *Drosophila* (Poernbacher et al. 2019). Initially developed by Baena-Lopez et al. (2013), two Cas9-induced DSB's leads to the removal of part, or all of the genomic locus, in doing so generating a null allele by removing either all, or the majority of the gene's coding sequence (**Fig. 1.9**). By providing a repair template containing an *attP*-landing site and *Pax-Cherry* marker, flanked by homology arms, the gene can be replaced by *attP*-containing repair construct via HR, and successful insertion events can be easily screened for by observing *Cherry* expression. The *attP* site allows integration of 'rescue' genetic elements into the engineered locus and hopefully restore gene function. Even though single base variations in gRNA sequence are enough to prevent Cas9 cleavage and thus the likelihood of off-target mutagenesis appears to be low (Gratz et al. 2014; Ren et al. 2014) , a successful rescue of null phenotypes with wild-type gene fragments provides a gold-standard level of certainty that the phenotypes associated with novel LOF alleles are due to 'on-target' effects. Given a successful rescue with wild-type gene fragments, such alleles could then provide an effective platform to probe the function of the endogenous allele through the rescue with modified gene constructs such as those with tags, or mutations in specified domains with possible functional significance. However, because it is not a scarless technique, reintegration of wild-type genetic elements into the null locus may not necessarily fully restore gene

function. Therefore, care must be taken in designing such 'insertion-ready deletion' alleles to minimise disruption to the endogenous gene locus upon reintegration. However, this remains a non-exact science, given the lack of understanding of the importance of all DNA sequence within a gene sequence.

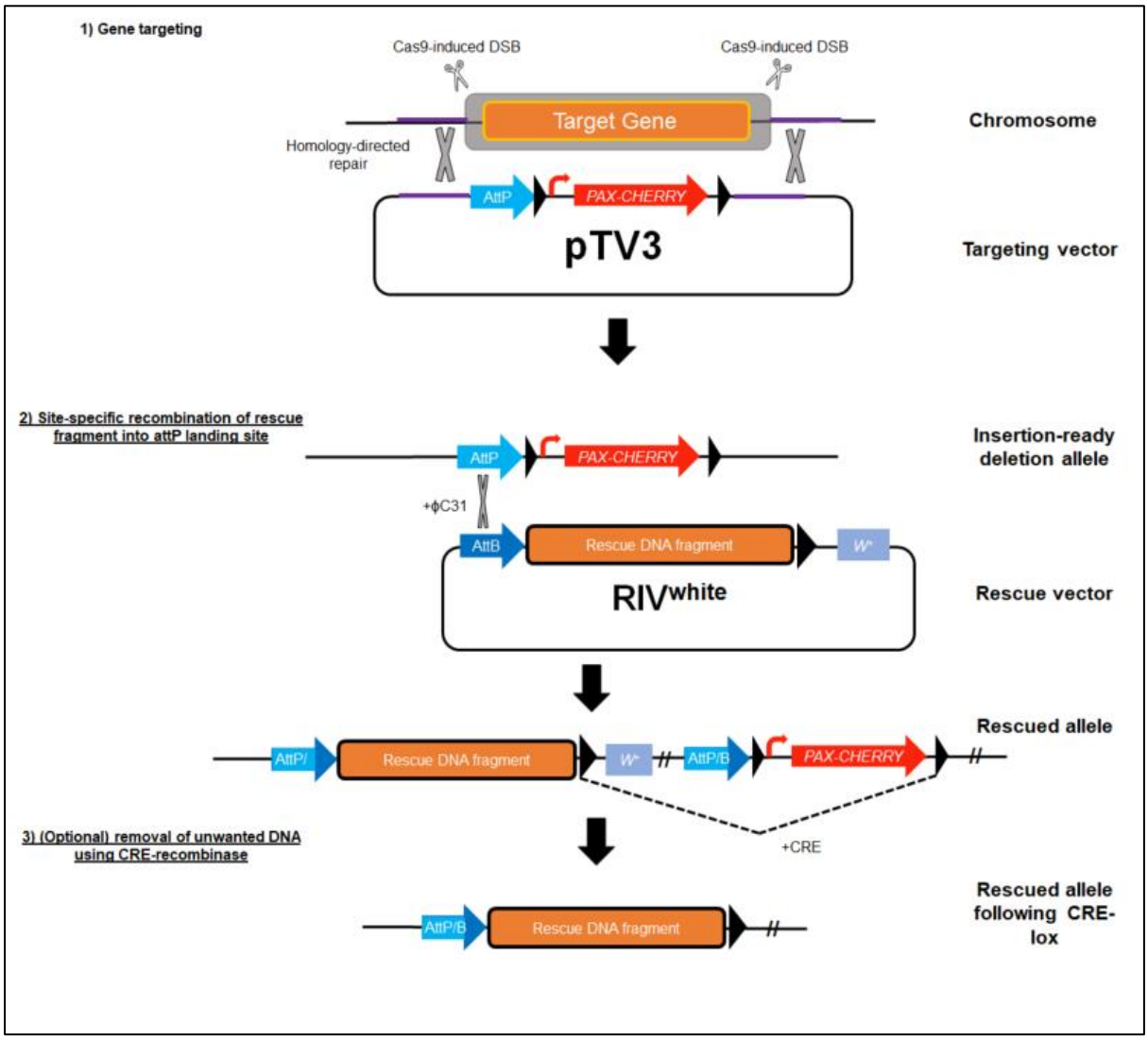


Figure 1. 9. Overview of the generation of insertion-ready deletion alleles.
(1) Two cas9 DSB promote deletion of the entire gene. A repair template vector (pTV3) contains homology arms flanking the cut sites, and promotes HDR of the locus, resulting in the insertion of a construct containing the attP landing site, followed by a *Pax-cherry* cassette flanked by loxP sites.
(2) Site-specific recombination (facilitated by the ΦC31 integrase) of a rescue vector (RIV^{White}) containing a complementary attB site, followed by a rescue DNA fragment designed to restore gene function, as well as a w^+ marker.

(3) (optional) Excess DNA integrated into the engineered locus (such as the marker genes) can be removed by using cre-recombinase, whose substrates are the loxP sites.

1.7 Project aims

In this project, I aimed to further the understanding of how Mef2 transcriptional activity *is* regulated *in vivo*, using *Drosophila melanogaster* as a model organism. In particular, I aimed to utilise CRISPR-Cas9 genome engineering to facilitate the investigation of potential Mef2-interacting proteins. The project can be refined into two key research aims:

- Utilize CRISPR-Cas9 genome engineering to develop cutting-edge tools for the study of Mef2, and Mef2-interacting proteins *in vivo*.

- Investigate the roles of the *Drosophila* Class IIa HDAC, HDAC4, and the p300/CBP HAT, nejire, in the context of Mef2 regulation in an *in vivo* model for muscle differentiation.

Chapter 2: Materials and Methods

2.1 Fly work

2.1.1 General Fly Work

Drosophila melanogaster stocks were maintained using standard methods (Stocker and Gallant 2008). Flies were reared in standard food vials consisting of agar, dextrose, yeast, cornflour, water, and Nipagen (10% p-hydroxy benzoic acid methyl ester). General fly stocks were maintained at 18°C and flipped into fresh food vials every 5 weeks. Two independent copies of each stock were maintained and flipped on alternate weeks. Working stocks were maintained at room temperature (approx. 21°C) and flipped into fresh food weekly. Where stock amplification was required for virgin collection or laying cages, flies were maintained at 25°C and flipped into fresh food every 2-3 days until a suitable number of vials were established.

Virgin collection was carried out by emptying tubes periodically 2-3 times daily. Tubes were emptied in the morning, and virgins were initially identified by their lighter cuticle colour and distinctive green spot on the abdomen (meconium). Subsequent collections throughout the day identified virgins by these visible characteristics, but also the knowledge that newly eclosed flies do not mate for approximately 8 hours post-eclosion (Stocker and Gallant 2008). Tubes were then moved to 18°C overnight to slow development to maximise subsequent collections the next day. Virgin females were stored in separate food vials until required. To set up crosses, approximately 10-20 virgin female flies were crossed with males at 2:1 – 3:1 ratio. For generation of new stocks, virgin females of the desired genotype were selected in the offspring and crossed with appropriate males. For laying cages, approximately 75-100 virgin female flies (if a cross) and 30 male flies were placed into a laying pot with a spot of yeast paste on an apple juice agar plate, and incubated at 25°C for 24 hours to acclimatize**.** Following this period, the apple juice plate was changed daily, or at intervals which varied depending on the stage, or age range of embryos to be collected.

2.1.2 Embryo Fixation

Embryos were collected from the apple juice plate and deposited in a cell strainer. Embryos were then dechorionated in 50% bleach for 2 minutes, before being washed with dH_2O to remove residual bleach, and subsequently blotted on tissue paper to dry. Embryos were then transferred into a petri dish in n-heptane (Fisher Scientific) and transferred into 1 ml fixing solution (200µl 37% formaldehyde (F8775, Sigma-Aldrich), 80µl 10X PBS (20-7400-10, Severn Biotech Ltd), 720µl dH_2O) and fixed on a rocking

platform for 20 minutes. Fixative was then removed, replaced with 1 ml methanol (Fisher Scientific) and the embryos were vortexed for approximately 45 seconds to devitellinise them. Residual heptane was then removed, and embryos were rinsed 3X in methanol before being stored in methanol at -20°C until required.

2.1.3 Embryo Immunostaining

Fixed embryos were transferred into wells of a 48-well plate, the methanol removed and replaced with 0.3% PBT solution (1X PBS containing 0.3% Triton X-100 (Sigma-Aldrich)). Embryos were rinsed 3X with PBT, washed for 5 minutes on a rocking platform, and subsequently rinsed once more. Embryos were then blocked for 30 minutes on a rocking platform in 300µl 0.3% PBT containing 1% w/v bovine serum albumin (BSA) (Sigma-Aldrich). Blocking solution was then removed and embryos were rocked at 4°C overnight in 250µl of diluted primary antibody (**Appendix.2)** in 0.3%PBT-BSA. The following day, embryos were rinsed 3X, and washed 3X in 0.3%PBT for 10 minutes each before embryos were rocked at room temperature for 1 hour in 200µl of diluted secondary antibody (**Appendix 2)** in 0.3%PBT-BSA, while protected from light. Secondary antibody was removed, and embryos were rinsed 3X and washed 3X in 0.3%PBT for 10 minutes each. Embryos were then transferred to a clean Eppendorf tube, PBT removed, and stored in 80% glycerol at 4°C. To mount, two coverslips were affixed to a glass microscope slide with nail varnish, with approximately 1cm space between them. Approximately 30-50µl of stained embryos (in 80% glycerol) were pipetted into the space, and covered with an additional coverslip. The slide was sealed with nail varnish and protected from light until imaged.

2.1.4 Fixation and Immunostaining of 3rd instar *Drosophila* wing discs

3rd instar wandering larvae were removed from the vial using a paintbrush and dissected in 1X PBS. To dissect, the posterior half of the larvae was removed and discarded by pulling the tissue away with forceps. The remaining tissue was then inverted by forcing the cuticle over itself, to expose the internal tissue. Dissected larvae were fixed in 4% paraformaldehyde (PFA) solution (4% w/v prilled PFA in 1X PBS) for 20 minutes at room temperature with gentle shaking. Sample preparation then varied depending on the assay. For the premature differentiation assay where antibody staining was not required, samples were protected from light to minimise bleaching of GFP signal. Fix was removed, followed by 3X rinses in 0.03% PBT solution (1X PBS

containing 0.03% Triton X-100). Tissue was then incubated with gently shaking in 0.03%PBT containing 1:3000 Hoechst for 30 minutes at room temperature (light protected), before dissection and mounting of discs.

For immunostaining, fixed samples were rinsed 3X and washed 3X for 15 minutes each in 0.03%PBT. Samples were then blocked in 250µl PBTN (2% v/v Normal Donkey Serum (NDS) in 0.03%PBT) for 30 minutes with gentle shaking. Samples were incubated in diluted primary antibody solution in PBTN overnight at 4°C with gentle shaking. The next day, samples were rinsed 3X and washed 3X (15 minutes each) in cold 0.03%PBT. Samples were then incubated with 250µl diluted secondary antibody solution (in PBTN) containing 1:3000 Hoechst, at room temperature for 1hour with gentle shaking. Samples were protected from light exposure to preserve fluorescence signal. Finally, samples were rinsed 3X and washed 3X (15 minutes each) in 0.03%PBT, before being dissected and mounted prior to imaging.

Wing discs were dissected in PBS from larval carcasses using forceps, on a plastic dish under a standard dissecting microscope. Discs were subsequently mounted on a glass microscope slide in 80% glycerol, and a coverslip placed over the top. To prevent crushing of wing discs, 4 larval brains were also mounted on the slide, one in each corner, to act as spacers. For the “premature differentiation” assay, wing discs were imaged immediately after mounting. Antibody-stained wing discs were stored at 4°C and protected from light until imaged. For the imaging of Mef2-EGFP in unfixed tissue, wing discs were immediately dissected from inverted larval carcasses and mounted in Schneider’s medium (Sigma, S0146). As with the “premature differentiation” assay, samples were imaged immediately following mounting.

2.1.5 Larval muscle assay

Wandering L3 larvae were removed from food vials using a paintbrush, and males were selected under a standard dissection microscope through identification of visible gonads (Chyb and Gompel 2013). Male larvae were then immobilised by placing on a glass microscope slide and incubating at -20 °C for 5-10 minutes. Larvae were then mounted (oriented laterally) in 80% glycerol, and flattened by gently placing a coverslip on top. Muscle structure was then visualised by microscopy.

2.1.6 Adult DLM muscle dissection and staining

Black pupae (approx. 90-96hr APF) were removed from vials using a paintbrush and dissected in PBS by cutting the most anterior end of the pupa case with fine dissecting scissors and removing enough such that the developing thorax was exposed. If required, flies of the desired genotype were then selected through the identification of balancer phenotypes. Pupae were then fixed in 17% formaldehyde solution (in PBS) at 4°C for 24 hours. Fixing solution was subsequently removed and the pupae rinsed 3X in PBS. Pupae were then emptied out into fresh drops of PBS on a plastic dissecting dish, and the rest of the pupa case was removed. Legs and wings were removed, before the thorax was cut between the first and second pairs of legs, and the head and abdomen were finally removed. To stain the DLM fibres, dissected thoracic segments were submerged in Mayer's haematoxylin (Sigma, MHS1-100ML) for 2 minutes, before being washed 2X for 2 minutes each in PBS. Stained thoraces were then stored in 80% glycerol solution until required for imaging. Stained thoraces mounted on a glass microscope slide between two strips of insulation tape to prevent crushing, and a glass coverslip was placed over the top and sealed with nail varnish.

For phalloidin staining, pupae were collected as described above, however were instead fixed for 20 minutes in 4% PFA solution for 30 minutes. The thorax was then prepared, and cut using the same method as for hematoxylin staining. Following this, dissected thoraces were rinsed 3X in 0.1% PBT, before being incubated in diluted phalloidin solution in 0.3% PBT (**Appendix. 2**) for 2 hours. Stained thoraces were mounted as for hematoxylin staining.

2.1.7 Hatching and survival assay

Approximately 80-100 females and 30 males were collected and placed in a laying cage. Young flies, less than 5 days old were chosen to maximise the collection of fertilised eggs. Flies were allowed to lay eggs on apple juice agar plates for 2 hours at 25°C, before the plate was changed and eggs were allowed to develop to approximately stage 15. Following this, fertilised embryos were selected (based upon identification of auto-fluorescent gut morphology) and, where necessary, the presence or absence of the fluorescent balancer chromosome. Using a paintbrush, embryos were carefully transferred to a yeast-paste supplemented apple juice plate, and incubated at 25 degrees. The following day, the percentage of hatched embryos were scored. Larvae were then incubated on the same apple juice plate until the surviving

larvae became L3, which was determined based upon mouthpart morphology, and darkening of the posterior spiracles. L3 larvae were transferred to a standard food vial and allowed to develop further. Percentage survival was calculated by determining the number of live individuals that reached one of four particular stages (hatched, L3, pupation, adult fly), as a proportion of the total population of embryos first selected. 100 Embryos were selected per genotype.

2.1.8 Microinjection of constructs for transgenesis

All injections to generate transgenic lines were outsourced. For generation of HDAC4/Mef2 insertion-ready deletion ($^{\Delta att}$) CRISPR alleles, miniprep DNA was sent to BestGene Inc. (Chino Hills, CA), who then prepared the DNA for injection (Service Z1). The premium genome engineering service (service RI) was used, which provided balanced heterozygous mutant alleles. Constructs were injected into embryos of the genotype *yw;;nos-Cas9(III-attP2)* (Kondo and Ueda 2013). For rescue of the Mef2 locus, rescue constructs were injected into their respective Mef2$^{\Delta Att}$ lines by The University of Cambridge Department of Genetics Fly Facility, due to logistical challenges associated with transporting live *Drosophila* to the USA. Midi-prepped rescue constructs (RIV^{white}) (>400 ng/ul) were sent with approximately 5-6 vials of each Mef2$^{\Delta Att}$ and PhiC31 integrase-expressing line (BDSC#40161). Embryos injected were offspring of a cross between virgin females of the PhiC31 integrase, and Mef2$^{\Delta Att}$ lines. Because the Mef2$^{\Delta Att}$ lines were homozygous lethal and maintained over the Cyo balancer, survivors displaying the dominant Cyo phenotype were not screened for integration by Cambridge. As with the initial allele generation by BestGene, transgenics were screened for, identified, and balanced to generate stable lines by Cambridge. For rescue of the HDAC4 locus, BestGene was used, with service I for these injections, with PhiC31 DNA injected directly into embryos of our HDAC4 null stock (HDAC4^{Att}/FM7).

For generation of novel UAS-HDAC4 and UAS-Mef2 (WT and mutant) lines, midi-prepped pUAST-attB vector containing HDAC4 or Mef2 CDS (>500 ng/ul) was sent to BestGene and inserted into the either the attP40 on ChrII (HDAC4), or attP2 on ChrIII (Mef2) landing sites. The associated landing site-containing stocks used for these injections were: y^1 w^{67c23}; P(CaryP)attP40 or y^1 w^{67c23};; P(CaryP)attP2 (Groth et al. 2004; Markstein et al. 2008). Embryos injected were that of a cross between the aforementioned landing site stocks and the germline-expressing PhiC31 expressing

stock M(vas-int.Dm)ZH-2A (BDSC#40161). We selected plan H, where we received unbalanced heterozygous transformant stocks. These stocks were then balanced in-house prior to experimental use.

For crossing schemes for both H, and I plans, see **Appendix 3 and Appendix 4**

2.1.9 List of fly stocks used in this project

Stock	Genotype (if available)	Reference/Source
Gal4 drivers		
Mef2-Gal4, Mhc-GFP	*yw;; mef2-Gal4,Mhc-GFP*	Gift from Dr Frank Schnorrer.
TwipTwip-Gal4	*w, twist-GAL4; twist-Gal4*	Baylies & Bate (1996)
1151-Gal4	*w, 1151-Gal4*	(Roy & VijayRaghavan 1997; Anant *et al*. 1998)
1151-Gal4;UAS-mCherry	*w, 1151-Gal4; UAS-CD8mCherry*	Generated by Rob Mitchell using 1151-Gal4 and BDSC#27391
1151-Gal4;;Mhc-GFP	*w,1151-Gal4;;Mhc-GFP*	Generated by Rob Mitchell using 1151-Gal4 and ftrg500
Act88F-Gal4	*w, Actin88F-Gal4*	Gift from Frank Schnorrer. Bryantsev *et al* (2012)
Daughterless-Gal4	*w*; KrIf-1/CyO; P(GAL4-da.G32)UH1*	BDSC#55850
UAS-lines		
UAS-HDAC4 lines		
UAS-HDAC4WT	*w;UAS-HDAC4 (attP40)*	Generated during this project
UAS-HDAC4 WT.Myc	*w;UAS-HDAC4 WT.Myc (attP40)*	Generated during this project, construct was a gift from Helen Fitzsimons (Massey, NZ)
UAS-HDAC4 Y1142H	*w; UAS-HDAC4ΔY1142H (attP40)*	Generated during this project
UAS-HDAC4 F171A	*w; UAS-HDAC4ΔF171A (attP40)*	Generated during this project
UAS-HDAC4 K165,L168,I172A.myc	*w; UAS-HDAC4ΔK165,L168,I172A (attP40)*	Generated during this project, construct was a gift from Helen Fitzsimons (Massey, NZ
UAS-HDAC4 L168A.Myc	*w;UAS-HDAC4ΔL168A.Myc (attP40)*	Generated during this project
UAS-Mef2 lines		
UAS-Mef2 10t4a (III)	*w;UAS-Mef2 10t4a (attP2)*	Generated by Rob Mitchell using sequence

		information from Gunthorpe et al (1999)
UAS-Mef2 10t4a (II)	*w;;UAS-Mef2 10t4a (attP40)*	Generated by Rob Mitchell using sequence information from Gunthorpe et al (1999)
UAS-Mef2 CDS.HA	*w;;UAS-Mef2 CDS.HA (attP2)*	Generated during this project
UAS-Mef2 VLL65-67ASR.HA	*w;;UAS-Mef2ΔVLL65-67ASR.HA (attP2)*	Generated during this project
UAS-Mef2 L66A.HA	*w;;UAS-Mef2ΔVLL65-67ASR.HA (attP2)*	Generated during this project
UAS-Mef2 L67A.HA	*w;;UAS-Mef2ΔL67A.HA (attP2)*	Generated during this project
UAS-Mef2 Y69A.HA	*w;;UAS-Mef2ΔY69A.HA (attP2)*	Generated during this project
	Additional lines	
UAS-Nejire.V5	*w[*]; P(w[+mC]=UAS-nej.wt-V5)3*	BDSC#32573
UAS-CD8-mCherry (II)	*w[*]; P{w[+mC]=UASmCD8.ChRFP}2*	BDSC#27391
UAS-CD8-mCherry (III)	*w[*]; P{w[+mC]=UASmCD8.ChRFP}3*	BDSC#27392
Mhc-GFP		VDRC#318471 (FTRG500)
RNA interference (RNAi) lines		
UAS-HDAC4 RNAi (20522)	*w1118; P(GD9446)v20522*	VDRC #GD20522 (Dietzl et al. 2007)
UAS-HDAC4 RNAi (330055)	*w1118; P(VSH330055)attP40*	VDRC #330055
UAS-HDAC4 RNAi (34774)	*y[1] sc[*] v[1] sev[21]; P(y[+t7.7] v[+t1.8]=TRiP.HMS00083)attP2/TM 3, Sb[1]*	BDSC#34774
UAS-HDAC4 RNAi (28549)	*y[1] v[1]; P(y[+t7.7] v[+t1.8]=TRiP.HM05035)attP2*	BDSC#28549
UAS-Mef2 RNAi	*w1118; P(GD5039)v15550*	VDRC #15550
UAS-Nej RNAi (KK)		VDRC#KK102885
UAS-Nej RNAi (TRiP)	*y[1] sc[*] v[1] sev[21]; P(y[+t7.7] v[+t1.8]=TRiP.HMS01507)attP2*	BDSC#37489
Mef2 CRISPR alleles		
$Mef2^{\Delta CDS}$	*w;Mef2$^{\Delta CDS}$/Cyo* **or** *w;Mef2$^{\Delta CDS}$/CyoActGFP*	Generated during this project
$Mef2^{\Delta ATG}$	*w;Mef2$^{\Delta ATG}$/Cyo* **or** *w;Mef2$^{\Delta ATG}$/CyoActGFP*	Generated during this project
$Mef2^{\Delta CDS\text{-}gDNA}$	*w;Mef2$^{\Delta CDS\text{-}gDNA}$/sm6a*	Generated during this project
$Mef2^{\Delta CDS\text{-}cDNA}$	*w;Mef2$^{\Delta CDS\text{-}cDNA}$/sm6a*	Generated during this project
$Mef2^{\Delta ATG\text{-}gDNA}$	*w;Mef2$^{\Delta ATG\text{-}gDNA}$/sm6a*	Generated during this project
$Mef2^{\Delta ATG\text{-}cDNA}$	*w;Mef2$^{\Delta ATG\text{-}cDNA}$/sm6a*	Generated during this project

Mef2-EGFP	*w; Mef2-EGFP*	Generated during this project
HDAC4 CRISPR alleles		
HDAC4$^{\Delta att}$	*y,w,HDAC4$^{\Delta Att}$/FM7* **or** *y,w,HDAC4$^{\Delta Att}$/FM7actGFP*	Generated during this project
HDAC4$^{\Delta att\text{-}gDNA}$	*y,w,HDAC4$^{\Delta Att\text{-}gDNA}$*	Generated during this project
HDAC4-mScarlet-I	*y,w,HDAC4-mScarlet-I*	Generated during this project
mScarlet-I-HDAC4	*y,w,mScarlet-I-HDAC4*	Generated during this project
Balancer chromosome lines		
If/Cyo;TM3/TM6	*w;If/Cyo;TM3/TM6*	
FM7actGFP	*FM7i, P(w[+mC]=ActGFP)JMR3/C(1)DX, y[1] f[1]*	BDSC#4559
CyoActGFP	*Tango11[1]/CyO, P{w[+mC]=ActGFP}JMR1*	BDSC#36320
	yw;Cyo/Sco	BestGene (Via BDSC#2555)
FM7i	*FM7i/C(1)DX, y1 f1*	*BestGene (via BDSC#5263*)
w/FM6;;Sb/TM3Ser	*w/FM6;;Sb/TM3Ser*	Gift from Dr Sonia Lopez de-Quinto

2.2 Molecular Biology

2.2.1 Genomic DNA (gDNA) extraction

Genomic DNA (gDNA) was extracted from adult flies using the Invitrogen ChargeSwitch™ (CS11203) gDNA Miro Tissue kit. The manufacturers protocol was adapted for use in *Drosophila*. 3-5 flies (cooled on ice) were crushed in a 1.5 ml microcentrifuge tube containing lysis buffer (L15+5 ul Proteinase K (20 mg/ml)) using a sterile pestle. Fly-lysis mix was then incubated at 55°C for 2 hours, with vortexing at regular intervals. Following incubation, RNA was degraded by incubating lysis mix for 5 minutes with 2.5 ul RNAse A (5 mg/ml). To bind DNA, 100µl Purification Buffer (N5) and 20µl ChargeSwitch™ beads were added, incubated for 1 minute, before placing the microcentrifuge tube in a magnetic rack (MagnaRack™, CS15000) for 1 minute. Supernatant was then removed and discarded. Beads were then washed twice by resuspending in 500µl wash buffer (W12), incubating in the magnetic rack for 1 minute, and then discarding supernatant. DNA was finally eluted by resuspending the beads in 75 ul Elution Buffer (Qiagen buffer EB, 10mM Tris-Cl), incubating for 5 minutes, and then incubating in the magnetic rack for 1 minute. Eluted DNA was then removed and stored at -20°C until required.

2.2.2 Polymerase Chain Reaction (PCR)

For DNA amplification, all PCR primers were designed using standard guidelines, including: approximately 18-30 bp in length (varied when overhangs were required), 40-60% GC content, T_m difference no greater than 5 degrees, and ideally a GC clamp at the 3' end of the primer. Primer characteristics were confirmed using the OligoEvaluator™ tool (available at: https://www.sigmaaldrich.com/GB/en/technical-documents/technical-article/genomics/pcr/oligo-evaluator-for-tm-calculation-primer-analysis), with primer-dimers and strong secondary structure also avoided where possible. All primers were ordered from Sigma. Standard PCR reactions were carried out using either gDNA or plasmids as templates using either BioRad T100 or Biometra TRIO thermal cyclers. Depending on the purpose, a variety of polymerases were used. As a preference, I used either NEB Q5 2X Master Mix (M0492) or Q5 HotStart 2X Master Mix (M0494) high-fidelity polymerases to amplify regions where sequence fidelity was important (homology arms, gRNA sequences, rescue constructs, cDNA sequences). For genotyping experiments, we used NEB OneTaq 2X Master Mix (M0482). DNA was amplified in 25 ul reactions, using reaction mixture composition, and cycling conditions suggested by the manufacturer. Annealing temperature was determined using NEB T_m calculator (https://tmcalculator.neb.com/#!/main). For more difficult amplicons which could not be amplified with NEB HiFi polymerases, we used PCRBIO VeriFi™ Mix (PB10.43-01). This was particularly important for the amplification of sequences from the *HDAC4* locus, as this genomic region could not be amplified using Q5 polymerases. DNA was amplified in 25 ul reactions, and following the manufacturers guidelines. As no T_a calculator was available, annealing temperature was optimized by gradient PCR, with T_a ranging from 60-72°C.

2.2.3 Agarose gel electrophoresis

PCR-amplified DNA or restriction digested vector fragments were size-separated and analysed on agarose gels. Gels were constructed using agarose concentrations between 0.8-1.2% (w/v) in 1X TBE buffer (89 mM Tris, 89 mM boric acid, 2 mM EDTA) containing SYBR® Safe DNA gel stain, diluted 10'000X (Life Technologies). Gel concentration was adjusted depending on the size of fragment. DNA mix was prepared by adding either 5x (GelPilot, Qiagen), or 6x (B7024, NEB/ TriTrack, ThermoFisher) loading dye. DNA was loaded into cast gels submerged in 1x TBE and ran at between 75-100V for the appropriate time. To determine size of DNA fragments, either 100 bp

(NEB, N3231), 1 kb (NEB, N3232) or GeneRuler 1 kb Plus (ThermoFisher, SM1331) ladder solutions, diluted appropriately in dH_2O and loading dye, was run alongside samples. DNA was visualized either using the GelDoc EZ™ Imager (BioRad), or on a Dark reader transilluminator (Clare Chemical Research) if the DNA was to be purified and used in downstream applications.

2.2.4 DNA Gel Extraction and Purification

DNA bands were excised from agarose gels using a clean scalpel and placed in a 1.5 ml microcentrifuge tube. DNA was then purified using the QIAGEN Qiaquick® gel extraction kit and following the manufacturers instructions. DNA was eluted in 30-50µl Buffer EB, depending on DNA yield and desired concentration of the eluate. In instances where DNA fragments did not require separation via agarose gel electrophoresis, DNA mixture was directly purified using the QIAquick® PCR purification protocol. Purified DNA was then stored at -20°C until required.

2.2.5 Restriction endonuclease digestion

All restriction enzymes were supplied by NEB, and general protocols for either single, or double digestions were used via the NEBcloner online tool (https://nebcloner.neb.com/#!/). For preparing digested vector samples for cloning, approximately 2ug vector was digested in 50 µl reactions and using 10 units of restriction enzyme per ug of vector. Digests were then incubated at 37°C for 1-2 hours. To prevent empty vector re-ligation downstream, digested vector ends were dephosphorylated by the addition of 1.5 ul Quick CIP (NEB, M0525) directly to the restriction digestion mix and incubated for a further 30 minutes at 37°C. For PCR products, approximately 500 ng-1 ug was digested in the same reaction volume and conditions as vectors. Digested PCR products were not dephosphorylated to enable ligation. Digested DNA was then run on an agarose gel and purified.

2.2.6 Ligation

Ligations were carried out in 20 µl reactions using T4 DNA ligase and following the manufacturers guidelines (NEB, M0202). Approximately 20 ng vector was mixed with insert at a 1:3-1:5 vector:insert molar ratio and incubated either for 1 hour at room temperature, or overnight at 16°C. Following incubation, ligation mixtures were stored at 4°C prior to transformation, without a heat-inactivation step. If long term storage was required, ligations were kept at -20°C.

2.2.7 Gibson assembly

Assembly of DNA fragments was carried out using the Gibson Assembly ® Master Mix (NEB, #E2611), and following the manufacturer's guidelines. Approximately 0.02-0.5 pmol total DNA fragments (ensuring a 3-fold molar excess of each insert relative to linearised vector) were mixed with Gibson Assembly 10X Master Mix and dH_2O up to a total volume of 20 µl. Sample mix was then incubated at 50°C for 15 minutes. 4ul reaction mix was then used to transform 50 µl DH5α competent cells, and constructed plasmids were propagated in bacterial culture and isolated, before being verified by DNA sequencing.

2.2.8 Competent cells, LB agar and bacterial growth media preparation

LB Agar (Miller's) (ThermoFisher, BP-9724) and LB Broth (ThermoFisher) granules were dissolved in the appropriate volume of dH_2O, sterilized by autoclaving, and left to cool to approximately 55°C before adding the appropriate antibiotic. LB Agar was poured into 100 mm petri dishes, and both LB Agar plates and LB Broth were stored at 4°C until required. Unused media was discarded after one month.

The Z-competent™ DH5α *E.coli* Transformation Kit (Zymo Research, T3001) was used for transformations. To prepare these aliquots of these cells, cells were thawed from the glycerol stock (maintained at -80°C) and streaked onto an LB Agar plate. A single colony was then picked and grown in LB Broth at 37°C for 4-6 hours, before 5 ul of that starter culture was inoculated in 50 ml LB-20mM glucose and grown at 18°C in a shaking incubator. Cells were incubated for approximately 3 days until the appropriate density was reached (OD^{600}=0.4-0.6). Cells were then cooled on ice before pelleting at 2°C for 10 minutes at 2000g. The supernatant was discarded and cells were resuspended gently in 5 ml ice-cold 1X Wash Buffer. Cells were then pelleted at 2°C, 2000g for 10 minutes, and the remaining supernatant discarded. Finally, cells were resuspended in in 5 ml ice-cold 1x Competency Buffer, aliquoted (50µl) into 1.5 ml microcentrifuge tubes, flash frozen in liquid nitrogen, and then stored at -80°C until required.

2.2.9 Bacterial Transformation

When required, competent cell aliquots were thawed on ice, and added directly to either 2-5 ul of ligation/Gibson assembly mixture, or 1-2 ng purified miniprep plasmid DNA. The transformation reaction was mixed by gently pipetting up and down, before

being incubated on ice for 5 minutes. Cells were then spread on to LB-Agar plates containing appropriate antibiotic (Ampicillin, 100 ug/ml or Chloramphenicol, 34 ug/ml) and incubated at 37°C overnight.

2.2.10 Plasmid preparation from Bacterial Culture

Plasmid DNA was isolated from cultures derived from single colonies to either screen for desired transformants, or to prepare plasmid DNA samples for *Drosophila* microinjection. For screening transformants. approximately 4-8 single colonies from streaked bacterial transformation reactions were picked using a sterile 10 µl pipette tip and cultured in 2-5 ml LB Broth (containing antibiotic) overnight (approximately 12-16 hours) in a shaking incubator at 37°C. Plasmid DNA was then isolated using the Qiagen QiaPrep Spin Miniprep kit DNA was purified according to the manufacturer's instructions, and eluted in 50 µl Buffer EB. For preparation of plasmid DNA for microinjection, miniprep DNA was retransformed, and single colonies were picked and cultured overnight in 50 ml LB Broth (containing antibiotic) at 37°C. Plasmid DNA was then purified using the Qiagen HiSpeed Plasmid Midi Kit and eluted in 1 ml Buffer TE.

Prior to sending for microinjection, eluted DNA yield was analysed using a ThermoFisher NanoDrop 1000. If the concentration was <500 ng/ul (BestGene requirement), ethanol precipitation was used to precipitate, and then the DNA resuspended in a lower volume: 1 ml plasmid DNA elution was mixed with 1ml 5M Ammonium Acetate and 5 ml 100% Ethanol, before centrifugation for 10 minutes at 10'000 rpm at 4°C. The supernatant was carefully removed, and the pellet washed with 1 ml 75% Ethanol. Ethanol was then removed, and the pellet left to air dry for approximately 10-15 minutes, before resuspending in an appropriate volume of Buffer TE to achieve the required concentration.

2.2.11 Site directed mutagenesis

Mutagenesis was carried out using the QuikChange II XL Site-Directed Mutagenesis Kit (Agilent, 200521), with primers designed using the QuikChange Primer Design tool (https://www.agilent.com/store/primerDesignProgram.jsp). The mutagenesis reaction was set up according to the manufacturer's instructions. Parental plasmid was digested by adding 1 ul DpnI to the completed reaction and incubating for 1 hour at 37°C, before mutant plasmid was transformed into Z-competent™ DH5α *E.coli.* Prior to subsequent cloning steps, mutagenesis was confirmed by DNA sequencing

covering the entire CDS to confirm that there had been no unwanted off-target mutagenesis elsewhere in the sequence.

2.2.12 Verification of constructs

Purified plasmid DNA samples were verified by restriction digestion. Enzymes used to clone DNA fragments into constructs were used to release it from the vector backbone, and reactions were then visualised on an agarose gel to check for presence of insert. Fidelity of cloned DNA sequence was then checked by DNA sequencing: All sequencing during this project was completed by Eurofins Genomics, using their TubeSeq service. Constructs which could not be verified by restriction digestion were verified by DNA sequencing alone. Where genomic DNA was used as a template for CRISPR-Cas9 genome engineering, sequence fidelity was verified through the sequencing of two independent PCR products, given that gDNA sequence may vary from the annotated *Drosophila* reference genome sequence.

2.3 Constructs generated in this project

See **Appendix. 1** for all primers used in this project. Primer names detailed below are identical to those found in the Appendix. Where restriction enzymes were used to clone in fragments, overhangs of between 3-6 nucleotides were included at the 5' end of the primer to optimise cleavage efficiency. Overhang length was determined from NEB's guidance on enzyme cleavage efficiency when near the end of a DNA fragment (available at: https://www.neb.com/en-gb/tools-and-resources/usage-guidelines/cleavage-close-to-the-end-of-dna-fragments). Restriction sites were followed immediately by the desired fragment to be amplified and inserted.

2.3.1 UAS-HDAC4 constructs

To probe the effect of HDAC4 overexpression on *Drosophila* myogenesis, HDAC4 coding sequence (CDS) was cloned into the pUAST-attB vector (gift from Dr Sonia Lopez de Quinto, **Appendix. 5,** (Bischof et al. 2007)) for subsequent injection and integration into the *Drosophila* genome at the attP40 (25C7) landing site (Markstein et al. 2008). We obtained HDAC4 CDS from the Berkeley Drosophila Genome Project (BDGP) Gold Collection (clone FI19806, isoform RD) (Stapleton et al. 2002) via the Drosophila Genomics Resource Center (DGRC). To clone HDAC4 CDS into pUAST-attB, CDS was amplified using NEB Q5 2X Master Mix with primers UAS-HDAC4_CDS_F-BglII and UAS-HDAC4_CDS_R-XbaI. PCR product was gel purified, before being directionally cloned into the multiple cloning site (MCS) of pUASt-attB

using BglII and XbaI. I also gratefully received a pUAST-attB vector containing the same HDAC4 cDNA clone (FI19806) conjugated to a 6xMyc tag from Dr Helen Fitzsimons (Massey University, NZ).

2.3.2 Site directed mutagenesis of the Mef2-binding and catalytic domains of HDAC4

I used site-directed mutagenesis to mutate specific residues within the Mef2 binding domain of HDAC4. To avoid sequencing the entire vector construct due to potential off-target effects during mutagenesis, I either mutated the HDAC4 CDS within the original FI19806 clone, before cloning the mutant CDS into pUASt-attB**,** or mutated the HDAC4 CDS within pUAST-attB before subcloning into a fresh copy of pUAST-attB. I used four UAS-HDAC4 mutant constructs in this project. Primer pairs HDAC4_SDM_F171A-F/R and HDAC4_SDM_LEGGY/H-F/R were used to generate the F171A and Y1142H mutations respectively. Mutagenesis was carried out using the original FI19806 cDNA clone as template. Following confirmation of the desired change, the mutant CDS was subsequently cloned into pUAST-attB as described. The second, UAS-HDAC4 L168A.Myc line was generated with primers HDAC4_SDM_L168A-F/R and using the UAS-HDAC4.Myc-containing pUAST-attB plasmid donated by Dr Helen Fitzsimons as template. The pUAST-attB plasmid containing the CDS for UAS-HDAC4 K165,I168,I172A DNA was kindly sent to us by Dr Helen Fitzsimons, which was generated in her lab but reinjected by BestGene to make our own independent line.

2.3.3 UAS-Mef2 constructs

Mef2 CDS fused to a C-terminal 3xHA tag was generated in a 2-step cloning procedure. The 5' 1090bp of Mef2 CDS (isoform III) fragment was PCR amplified with NEB Q5 2X Master Mix using primers UAS-Mef2_CDS_F+EcoRI and UAS-Mef2_CDS_R+BglII. CDS was amplified from an existing vector within the lab containing the desired sequence (Gunthorpe et al. 1999). PCR product was purified and cloned into the pUAST-attB vector using EcoRI and BglII. Vector and PCR product was digested sequentially due to buffer incompatibility between the two enzymes. The C-terminal fragment of the Mef2 CDS (including the internal BglII site of the Mef2 CDS) plus sequence encoding a 3xHA epitope tag (separated by a Gly-Gly-Ser linker peptide) was synthesised as a gene fragment from Twist Bioscience (San Francisco, CA) (**Appendix 13**). Sequence for the 3xHA epitope tag was obtained from the

existing pUASG.HA vector (Bischof et al. 2012). This synthesised fragment was then directionally cloned into pUAST-attB vector (containing the 5' fragment) using the BglII and XbaI restriction sites, to generate the full length UAS-Mef2.3xHA construct. The UAS-Mef2 10t4a construct was generated separately in the lab by Rob Mitchell.

2.3.4 Site-directed mutagenesis of the Mef2 domain

Site directed mutagenesis was used to generate four separate mutations, L66A, L67A, Y69A, and VLL65-67ASR using primers: Mef2_SDM_L66A-F/R; Mef2_SDM_L67A-F/R; Mef2_SDM_Y69A-F/R; and Mef2_SDM_VLL65-67ASR-F/R. These mutations are all contained within the 5' 1090bp of the Mef2 CDS. Site-directed mutagenesis of the Mef2 cDNA was carried out using these primers and the original template vector containing Mef2 cDNA. Following site-directed mutagenesis and verification of the desired change, the UAS-Mef2 mutant line was generated as described in **2.3.3**: The N-terminal 1090bp of CDS (containing the mutation) was cloned into pUAST-attB, followed by the C-terminal, synthesised fragment.

2.3.5 CRISPR target site selection

Suitable CRISPR sites (or protospacers) were first determined based on their location. For the direct tagging of *HDAC4* or *Mef2*, optimal CRISPR sites were determined to be as close as possible to the desired insertion site (i.e the start/stop codon). For the generation of insertion-ready deletion alleles, optimal target site selection was less obvious and required careful consideration, taking into account two major considerations of the system. Firstly, to remove enough of the locus to generate a null allele, but not too much to compromise the probability of generating the desired allele. Secondly, to minimise permanent disruption to the locus following rescue (i.e such that the rescued alleles have fully restored gene function). There is a lack of ability to predict this effectively, due to gene level variations and the contributions of, in particular non-coding sequence to gene function *in vivo*.

Suitable protospacer sequences for genome engineering were selected using the CRISPR optimal target finder online tool (http://targetfinder.flycrispr.neuro.brown.edu/) (Gratz et al. 2014). Sequence for the desired region of the genome to be engineered was obtained via Flybase (Gramates et al. 2022), and CRISPR targets were identified using the maximum stringency setting. CRISPR sites with predicted potential off-targets on the same chromosome were not selected. CRISPR sites were verified in the Cas9-expressing line by PCR followed by DNA sequencing. This was to confirm

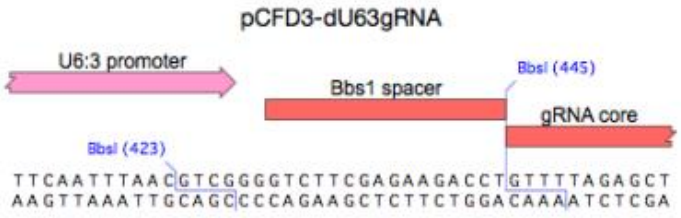
there were no unique SNPs in the Cas9-expressing line not present in the reference genome, since even small mismatches in the gRNA sequence have a significant effect on the efficiency of Cas9-induced DSB generation (Ren et al. 2013; Gratz et al. 2014).

2.3.6 Vectors used for CRISPR-Cas9 genome engineering

2.3.6.1 gRNA expressing vectors

gRNA-expressing vectors used during this project are the publicly available, pCFD vectors generated by Port et al (2014). For the generation of insertion-ready deletion (or IRD) alleles, pCFD4 was used (**Appendix. 7**). For the generation of direct-tagged alleles, pCFD3 was used (**Appendix. 6**). The cloning approach for these vectors is slightly different from conventional restriction cloning, and therefore schematics are indicated below.

For pCFD3 cloning (**Fig. 2.1**), two complimentary oligos (Sense = CRISPR site/protospacer sequence, antisense = reverse compliment) were annealed and phosphorylated in a single 10µl reaction containing: 1 ul of each oligo stock solution (100uM), 1 ul T4 10X ligation buffer, 0.5 ul T4 PNK (NEB, M0201), and 6.5 ul dH_2O. Oligos were incubated at 37C for 30mins, followed by 95C for 5 minutes. Sample was then ramped down to 25°C at a rate of 5 °C/min. Annealed oligos were subsequently ligated in to BbsI-digested pCFD3 (using T4 DNA ligase).



Sense:	5′ – GTC**G**-N19/20
Anti-sense:	5′ – AAAC-N19/20 reverse complement

Figure 2. 1. Characteristics of the pCFD3 vector for expression of a single gRNA
(top) Sequence map of the U6:3 promoter, BbsI spacer, and gRNA core sequences in the pCFD3 vector used for cloning and subsequent expression of a single gRNA. (bottom) Template for the design of sense and antisense oligos to clone gRNAs into BbsI-digested pCFD3. The desired gRNA sequence is 20bp, or 19bp if it begins with a G. The GTCG and AAAC overhangs allow sticky-end ligation of annealed oligos into digested pCFD3.

For cloning two gRNAs into pCFD4, primers specific to protospacer sequences were used in a PCR reaction, using undigested pCFD4 as template. The 5' protospacer was incorporated in the forward primer (U6-1 promoter) and 3' in the reverse primer (U6-3 promoter) for no specific reason, since both promoters effectively drive gRNA expression in the *Drosophila* germline (Port et al. 2014). PCR product was purified by gel extraction, and cloned into BbsI-digested pCFD4 vector by Gibson assembly.

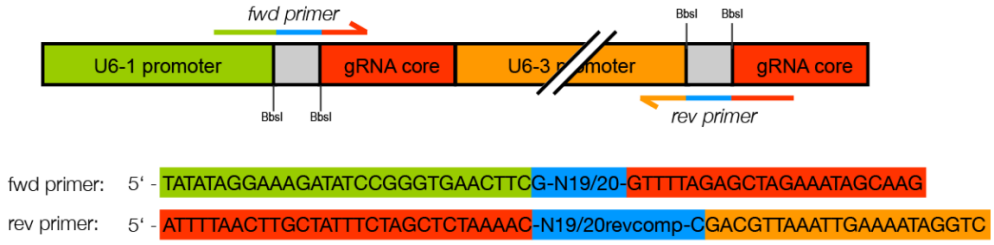


Figure 2. 2. Characteristics of the pCFD4 vector for expression of two gRNAs. schematic of the gRNA-expressing cassette in the pCFD4 vector. One gRNA is driven by the U6-1, while the other the U6-3 promoter. Two primers are used to amplify a cassette containing sequences for the two gRNAs of choice, using pCFD4 as template. If the gRNA begins with a G, then N=19, otherwise N=20. The reverse primer sequence has the reverse complement of the desired gRNA sequence. Primer overhangs allow for subsequent Gibson assembly of PCR product into BbsI-digested pCFD4.

2.3.6.2 Homology-directed repair and rescue vectors used in insertion-ready deletion allele generation

A homology-directed repair (HDR)-repair template is also required for initial gene targeting, and a rescue vector is required for subsequent DNA reintegration during locus rescue. Both vectors were kindly provided by Dr Cyrille Alexandre.

The HDR-repair template vector used for generation of these alleles was pTV3 (**Fig. 2.3A, Appendix 8**) (Baena-Lopez et al. 2013; Poernbacher et al. 2019). There are many significant components of this vector: two multiple cloning sites, which allow the cloning of 5' and 3' homology arms. These are separated by a targeting cassette designed to be incorporated in the genome, which includes: an attP landing site to allow the subsequent site-specific recombination of rescue constructs into the

generated IRD alleles; a *Pax-cherry* marker gene, which is the CDS for mCherry under 3XP3 regulatory sequences. This drives expression of cherry in the developing larval CNS and photoreceptors of the adult fly (Horn et al. 2002; Baena-Lopez et al. 2013); Finally, loxP sites flank the *Pax-cherry* cassette and allow its excision through the expression of cre-recombinase.

The vector used to rescue IRD alleles was RIV^{White} (**Fig. 2.3C**, **Appendix. 9**). This vector contains an attB landing site complimentary to that of attP in pTV3, which is immediately followed by an MCS used to clone in the desired rescue construct. This is followed by a loxP site, which again allows cre-recombinase mediated excision of excess vector DNA if required. The marker gene in this vector is *white*. Therefore, given that IRD alleles are generated in a w^- background, successful integration of the rescue construct is marked by red eye colour.

A technical note in the generation of these alleles is related to the fact that this technique is not scarless, and how this relates to the enzyme choice when cloning homology arms into pTV3; or rescue DNA fragments in to RIV^{White} (**Fig. 2.3)**. For the former, it relates to the enzymes used to clone the 3' end of the 5'HA, and the 5' end of the 3'HA. For the latter, it relates to the enzymes used to clone in the rescue fragment at the 5' end. This is because these DNA sequences will ultimately increase the size of the 'scar' sequence present in the finally rescued allele, and this could have an effect on gene function. Therefore, where possible enzymes were chosen to minimise the effect of this. However, this was limited by whether restriction sites were present in the fragments we were attempting to clone.

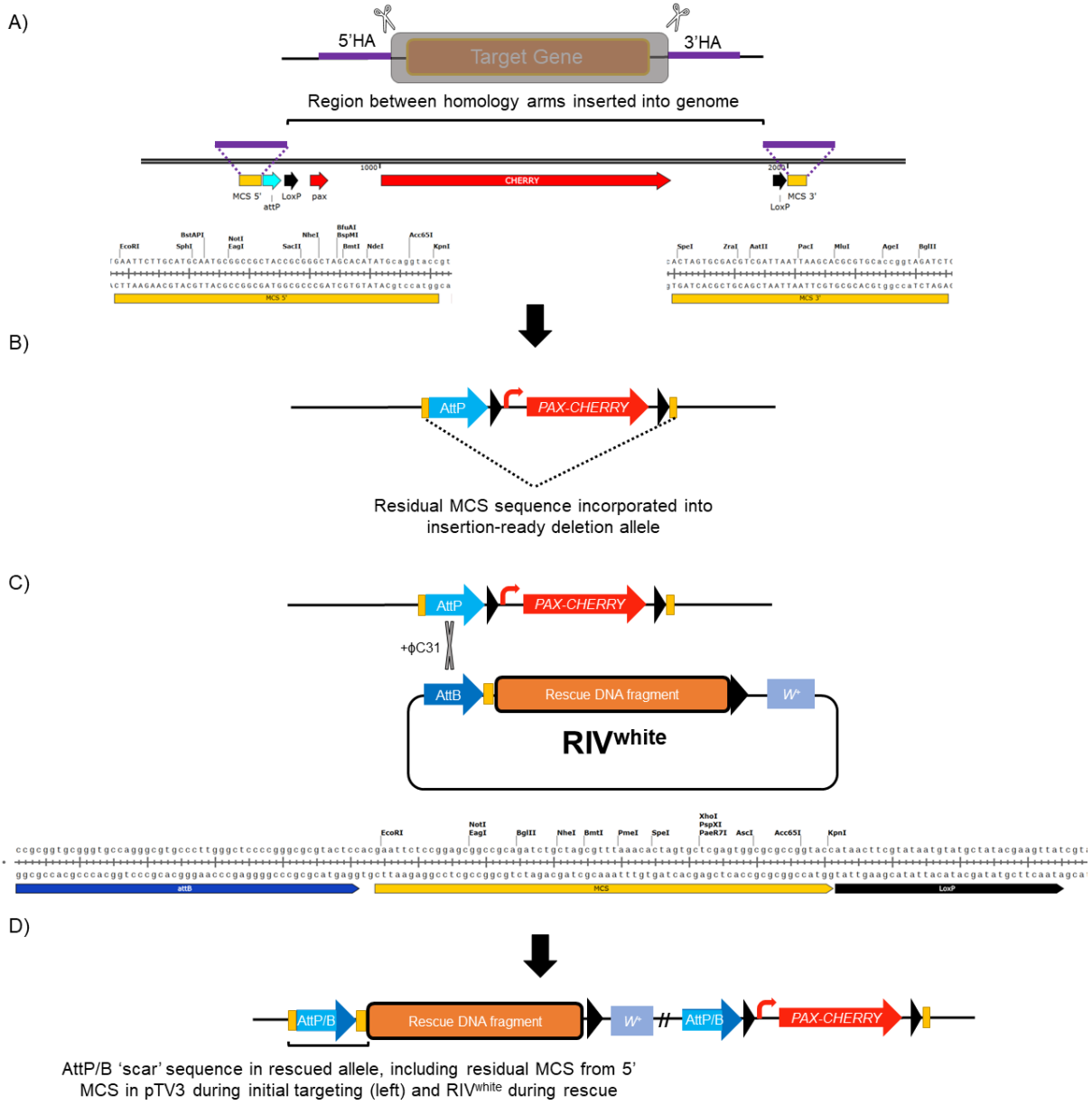


Figure 2. 3. Influence of multiple cloning site selection on allele generation.

(**A**) top: Schematic of target gene target to generate insertion-ready deletion allele. Homology arms (purple) are designed to flank to Cas9-induced DSB (✂). Bottom: schematic of the pTV3 targeting cassette and multiple cloning sites (MCS) where the homology arms are cloned. Restriction enzymes within each MCS are indicated. (**B**) Schematic showing, depending on the restriction enzyme used to clone in the homology arms, additional sequence (yellow) can be incorporated into the IRD allele. (**C**) top: schematic showing recombination of rescue vector, RIV^{white} into the IRD allele. Bottom: Schematic of the attB sequence followed by the MCS sequence. Depending on the choice of restriction enzyme used, residual MCS sequence (yellow) may also be retained in the rescue vector (**D**) Additional sequence, derived from the choice of restriction enzyme used to clone in the rescue DNA fragment, ultimately influences the size of the AttB/P 'scar' sequence present in the rescued allele. This has most influence at the 5' end of the engineered gene.

2.3.6.3 constructing pTV3 homologous recombination targeting vectors for generation of Mef2/HDAC4 insertion-ready deletion alleles

Homology arms were cloned into pTV3 using restriction digest-ligation cloning. For Mef2, 1.6kb or 1kb homology arms, for the ΔCDS and ΔATG deletions respectively were amplified by PCR with Q5 HiFi 2X master mix. Homology arm length was determined based on previous use of CRISPR-Cas9 in *Drosophila* (Gratz et al. 2013; Alexandre et al. 2014; Gratz et al. 2014; Poernbacher et al. 2019) and increased accordingly due to the predicted size of the Cas9-mediated DSB (Alexandre, C personal communication). Primers were designed such that the homology arm ended at the location of the Cas9-induced double-strand break, which occurs between nucleotides 17 and 18 of the protospacer sequence (**Fig. 2.6**) (Ran et al. 2013). Homology arms were cloned sequentially into digested pTV3, using NotI and NheI to clone the 5' arm, and SpeI and AatII for the 3' arm (see **Fig. 2.3A**) . For HDAC4, 1.6kb homology arms were amplified using PCRBIO VeriFi™ 2x Mix and cloned into pTV3 using the same procedure, however AatII and AgeI were used to clone the 3' homology arm. Completed vectors were verified by restriction enzyme digestion and DNA sequencing prior to microinjection.

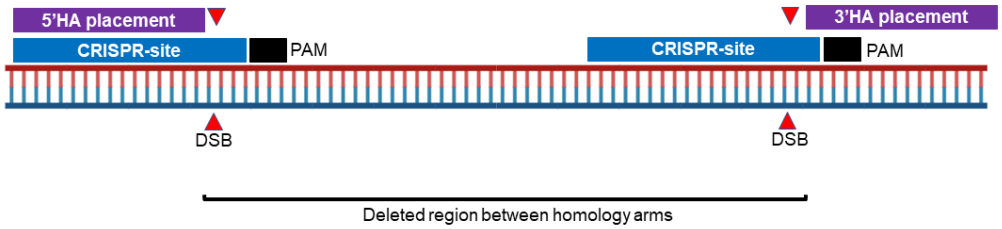


Figure 2. 4 Homology arm placement design in the generation of insertion-ready deletion alleles

Schematic showing the relative placement of homology arms (HA) when designing insertion-ready deletion alleles. HA's were designed adjacent to the Cas9 induced double-strand break (DSB) DSB break-point.

2.3.6.4 constructing RIV^{White} rescue constructs for recombination into insertion-ready deletion alleles

To clone rescue DNA fragments into RIV^{White}, I used gDNA isolated from the original injection line when inserting a gDNA rescue fragment. For the cDNA based rescue of

Mef2 IRD alleles, I used cDNA from Mef2 isoform C (which was also used in the generation of all UAS-Mef2 overexpression constructs), combined with gDNA amplified from the injection line. Most fragments were cloned in to RIV^{White} in a single step, using the restriction enzymes indicated in the primer name. However, for Mef2 cDNA rescue, two independent PCR fragments were cloned in sequentially, again utilising the restriction enzymes in the primer names.

For gDNA rescue of *Mef2*$^{\Delta CDS}$ and *Mef2*$^{\Delta ATG}$, gDNA was amplified using Mef2_rscu_5'UTR-F+EcoRI forward primer, and either the Mef2_gDNArscu_CDS-R+XhoI or Mef2_gDNArscu_ATG-R+XhoI reverse primer for the *Mef2*$^{\Delta CDS}$ or *Mef2*$^{\Delta ATG}$ respectively.

For the cDNA rescue, the same forward primer was used as for gDNA rescue, however the construct was generated via the independent amplification of two fragments. One fragment, termed the CDS fragment, was amplified from Mef2 cDNA using the Mef2_cDNArscu_CDS-R+XhoI primer. The second, termed the UTR fragment, was amplified from gDNA using the Mef2_cDNArscu_3'UTR-F+XhoI and Mef2_cDNArscu_3'UTR-R+KpnI primers.

The HDAC4 gDNA rescue fragment was amplified using primers HDAC4_gDNArscu_F+NotI and HDAC4_gDNArscu_R+NheI.

2.3.6.5 homology-directed repair vectors for direct tagging

HDR vectors for direct tagging of *HDAC4* and *Mef2* were also kindly provided by Dr Cyrille Alexandre. There are two vectors, used for either the addition of an N-terminal or a C-terminal tag (**Fig. 2.4**). By contrast to pTV3, the HDR vectors for direct tagging were digested with BsaI and SapI, and the 1kb homology arms were cloned in using Gibson assembly rather than restriction cloning. Homology arms were PCR amplified from gDNA extracted from the injection stock. These homology arms flank a targeting cassette, which differ between both vectors.

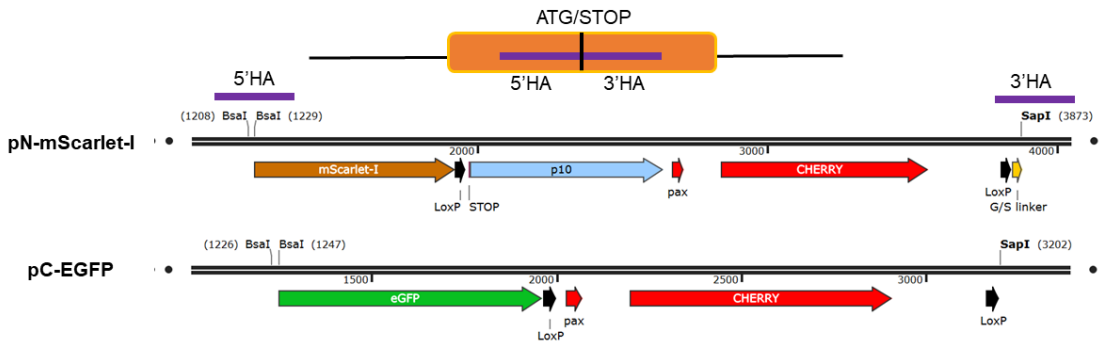


Figure 2. 5 General characteristics of the N- and C-terminal tagging vectors
(Top) general schematic showing an exon (orange), and a position indicating either an ATG start, or a STOP codon. Homology arms (HA, purple) are placed either side of the codon to ensure in-frame insertion of the desired tag. (Bottom) Tagging cassette maps for pScarlet-I or pC-EGFP. Homology arms are cloned into their respective vectors by Gibson assembly, following vector digestion with BsaI (5' HA) and SapI (3'HA).

For the N-terminal tagging vector, termed pN-mScarlet-I (**Appendix. 10**), the arms flank the cassette containing: the coding sequence for the mScarlet fluorophore, which is followed by the 3'UTR from the p10 baculovirus. This is followed by the same *Pax-cherry* transgenic marker as used in pTV3. The p10 3'UTR and *Pax-cherry* are flanked by loxP sites. This holds particular significance in this allele. The p10 3'UTR promotes transcription termination, and thus the initial allele generated using this vector disrupts transcription of the endogenous gene, theoretically generating a null allele. At the same time, the mScarlet-I fluorophore is transcribed and expressed under the regulation of the endogenous gene which has been targeted, and thus should also generate a transcriptional reporter. Generation of an in-frame N-terminal tag is dependent upon cre-mediated excision. Between the end of the mScarlet CDS, and the tagged gene CDS is a glycine-serine linker, formed through the residual loxP scar, and some additional DNA. In this project, the mScarlet-H fluorophore within the initial pN-mScarlet vector was modified to the mScarlet-I fluorophore using site-directed mutagenesis. Residues which required mutating were identified previously during original fluorophore synthesis (Bindels et al. 2016) . The primers used were: Scarlet-I T74I-F/R and Scarlet-H H164M-F/R.

The C-terminal tagging, termed pC-EGFP or pC-mScarlet vector is more straightforward, consisting of the CDS for EGFP, which is followed by a loxP-flanked *Pax-cherry* targeting cassette (**Appendix. 10**). In this project, the EGFP fluorophore

was also swapped out for the mScarlet-I fluorophore to generate pC-mScarlet-I (**Appendix. 11**).

2.3.6.6 Construction of a novel vector, pC-mScarlet-I for C-terminal tagging with mScarlet-I

To construct a novel vector, pC-mScarlet-I for ability to generate C-terminally tagged mScarlet-I proteins, the EGFP-tag within pC-EGFP was removed through restriction digestion with the enzymes BsiWI and XmaI (**Fig. 2.5**), and the vector backbone was gel-purified. Independently, the equivalent region within pN-mScarlet-I vector was PCR amplified using primers Scarlet_Fwd+XmaI and Scarlet_Rev+BsiWI. Purified PCR product was digested with BsiWI and XmaI, and then cloned into the pC vector backbone to generate pC-mScarlet-I (**Appendix. 11**).

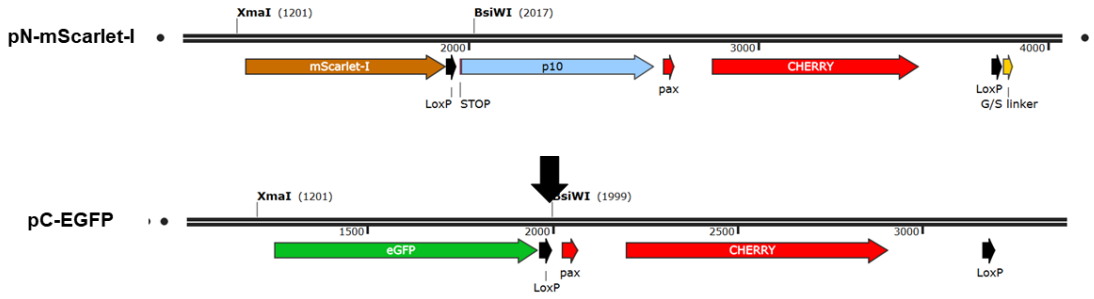


Figure 2. 6 Enzymes used to construct a novel vector, pC-mScarlet-I, for C-terminal tagging with mScarlet-I.
Schematics over the original cassette maps for the pN-mScarlet-I and pC-EGFP vectors. The enzymes XmaI and BsiWI were used to cut the EGFP CDS out of pC-EGFP, and replace it by ligating in a XmaI and BsiWI-digested PCR product encoding the mScarlet-I fluorophore.

2.4 Microscopy and data analysis

2.4.1 Imaging systems used to image *Drosophila* samples

Routine fly pushing was done using either a Nikon SMZ-2B or Nikon SMZ 745 stereo microscope. For data collection, a combination of compound fluorescence, and confocal microscopy was used during this project. An Olympus BX53 microscope was used for all standard fluorescence microscopy imaging. For genotype selection whereby fluorescent markers were used, a Leica MSV269 microscope was used in conjunction with a Leica Kubler Codix light source. A Zeiss Axioskop 2 was used for the imaging of thoracic cross sections stained with hematoxylin. All confocal microscopy was done using in the Cardiff University School of Biosciences Bioimaging

Hub, using a Zeiss LSM880 laser-scanning confocal microscope. Image analysis was done using Fiji (Schindelin et al. 2012). Image reconstruction required to generate single images of somatic muscle structure in L3 larvae was done using “stitching” plugin (Preibisch et al. 2009).

2.4.2 Protein structures

All protein structures analysed in this project were obtained from the Protein Data Bank (PDB, available at https://www.rcsb.org/). The files used were: 1TQE, for the analysis of the Class IIa HDAC/Mef2 interaction (Han et al. 2005); and 3P57 for the analysis of the p300/Mef2 interaction (He et al. 2011). Protein structures were downloaded in PDB format and analysed using the molecular visualisation software, PyMOL.

2.4.3 DNA sequence analysis

The majority of DNA sequence analysis was done using SnapGene (https://www.snapgene.com/). This includes: the generation and annotation of sequence and vector maps, analysis of raw DNA sequence traces, and alignments. For some DNA sequence alignments, online sequencing alignment tools from the European Bioinformatics Institute (EBI) were used: EMBOSS stretcher (available at: https://www.ebi.ac.uk/Tools/psa/emboss_stretcher/), or Clustal omega (available at: https://www.ebi.ac.uk/Tools/msa/clustalo/) programs were used for either pairwise, of multiple-sequence alignments respectively. The DNA analysis software, Jalview was used for the analysis of outputs using the EBI sequence tools (Procter et al. 2021).

2.4.4 Figure generation, sample sizes and statistical analysis

All graphs were generated using Microsoft Excel. Figures were mainly synthesised in Microsoft Powerpoint, saved as PNG files, and directly inserted into the thesis document. Some figures were generated using BioRender. Statistical analysis to quantify severity of larval muscle, and DLM phenotypes was carried out in GraphPad Prism. Statistical significance was obtained using the non-parametric Kruskal-Wallis test followed by Dunn’s multiple comparisons test. A minimum n of 10 (either larvae or flies) was used where statistical analysis was undertaken. For the analysis of premature differentiation of L3 AMPs between genotypes, a minimum of 20 discs were imaged per genotype. Embryonic muscle phenotypes were analysed qualitatively and representative images are presented.

Chapter 3: Investigating the ability of HDAC4 to regulate muscle differentiation *in vivo*

Introduction

As introduced in section **1.5.4,** previous reports *in vitro* have identified Class IIa HDACs as a potential regulator of somatic muscle differentiation. By mainly investigating two of the vertebrate Class IIa HDACs, HDAC4 and HDAC5, it has been shown that they physically interact with the transcription factor Mef2, and repress its transcriptional activation of target genes. Furthermore, the expression of these Class IIa HDACs can directly inhibit the differentiation of muscle cell line myoblasts into differentiated myotubes *in vitro*. Thus, these data suggest that Class IIa HDACs can act as negative regulators of muscle differentiation *in vitro* by physically interacting with, and inhibiting Mef2. However as yet, there has been no systematic investigation into the possible role of Class IIa HDACs in regulating Mef2 in an *in vivo* model for muscle differentiation, possibly contributed to by the presence of four Mef2, and four Class IIa HDAC genes in vertebrates. Given the presence of only one Mef2 and Class IIa HDAC genes, *Drosophila* represents an ideal model system for an investigation into Class IIa HDAC function during muscle differentiation *in vivo*. This chapter is dedicated towards initially furthering our understanding of Class IIa HDAC function during muscle development to investigate the possible function of the *Drosophila* Class IIa HDAC, *HDAC4*. I use the development of both the larval somatic musculature during embryogenesis, and adult DLM muscle fibers as models to investigate whether HDAC4 possesses the same ability to inhibit muscle differentiation *in vivo*, addressing these key research questions:

1. Akin to *in vitro* models, can *HDAC4* overexpression inhibit muscle differentiation *in vivo*?
2. If so, does HDAC4 inhibit muscle differentiation by physically interacting with, and repressing Mef2?

3.1 HDAC4 expression in primordial muscle tissue during embryogenesis inhibits their formation

The first aim was to address the question of whether, akin to previous *in vitro* model systems (McKinsey et al. 2000a; Miska et al. 2001), Class IIa HDACs can inhibit muscle differentiation in the context of a developing organism *in vivo*. To do so, a new transgenic fly line was generated, containing the coding sequence for the *Drosophila* Class IIa HDAC, HDAC4, under the control of Gal4-responsive UAS sites and a minimal promoter (**Appendix 5**). This UAS-HDAC4 construct was inserted into the attP40 landing site, which displays high expression in muscle tissue and limited 'leaky' expression in the absence of Gal4 (Markstein et al. 2008).

I chose to overexpress HDAC4 in the context of both embryonic and adult somatic muscle development. Not only are these two discrete stages of myogenesis, but Mef2 is required for both embryonic (Bour et al. 1995; Ranganayakulu et al. 1995; Taylor et al. 1995) and adult myogenesis (Bryantsev et al. 2012; Soler et al. 2012). Thus, both stages of myogenesis may also be suitable for investigating HDAC4 function in the context of Mef2 regulation. To overexpress *HDAC4* during embryonic muscle development, the TwipTwip-Gal4 driver line was used, which harbours a Gal4 transgene under the control of the regulatory sequences for *Twist* on both chromosomes II and III to maximise transgene expression (Baylies and Bate 1996). Akin to *Twist* expression, this line expresses Gal4 from early in embryogenesis, from before subdivision of the mesoderm occurs (Baylies and Bate 1996). Because the muscle pattern of late-stage embryo is well characterised (Beckett and Baylies 2006), it provides an ideal model system whereby defects in muscle pattern because of *HDAC4* expression can be visualised through antibody staining.

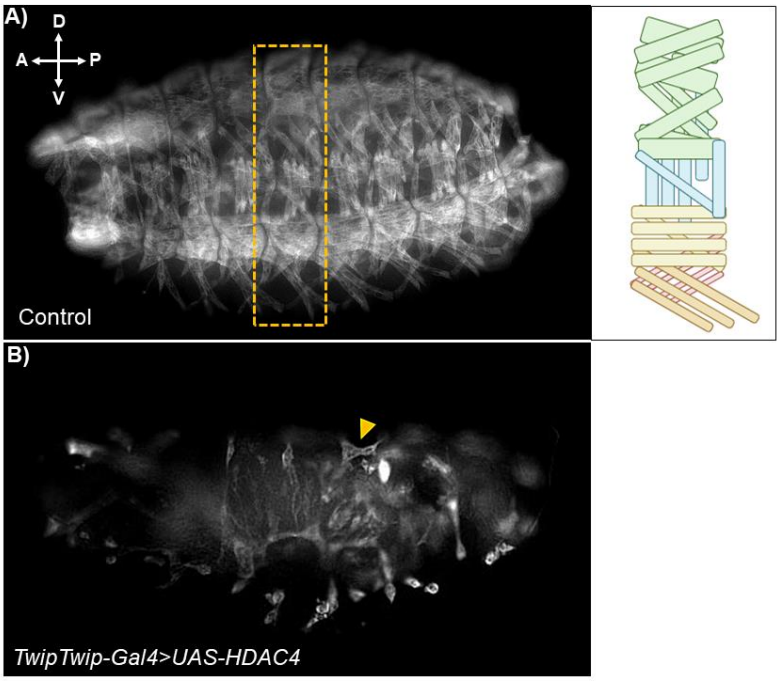


Figure 3. 1. Overexpression of *HDAC4* in the developing somatic muscle primordia inhibits formation of the embryonic somatic musculature.
Late stage 16 control embryos (**A**) or embryos in which HDAC4 is overexpressed using the TwipTwip-Gal4 driver (two copies of Twist-Gal4) (**B**). Embryos were stained using an antibody against the differentiated muscle marker, Myosin heavy chain (Mhc). (**A**) Control embryos display a highly organised and well characterised pattern of 30 individual muscle fibres per abdominal hemisegment. (**B**) When HDAC4 is overexpressed, somatic muscle formation is strongly inhibited and embryos fail to form any differentiated muscle fibres. A small amount of Mhc staining is observed in what appear to be small, partially differentiated syncytia which have failed to differentiate further (arrow.).

To visualise the differentiated muscle pattern, embryos were stained with an antibody against Myosin heavy chain (Mhc) (**Fig. 3.1**). Compared to control embryos (**Fig. 3.1A**), *HDAC4* overexpression completely inhibited the formation of differentiated muscle in all embryos (**Fig. 3.1B**): Embryos lack differentiated muscle fibres resembling that of a wild-type embryo, suggesting a strong inhibition to the myogenic differentiation pathway. There appear to be small, multinucleate, Mhc-expressing syncytia (**Fig. 3.1B arrow**), possibly representing fused, partially differentiated myoblasts which have been unable to differentiate further. However, they are present at random locations throughout the embryo, not consistently located in a position of where embryonic muscles would normally be. This suggests that none of the muscles may be more resistant to HDAC4-mediated inhibition than others.

3.2 HDAC4 expression inhibits formation of the DLM indirect-flight muscles

Given that HDAC4 can strongly inhibit the development of the larval somatic musculature during embryogenesis, I next asked whether HDAC4 can similarly inhibit the formation of the somatic musculature during adult myogenesis. To do so, the dorso-longitudinal indirect-flight muscles within the adult thorax (DLMs) were used as a model. They form through the fusion of wing imaginal disc-associated AMPs with larval muscle templates, which escape histolysis during metamorphosis. To drive *HDAC4* expression, the 1151-Gal4 was used, which is expressed in all AMPs associated with the wing imaginal disc from the second instar onwards and continues to be expressed in AMPs and immature, developing fibres until approximately 40hrs APF (Anant et al. 1998; Weitkunat and Schnorrer 2014). As there is a consistent pattern of six individual muscle fibres per hemithorax, phenotype severity was quantified by counting the number of fibres per hemithorax in flies of each genotype. This has been done previously to quantify DLM phenotypes (Soler et al. 2012).

The regimented DLM fibre pattern is seen in the 1151-Gal4 control, where the fibres fill a large proportion of the adult thorax and are aligned either side of the midline of the thorax (**Fig. 3.2A**). However, when *HDAC4* is overexpressed, DLM formation is strongly inhibited and the mean DLM number per hemithorax is significantly reduced (**Fig. 3.2E**). A range of severity in the DLM phenotype between individual flies is observed: In most hemithoraces (64%) there is a complete inhibition of muscle differentiation, with DLMs not forming at all (**Fig. 3.2B)**. In other cases, some DLMs are present, however they appear irregularly shaped and poorly positioned relative to the control (**Fig. 3.2C-D**).). In 10% of hemithoraces, DLM development appeared to be unaffected.

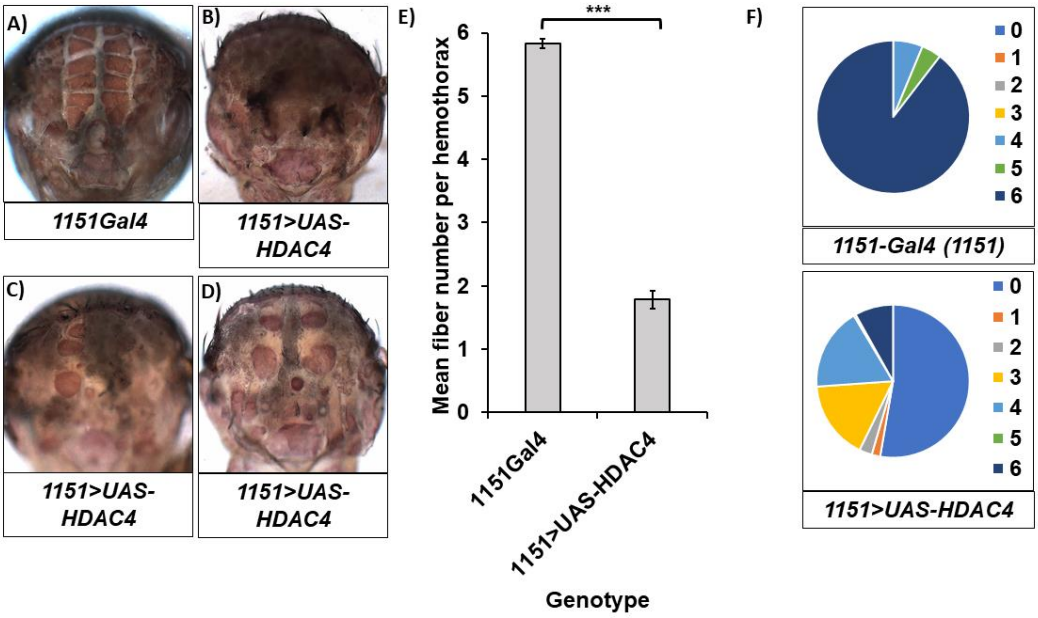


Figure 3. 2. Overexpression of HDAC4 in adult muscle progenitors inhibits formation of the dorso-longitudinal flight muscles (DLMs).
(A-B) Transverse thoracic cross sections of either control black pupae **(A**), or pupae in which *HDAC4* is overexpressed using the 1151-Gal4 driver **(B). (A)** In control flies, an organised pattern of 6 pairs of DLM fibres form within the adult thorax, aligning either side of the thorax midline. **(B-D)** When *HDAC4* is overexpressed, DLM formation is strongly inhibited. We observe variable penetrance of this phenotype in terms of number of fibres formed, including: **(B)** complete ablation of any fibres; or **(C-D)** a reduced number of fibres. **(E)** Mean DLM fibre number is significantly reduced upon *HDAC4* overexpression. Bar graph showing mean DLM fibres present in black pupae (~96hr APF) of each genotype. Error bars represent S.E.M. Significance *** represents $p<0.001$ from Kruskal-wallis followed by Dunn's multiple comparisons test. **(F)** Distribution of DLM fibres observed per hemithorax in control and *HDAC4* overexpression.

3.3 Investigating Mef2 regulation in the context of myogenesis: Mef2 induces premature differentiation of L3 AMPs

I next wanted to investigate whether the phenotypes associated with *HDAC4* overexpression were caused by negatively regulating Mef2 activity. To do so, I first investigated the effect of overexpressing Mef2 in an undifferentiated myoblast population. The AMPs associated with the wing imaginal disc are an attractive model to study Mef2 function during muscle differentiation for three reasons. Firstly, endogenous Mef2 is expressed in these cells prior to their differentiation (Cripps et al. 2004; Soler and Taylor 2009). Secondly, the 1151-Gal4 driver can be used to

manipulate Mef2 expression. And finally, the relative ease of accessing and analysing the tissue (Tripathi and Irvine 2022).

Here, the functional consequences of overexpressing two different UAS-Mef2 constructs, inserted into either the attP40 (ChrII) or attP2 (ChrIII) landing sites, were tested (**Fig. 3.3A**). The first, UAS-Mef2 10t4a, contains the coding sequence for Mef2 isoform III including 532bp of 5' UTR and 1.2kb of 3' UTR. This is a construct that has been used to study Mef2 function in our lab previously (Gunthorpe et al. 1999; Elgar et al. 2008). This line has been used most often and is why it was inserted onto most chromosomes to ensure most flexibility when working with other transgenic lines. The second construct, UAS-Mef2 CDS contains only the coding sequence for UAS-Mef2 isoform III, but is followed by a C-terminal, 3xHA epitope tag. I generated this line only on ChrIII and was used predominantly for additional experiments in **Chapter 6**.

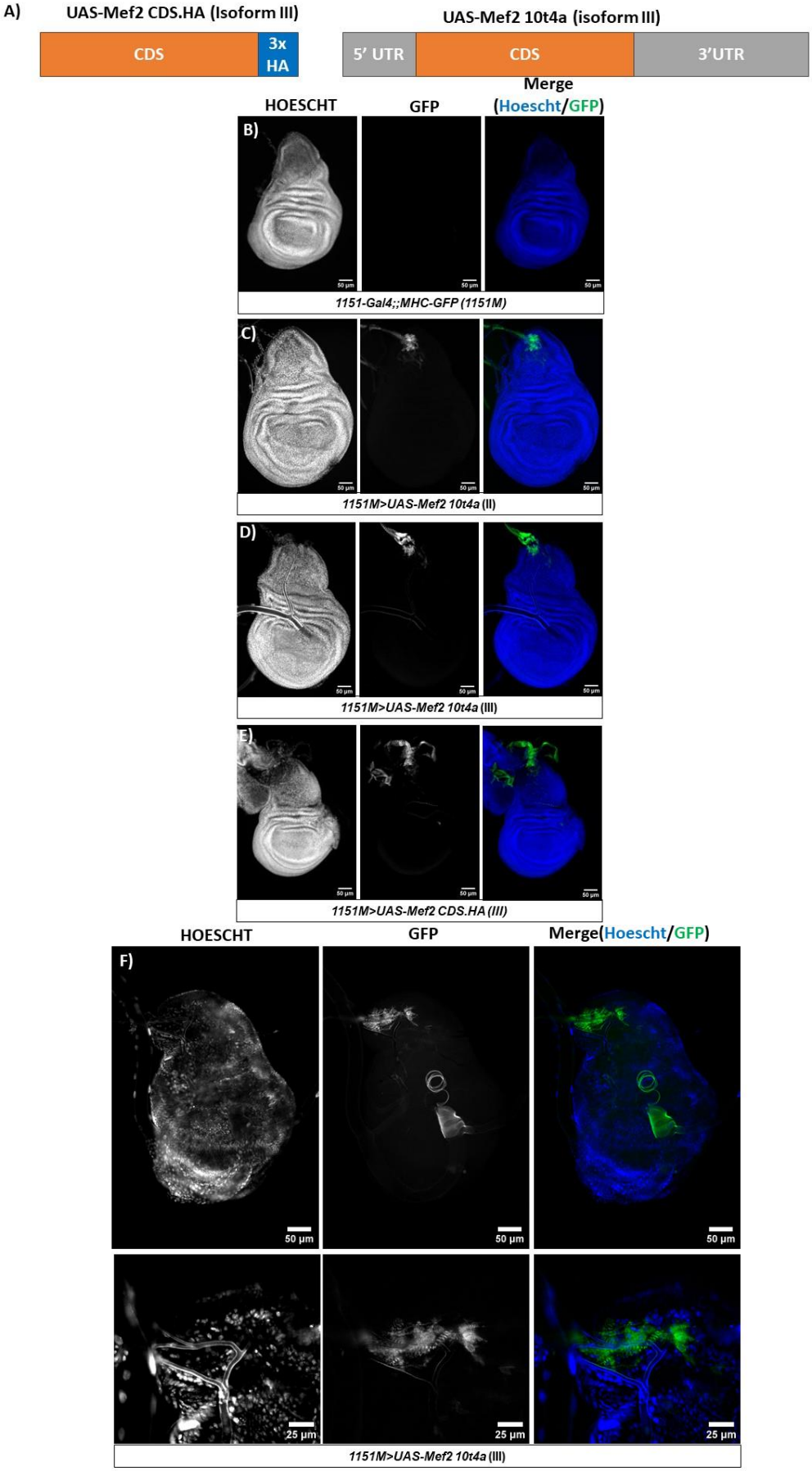

A)
UAS-Mef2 CDS.HA (Isoform III)
CDS
3x HA
UAS-Mef2 10t4a (isoform III)
5' UTR
CDS
3'UTR
HOESCHT
GFP
Merge (Hoescht/GFP)
B)
1151-Gal4;;MHC-GFP (1151M)
C)
1151M>UAS-Mef2 10t4a (II)
D)
1151M>UAS-Mef2 10t4a (III)
E)
1151M>UAS-Mef2 CDS.HA (III)
HOESCHT
GFP
Merge(Hoescht/GFP)
F)
50 µm
25 µm
1151M>UAS-Mef2 10t4a (III)

Figure 3. 3. Overexpression of Mef2 induces the premature differentiation of Adult muscle progenitors (AMPs).
(A) Schematic of two different UAS-Mef2 constructs generated and tested. (left) UAS-Mef2 CDS harbours the Mef2 coding sequence (CDS) for isoform III, plus a C-terminal 3xHA epitope tag. (Right) UAS-Mef2 10t4a harbours the same CDS, however is flanked by 532bp and 1.2kb of 5' and 3'UTR respectively, with no epitope tag. UAS-Mef2 CDS was inserted only onto ChrIII, while UAS-Mef2 10t4a was inserted onto both ChrII and ChrIII. **(B-E)** Premature differentiation of L3 AMPs, observed through the expression of Myosin heavy chain (Mhc), was observed in all UAS-Mef2 transgenic lines **(C-E),** but not in control (**B**). **(F)** Mhc forms striations in these AMPs, indicating the formation of sarcomeres in these prematurely differentiating AMPs.

Previous evidence has shown that Mef2 overexpression in the AMPs induces the premature expression of *Mhc*, a terminal muscle differentiation marker and a known Mef2 target gene (in the embryo), or the expression of a tau-GFP reporter under control of *mhc* regulatory sequence (Lovato et al. 2005; Sandmann et al. 2006; Soler and Taylor 2009). I used a more recently generated Mhc-GFP, which is a C-terminally tagged *mhc-GFP* transgene under control of endogenous regulatory sequences (Sarov et al. 2016). This method in theory combined the benefits of the published techniques: The ability to detect Mhc, expressed under its endogenous regulatory mechanisms; but, if the GFP signal is strong enough to detect after fixing alone, without the requirement to antibody stain. Indeed, this technique was successful in detecting GFP-signal after fixation alone, without the need for antibody staining. Expression of either UAS-Mef2 10t4a or UAS-Mef2 CDS induced premature Mhc expression in the AMPs, which was not present in the control (**Fig. 3.3B-E**). This phenotype was observed in 100% of wing discs, regardless of construct, or insertion site. Interestingly, and which has not been previously reported, striations were observed in the Mhc-GFP signal (**Fig. 3F**). This suggests these AMPs are not only expressing Mhc, but have begun to terminally differentiate and undergo sarcomerogenesis. Further investigation in the lab has also confirmed expression of other terminal differentiation markers (Rob Mitchell, unpublished data in MVT lab). Interestingly, while this phenotype is 100% penetrant, the premature differentiation is consistently located in the dorsal-most region of the notum, despite Mef2 overexpression throughout the AMP population.

3.4 HDAC4 expression suppresses the Mef2-induced differentiation of L3 AMPs

I next asked whether the phenotype caused by Mef2 overexpression could be inhibited by HDAC4, which could indicate that HDAC4 may negatively regulate Mef2 activity *in vivo (***Fig. 3.4***)*. To ensure *Mhc* expression is not affected by dilution of Gal4 protein, caused by the presence of two UAS-transgenes, UAS-mCherry was also expressed in addition to UAS-Mef2. Unlike UAS-Mef2 alone which, as previously mentioned, induced the premature differentiation of AMPs in 100% of discs (**Fig.3.3, Fig.3.4B**), co-expression of UAS-HDAC4 completely suppressed the premature differentiation phenotype, with 0% discs presenting with the Mef2-induced differentiation phenotype (**Fig.3.4C**). Moreover, co-expression of UAS-mCherry with UAS-Mef2 did not affect the ability for Mef2 to induce Mhc expression in AMPs (**Fig. 3.4D**). This suggests that HDAC4 can directly inhibit a Mef2-dependent phenotype *in vivo*.

I also asked whether, given the co-expression of Mef2 and HDAC4 suppresses Mef2-induced premature differentiation phenotype, this co-expression of Mef2 and HDAC4 can rescue the HDAC4-induced inhibition of DLM formation. Interestingly, co-expression of Mef2 only partially restored the DLM phenotype. The mean DLM number of 3.55 was significantly higher than UAS-HDAC4 alone, however it was also significantly lower than controls (**Fig. 3.4E)**. In agreement with this, the proportion of hemithoraces scored with a wild-type number of DLM fibres rose from 10% when UAS-HDAC4 is expressed alone, to 40% when UAS-Mef2 is co-expressed with UAS-HDAC4 (**Fig. 3.4F**).

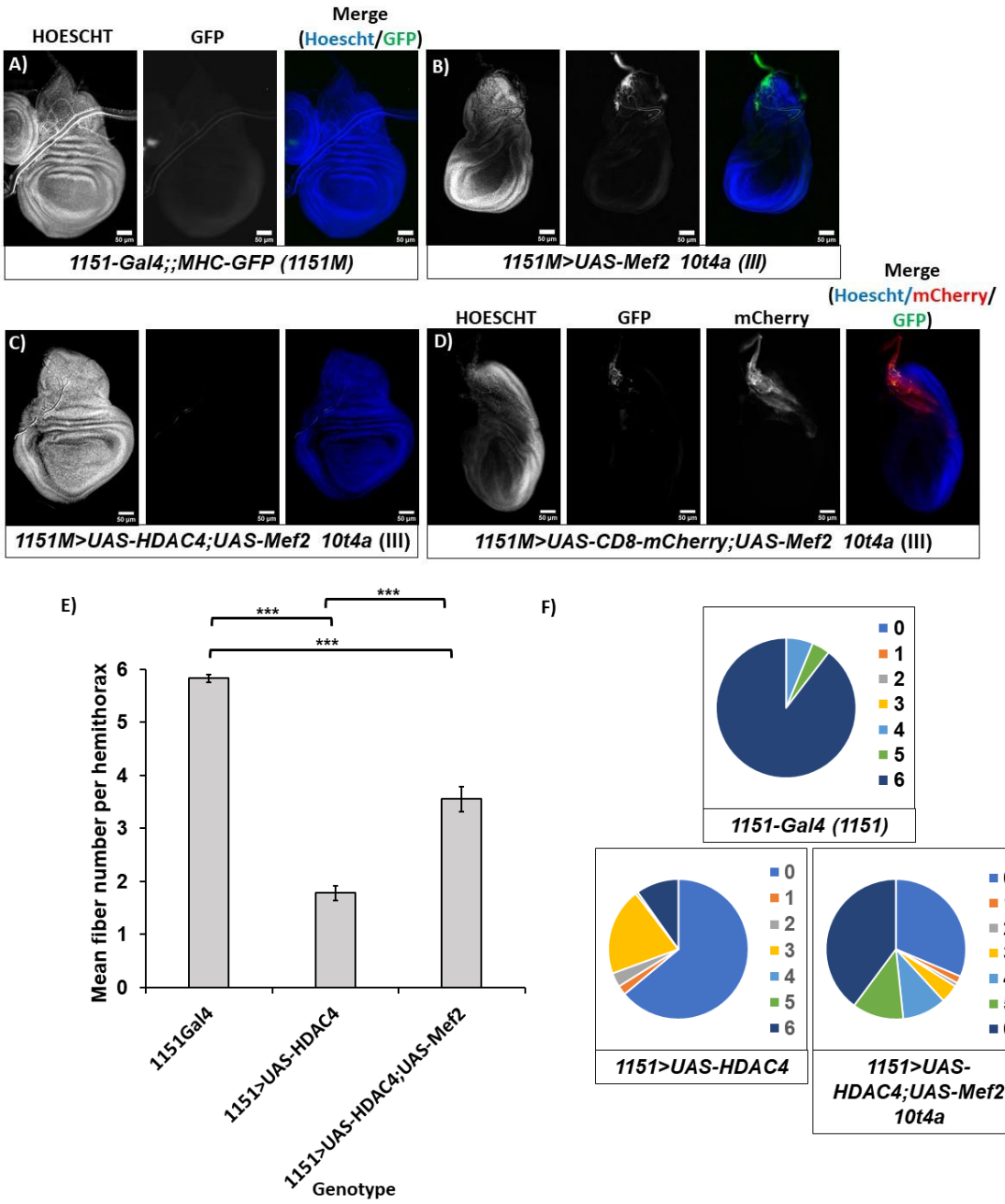


Figure 3. 4. HDAC4 can suppress the Mef2-induced premature differentiation of adult muscle progenitors (AMPs).

Premature differentiation visualised through the expression of Mhc-GFP not seeing in controls **(A),** but observed following overexpression of Mef2 in L3 AMPs using the 1151-Gal4 driver (**B**) . **(C)** When HDAC4 is co-expressed with Mef2, premature Mhc expression is repressed. **(D)** Co-expression of mCherry and Mef2 does not inhibit the Mef2-induced differentiation phenotype. **(E-F)** Mef2 co-expression only partially rescues the HDAC4-induced inhibition of DLM development. (**E**) Bar graph showing mean DLM fibres present in black pupae (~96hr APF) of each genotype. Error bars represent S.E.M. Significance *** represents p<0.001 from Kruskal-wallis followed by Dunn's multiple comparisons test. **(F)** Pie charts showing distribution of DLM fibres observed for each genotype.

3.5 Targeted mutagenesis of the Mef2-interacting domain of HDAC4

The ability to suppress a Mef2-induced phenotype in L3 AMPs demonstrates HDAC4 can function as a negative regulator of a differentiation phenotype caused by Mef2. However, it does not show whether the HDAC4-induced inhibition of both embryonic and somatic muscle differentiation is due to its interaction with Mef2, especially given that UAS-Mef2 co-expression did not fully rescue the UAS-HDAC4-induced DLM phenotype. To address this, site-directed mutagenesis was used to mutate residues within the conserved Mef2-interacting domain of HDAC4 (**Fig. 3.5A**). A pymol render of the published HDAC/Mef2 interaction, and residues mutated during this project, is shown in **Fig. 3.5B-D**. Residues were targeted based on *in vitro* characterisation of important residues for this interaction (Wang and Yang 2001; Han et al. 2005; Jayathilaka et al. 2012; Main et al. 2021). Leucine-168 (L168) lies at the centre of the Mef2-HDAC interaction (**Fig. 3. 5C,D**), and mutation of the homologous residue in human HDAC4 (Leucine-175) diminished the interaction with Mef2D when assayed by a mammalian-2-hybrid (Jayathilaka et al. 2012). The homologous residue of *Drosophila* Phenylalanine-171 (F171), mammalian F178 has also been shown biochemically to reduce the interaction between HDAC4 and Mef2B by approximately tenfold (Han et al. 2005). Thus, I generated novel transgenic lines harbouring mutant UAS-HDAC4 constructs to assess their ability to inhibit muscle differentiation *in vivo*. I generated UAS-HDAC4 L168A and UAS-HDAC4 F171A, while I also received a premade construct harbouring the mutations K165A, L168A and I172A. This was designed to abolish the Mef2-HDAC interaction and has been recently published as a Mef2-binding mutant in *Drosophila* (Main et al. 2021). Prior to microinjection, new constructs were fully sequenced to confirm the desired change (**Fig. 3.5E**).

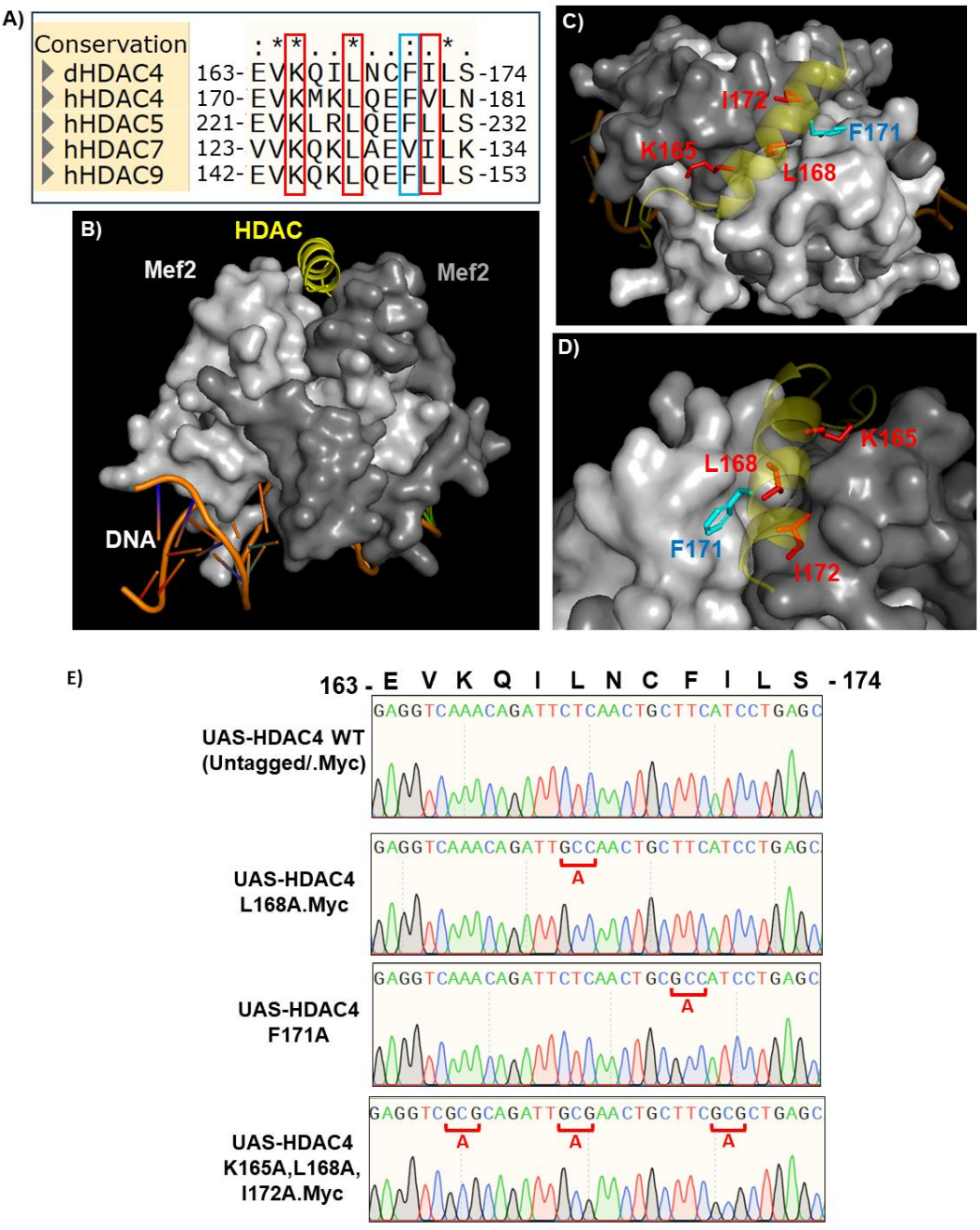


Figure 3. 5. Conservation and mutagenesis of the conserved Mef2-interacting domain of Class IIa HDACs.

(A) Protein sequence alignment of the Mef2-binding domain of different *Drosophila* and vertebrate Class IIa HDACs. Outlined residues indicate those residues mutated in UAS-HDAC4 constructs and are further described in **(B-D).** Crystal structure representation of the Mef2B:HDAC9 complex solved by Han et al (2005). **(B)** A Mef2 dimer (light grey/dark grey) is bound to DNA, while the Mef2-binding domain of HDAC9 (yellow) forms an amphipathic helix that sits in a groove formed between the two Mef2 monomers. **(C-D)** Detailed residues important for the Mef2:HDAC physical interaction. Numbers indicated are those residues for the equivalent *Drosophila* HDAC4 protein, and for graphical representation, L172 of HDAC9 has been artificially replaced by an isoleucine present in *Drosophila* HDAC4. The side chains of K165, L168, F171, and I172 protrude into the groove formed between the Mef2 monomers,

and make physical interactions with Mef2. Colours relate to the specific mutations which were made during the project: UAS-HDAC4 L168A and UAS-HDAC4 K165,L168,I172A (red), and UAS-HDAC4 F171A. **(E)** DNA sequence traces for UAS-HDAC4 and UAS-HDAC mutant lines generated during this project. Alanine mutations are highlighted in red.

I first confirmed whether UAS-HDAC4 mutants were expressed *in vivo*. As all were inserted into the same genomic location, transgene expression was controlled for. However, to test whether mutant proteins were stably expressed, the 1151-Gal4 driver was used to analyse transgene expression in the L3 AMPs. Because there is no commercially available antibody for HDAC4, most UAS-HDAC4 constructs were generated with a C-terminal myc epitope tag. Antibody staining showed strong expression throughout the AMP population for UAS-HDAC4 WT (**Fig. 3.6B**), UAS-HDAC4 L168A (**Fig. 3.6C**), and UAS-HDAC4 K165,L168,I172A (**Fig. 3.6D**), but not in 1151-Gal4 control (**Fig. 3.6A**), suggesting all UAS-HDAC4 lines are stably expressed *in vivo*. This supports additional data suggesting UAS-HDAC4 WT and UAS-HDAC4 K165,L168,I172A are expressed at similar level, when analysed by immunoblotting of cell lysates (Helen Fitszimons, unpublished). One construct, UAS-HDAC4 F171A, was not tagged and thus I could not confirm protein expression with this line. This was not tagged because it was generated during the first round of construct generation, which included UAS-HDAC4 WT (untagged), UAS-HDAC4 F171A, and UAS-HDAC4 K165,L168.I172A.myc (pre-synthesised). Other UAS-HDAC4 transgenics described above (UAS-HDAC4.myc, L168A.Myc) were generated later. Nevertheless, given that the UAS-HDAC4 K165,L168,I172A mutant is stably expressed, it is unlikely that the F171A mutation alone would cause drastic stability effects not observed when other residues within this domain are mutated.

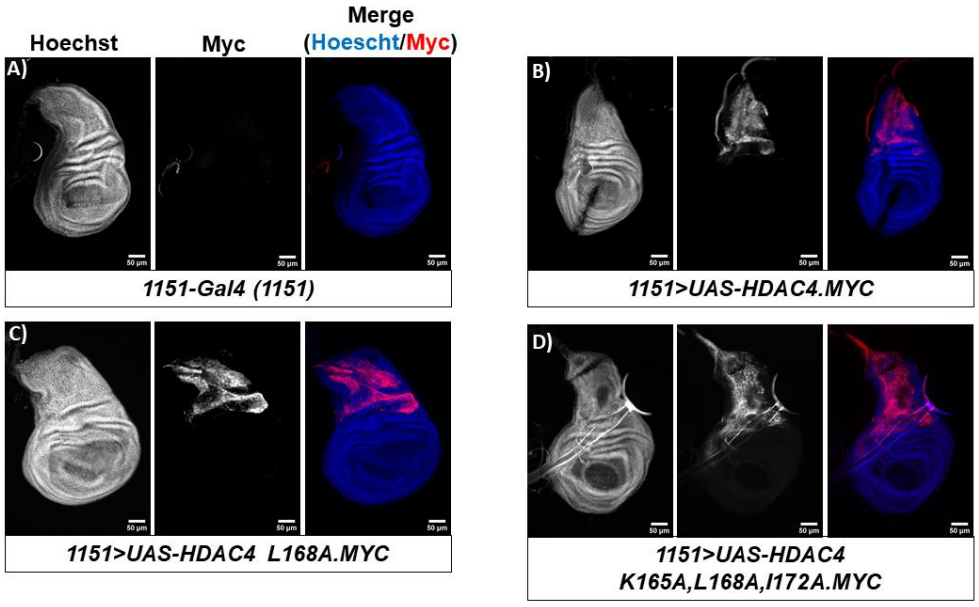


Figure 3. 6. UAS-HDAC4 and UAS-HDAC4 mutants are highly expressed.
Overexpression of HDAC4 and HDAC4 mutants in adult muscle progenitors, detected using an antibody against the c-myc epitope tag. (**A**) No positive staining was detected in the driver-only control, while HDAC4 WT (**B),** HDAC4 L168A **(C),** and K165,L168,I172A (**D**) are detected throughout the AMP population, suggesting that UAS-HDAC4 and UAS-HDAC4 mutant transgenes are stably expressed *in vivo*.

3.6 UAS-HDAC4 mutants form an allelic series when expressed in the developing larval somatic musculature.

To investigate whether the HDAC4-mediated inhibition of embryonic myogenesis (**Fig. 1**) was a result of inhibiting Mef2, UAS-HDAC4 mutants were expressed using the TwipTwip-Gal4 driver. While UAS-HDAC4 WT strongly inhibited formation of the embryonic somatic musculature (**Fig. 3.7B**), expression of UAS-HDAC4 K165,L168,I172A did not affect somatic muscle differentiation, with the differentiated muscle patter in late stage embryos unaffected (**Fig. 3.7D**). Furthermore, these embryos were viable, whereas those expressing UAS-HDAC4 WT were embryonic lethal (not quantified). Interestingly, UAS-HDAC4 F171A expression caused an intermediate phenotype. Muscle differentiation was still severely affected, with the gross muscle architecture failing to form (**Fig. 7D**). However, this phenotype was not as severe as when UAS-HDAC4 WT was expressed. Some muscle fibres which appear to be in the correct positions within the embryo do form. In particular, the lateral-transverse (LT) and ventral-lateral muscles (VL) muscles appeared to be less severely affected than that of other muscles throughout the embryo. This contrasts to

UAS-HDAC4 WT, where no muscle fibres resembling any of the 30 well-characterised fibres form (**Fig. 3.7B**). I did not analyse the effect of UAS-HDAC4 L168A during embryonic myogenesis.

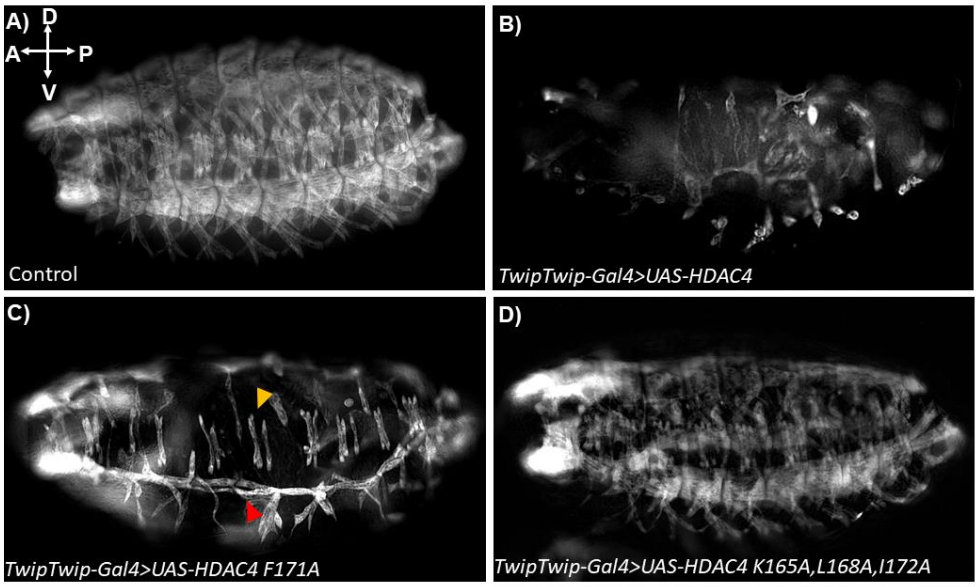


Figure 3. 7. HDAC4 mutants with defective Mef2 binding domains form an allelic series when expressed in the developing embryonic somatic musculature.
Different UAS-HDAC4 lines present with a differing phenotypic severity in their ability to inhibit the differentiation of the embryonic somatic musculature when expressed with the TwipTwip-Gal4 driver (2 copies of Twist-Gal4). Embryonic somatic muscle pattern was visualised in late stage 16 embryos using an antibody against Myosin heavy chain (Mhc) (**A**) Control, (**B)** UAS-HDAC4 WT, (**C**) UAS-HDAC4 F171A, (**D**) UAS-HDAC4 K165,L168,I172A. While UAS-HDAC4 WT can strongly inhibit muscle differentiation (**A**), the triple mutant K165,L168,I172A fails to do so and the embryonic somatic musculature appears phenotypically normal **(D). (C)** UAS-HDAC4 F171A displays an intermediate muscle phenotype in which muscle differentiation is affected, however more Mhc staining is observed than when UAS-HDAC4 WT is overexpressed. Arrows indicate either the lateral transverse (orange) or ventral lateral (red) muscles which appear least affected.

3.7 UAS-HDAC4 mutants form an allelic series when expressed in the developing adult somatic musculature.

I next investigated the effect of expressing these same UAS-HDAC4 mutants on the development of the DLMs, using the 1151-Gal4 driver. Similarly to that observed during embryonic myogenesis, expression of UAS-HDAC4 mutants differentially affected DLM formation (**Fig. 3.8**). Firstly, HDAC4 WT expression strongly inhibited DLM development (**Fig. 3.8B**). Importantly, UAS-HDAC4 WT.myc also strongly inhibited DLM development and did not yield a significantly different phenotype when

analysed by mean DLM fibre number (**Fig. 3.8B,C,G**). By contrast, UAS-HDAC4 K165,L168,I172A expression did not affect DLM formation, and instead the fibres were indistinguishable from controls (**Fig. 3.8F**). This result is analogous to when this mutant was expressed in the developing embryonic somatic musculature (**Fig. 3.7D**). Interestingly, mutation of L168 alone was also sufficient to completely suppress the inhibition of DLM formation. UAS-HDAC4 L168A expression also yielded a DLM phenotype indistinguishable from controls (**Fig. 3.8E**). Similar to in the embryo, UAS-HDAC4 F171 expression yielded an intermediate phenotype. DLM development was still strongly inhibited (Mean fibre number = 4.5), however this phenotype was not as severe as after expression of UAS-HDAC4 WT. Together, along with the phenotypes observed when HDAC4 and HDAC4 mutants are expressed in developing embryonic muscle, these data suggest that, when expressed in undifferentiated myoblasts, Mef2 binding is likely necessary for HDAC4-mediated inhibition of myogenesis *in vivo*.

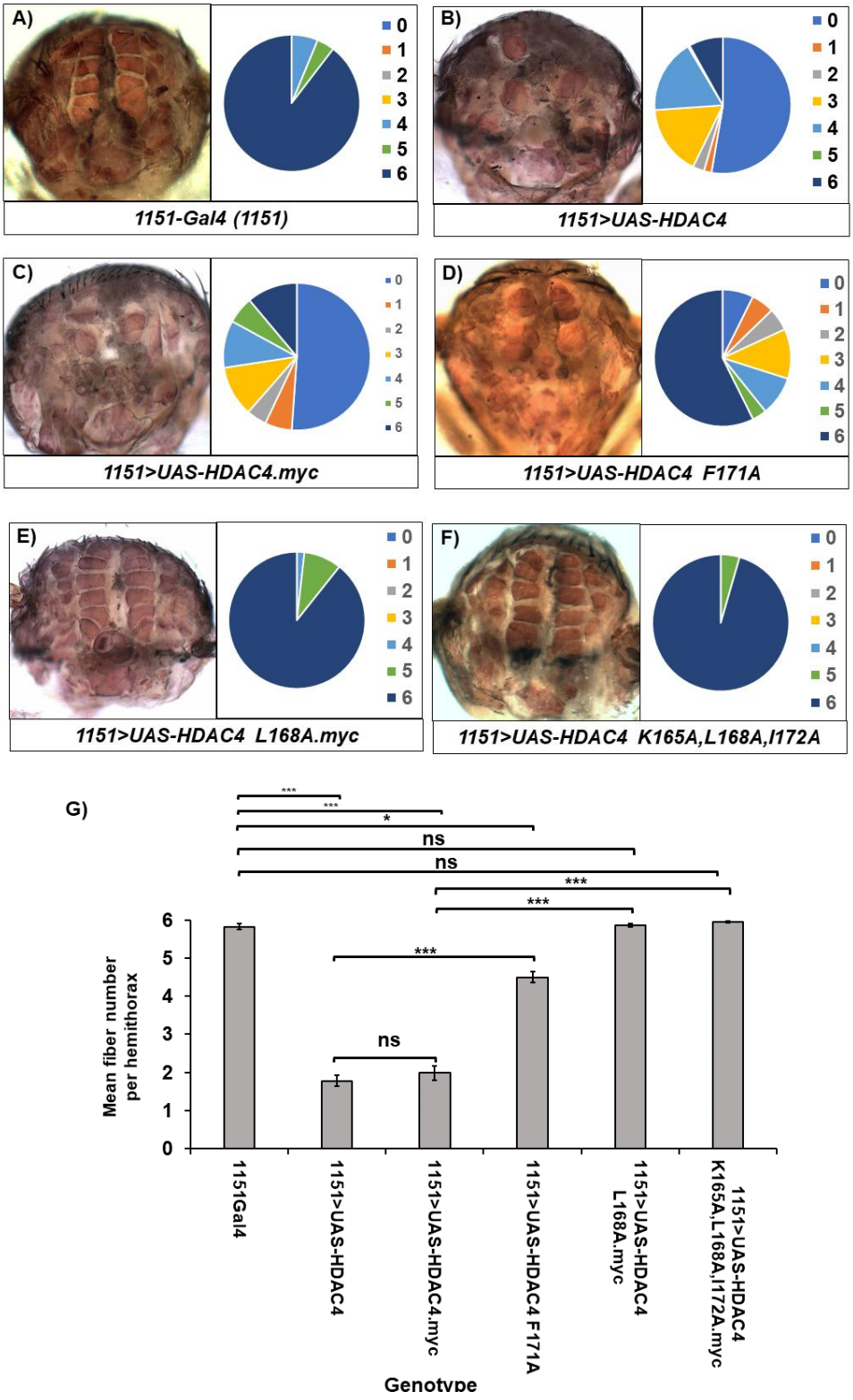


Figure 3. 8. HDAC4 mutants with defective Mef2 binding domains form an allelic series when expressed in the developing dorso-longitudinal flight muscles (DLMs).
(A-F) Transverse thoracic cross sections of black pupae at 96h APF to visualise DLM flight muscles, and associated fibre count distribution per hemithorax of different genotypes of fly expressing UAS-HDAC4 transgenes with 1151-Gal4. **G)** Mean DLM fibre number per genotype. Error bars represent S.E.M. *** p<0.001, ** p<0.005, * p<0.01, ns = not significant.

calculated from Kruskal-wallis followed by Dunn's multiple comparisons test. Overexpression of HDAC4 WT **(B)** or HDAC4 WT.myc **(C)** both strongly inhibit DLM development, however overexpression of the L168A mutation **(E)** or the K165,L168,I172A triple mutant **(F)** both fail to inhibit DLM development. Overexpression of the F171A mutant **(C)** causes an intermediate phenotype less severe than HDAC4 WT overexpression, but more severe than controls or the L168A/K165,L168,I172A mutants.

3.8 An in-tact Mef2 binding domain is not required for HDAC4 to inhibit muscle differentiation when expressed in immature muscles, post-fusion.

These data extend previous *in vitro* data (McKinsey et al. 2000a; Miska et al. 2001) to an *in vivo* model, whereby HDAC4 likely negatively regulates Mef2 activity in undifferentiated myoblast populations during *Drosophila* muscle differentiation. However, these data do not consider the role of Class IIa HDACs during myogenesis during maturation following myoblast fusion, whereby the developing DLM myotube fibres split (from 3 per hemithorax to 6), attach to tendons, grow dramatically to fill the entire thorax, and undergo myofibrillogenesis (Spletter et al. 2018). To investigate whether HDAC4 may regulate myogenesis in the events following myoblast fusion, Actin88F-Gal4 was used, an IFM-specific driver that expresses in myotubes, but not unfused myoblasts, from approximately 25hr APF onwards (Bryantsev et al. 2012). *HDAC4* overexpression in this context strongly inhibited normal DLM development, with almost a complete lack of differentiated DLM fibres (**Fig. 3.9B**). To further investigate whether, in this context, the inhibition is dependent upon an in-tact Mef2 binding domain, UAS-HDAC4 L168A and UAS-HDAC4 K165,L168,I172A were expressed using the same driver. Interestingly, and by contrast to their effect when expressed using 1151-Gal4, these two HDAC4 mutants still strongly inhibited normal DLM development (**Fig. 3.9C-D**), and the severity of this phenotype was not significantly different to UAS-HDAC4 WT overexpression. To confirm whether these phenotypes were a result of non-specific transgene expression within the immature myotubes adversely affecting their development, as opposed to a HDAC4-mediated effect, a UAS-mCherry transgene was also expressed as a control, which did not affect DLM development (**Fig. 3.9E**).

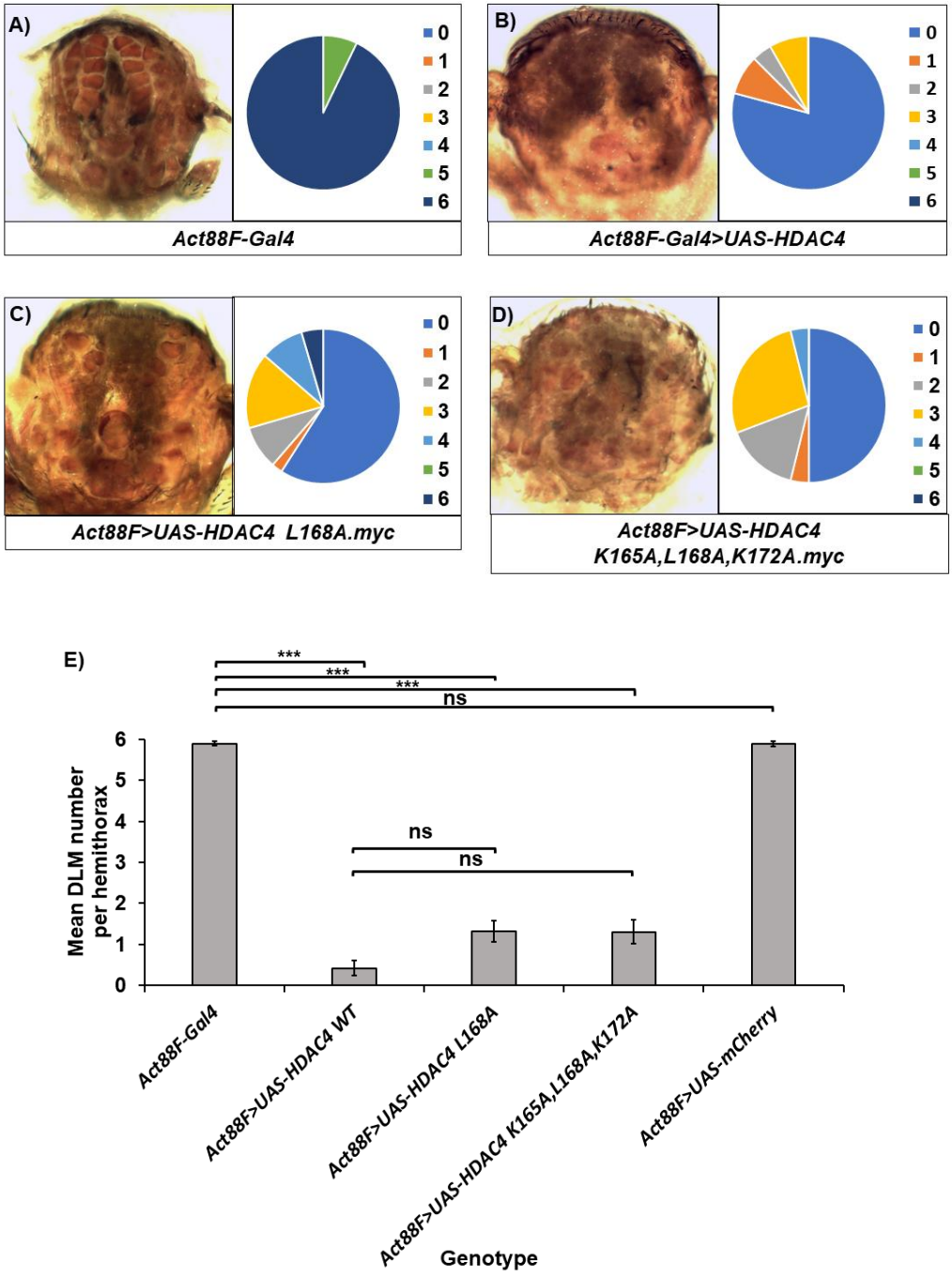


Figure 3. 9. An in-tact Mef2 binding domain is not required for HDAC4 to inhibition DLM formation when expressed post-fusion.

(A-D) Transverse thoracic cross sections of black pupae at 96h APF to visualise DLM flight muscles, and associated fibre count distribution per hemithorax of different genotypes of fly expressing UAS-HDAC4 transgenes with Act88F-Gal4. Mean DLM fibre number per genotype. Error bars represent S.E.M. *** $p<0.001$, ** $p<0.005$, * $p<0.01$, calculated from Kruskal-Wallis followed by Dunn's multiple comparisons test. HDAC4 WT and HDAC4 Mef2-binding domain mutants all strongly inhibit DLM formation.

3.9 Investigating the requirement of a conserved catalytic residue for HDAC4-mediated repression of myogenesis:

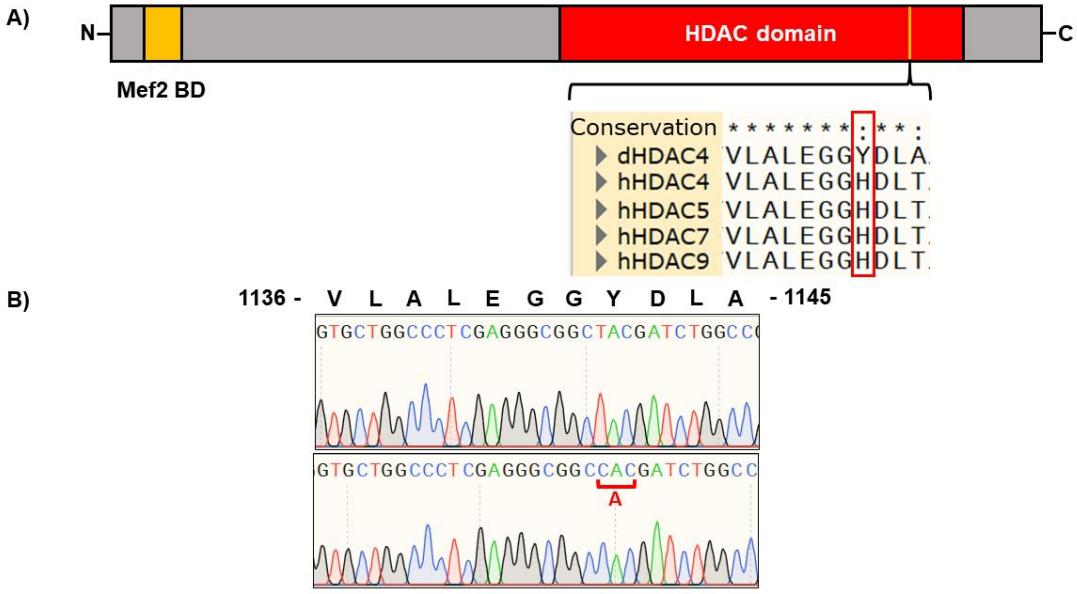
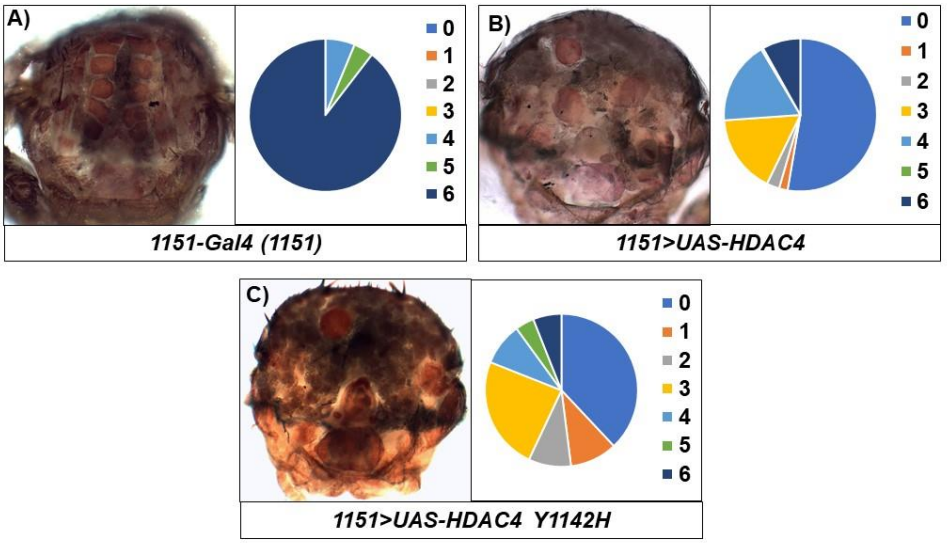


Figure 3. 10. Conservation and mutagenesis of a conserved residue in the HDAC domain of HDAC4.
(A) Schematic of the HDAC4 protein structure, with the Mef2-binding and HDAC domains shown. Below is a protein sequence alignment of the 'LEGGX' catalytic motif within the HDAC domain. The catalytic tyrosine residue is conserved in *Drosophila*, but mutated to a histidine in vertebrate Class IIa HDACs. (**B**) DNA sequence traces showing of the UAS-HDAC4 Y1142H mutant construct generated in this project. Codon mutated to an alanine is indicated.

Vertebrate Class IIa HDACs are catalytically inactive enzymes. This appears to be a direct result of is a tyrosine-histidine mutation within the catalytic domain, which causes a 1000-fold decrease in catalytic activity *in vitro* compared to Class I HDACs (Lahm et al. 2007). Interestingly, *Drosophila* HDAC4 retains this catalytic tyrosine, and has been postulated to thus retain catalytic activity (**Fig. 3.10A**) (Lahm et al. 2007). To investigate the potential functional significance of this sequence conservation on the ability for HDAC4 to inhibit myogenesis *in vivo*, I generated a UAS-HDAC4 Y1142H mutant by site-directed mutagenesis, and inserted the transgene into the same attP40 landing site as all other UAS-HDAC4 transgenes generated during the project (**Fig. 3.10B**). To assess whether this residue is required for the inhibition of myogenesis, I expressed UAS-HDAC4 Y1142H using the 1151-Gal4 driver. UAS-HDAC4 Y1142H retained the ability to strongly inhibit DLM formation and yielded a phenotype not

significantly different from the overexpression of UAS-HDAC4 WT alone (**Fig. 11**). This indicates that this conserved tyrosine is not required for the inhibitory phenotype associated with HDAC4 over-expression.



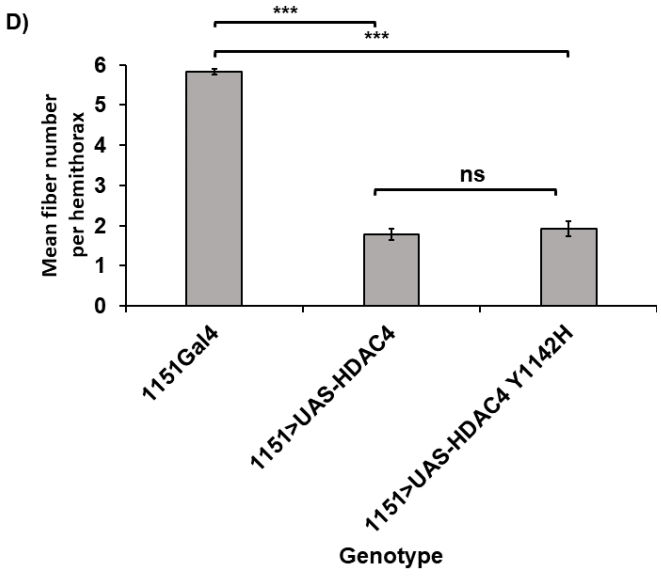


Figure 3. 11. A conserved tyrosine is not required for HDAC4-mediated inhibition of DLM formation

(A)Transverse thoracic cross sections of black pupae at 96h APF to visualise DLM flight muscles, and associated fibre count distribution per hemithorax of different genotypes of fly expressing UAS-HDAC4 transgenes with 1151-Gal4. (**B**) Bar graph showing mean DLM fibres present in black pupae (~96hr APF) of each genotype. Error bars represent S.E.M. Significance *** represents $p<0.001$ from Kruskal-wallis followed by Dunn's multiple comparisons test.

Results summary

By generating new UAS-HDAC4 transgenic fly lines, I have shown here that the *Drosophila* Class IIa HDAC, HDAC4, can strongly inhibit muscle differentiation when expressed in undifferentiated primordial muscle tissue or myoblasts. This phenotype is also dependent upon an in-tact Mef2-binding domain, suggesting HDAC4 inhibits myogenesis in this context by repressing Mef2. Interestingly, when expressed later in IFM development, after myoblast fusion and during muscle growth and maturation, both wild-type and mutant HDAC4 defective in Mef2 binding can also inhibit muscle differentiation, suggesting HDAC4 may inhibited muscle differentiation via an alternative mechanism later in development. Independent of the Mef2-interacting domain, I also show that a conserved tyrosine residue implicated in Class IIa HDAC catalytic activity is not required for the ability of HDAC4 to inhibit myogenesis when expressed in undifferentiated myoblasts during IFM development. These data in this chapter thus support *in vitro* models of muscle differentiation whereby Class IIa HDACs *can* inhibit muscle differentiation *in vivo*, however does not establish the function of endogenous HDAC4 in muscle differentiation. These data support the further investigation into the role of endogenous HDAC4 during *Drosophila* myogenesis *in vivo,* which will be addressed in Chapter 4.

Chapter 4: Development of a novel, rescuable HDAC4 null allele for the study of HDAC4 function *in vivo*

Introduction

In chapter 3, I have shown that overexpression of *HDAC4* in various contexts can inhibit the normal differentiation of the *Drosophila* somatic musculature *in vivo*. However, while these experiments further the established *in vitro* work to an *in vivo* model, these data are limited to being able to conclude that HDAC4 *can* inhibit muscle differentiation, rather than providing evidence as to what the function of endogenously expressed *HDAC4* is during muscle differentiation *in vivo*. Indeed, while more technically challenging, understanding the function of endogenous loci has been dramatically aided through the development of genome engineering tools. Most notably and widely used has been CRISPR-Cas9, which has been introduced in section **1.6.2** and **1.6.3**. CRISPR-Cas9 utilises the Cas9 nuclease isolated from Streptococcus pyogenes to induce double-stranded breaks at specific genomic locations. This specificity is achieved using an RNA-guided mechanism, whereby Cas9 forms a complex with a sgRNA molecule containing a 20bp 'protospacer' sequence that guides Cas9 to the desired genomic location and thus determining the double-strand break site. The breaks can then either be repaired through the error-prone NHEJ, or through homology-directed repair which, when a specific repair template is provided, may be used to induce very specific genetic changes within the genome. Indeed, since first being reported in 2012, CRISPR-Cas9 has been adopted for use in many model organisms, including *Drosophila*, where many genetic tools now exist for its effective implementation for the study of gene function. One adapted use of CRISPR-Cas9 developed in *Drosophila* has been the development of insertion-ready deletion alleles (introduced fully in section **1.6.3**), which link CRISPR-Cas9 genome engineering to site-specific recombination. By doing so, this system first allows the generation of a null allele, before rescuing this null allele through site specific recombination of genetic elements into the engineered locus in order to rescue gene function. Therefore, this system could allow a high level of versatility with regards to studying the function of any endogenous gene, however in particular *HDAC4*. In this chapter, I utilise CRISPR-Cas9 genome engineering to address these research aims:

1. Develop a rescuable *HDAC4* insertion-ready deletion allele which functions as a novel *HDAC4* null, rescuable by site-specific recombination of genetic elements into the affected *HDAC4* locus.

2. Use this *HDAC4* insertion-ready deletion allele to investigate the role, and potential functions of endogenous HDAC4 during muscle differentiation *in vivo*.

4.1 Targeting the *HDAC4 locus:* design of a targeted deletion to generate a *HDAC4* loss-of-function allele

The *HDAC4* locus is located on the X chromosome (11E8-11E9) and is large and complex. There are eight annotated isoforms, two transcription start sites, and three in-frame ATG start codons (**Fig. 4.1**). In total, the gene spans a region of over 23kb, and the region of coding DNA, that is, coding exonic sequence, spans over 17kb for isoforms utilising ATG1 (**Fig. 4.1**). However, coding exons are not equally distributed along the gene's length, with most coding sequence being clustered within a region spanning less than 5.75kb toward the 3' end of the locus. This region also encodes many of the functionally relevant domains of the HDAC4 protein, including the Mef2-binding domain, deacetylase domain, NLS, and 14-3-3 protein binding sites (**Fig. 4.1B**).

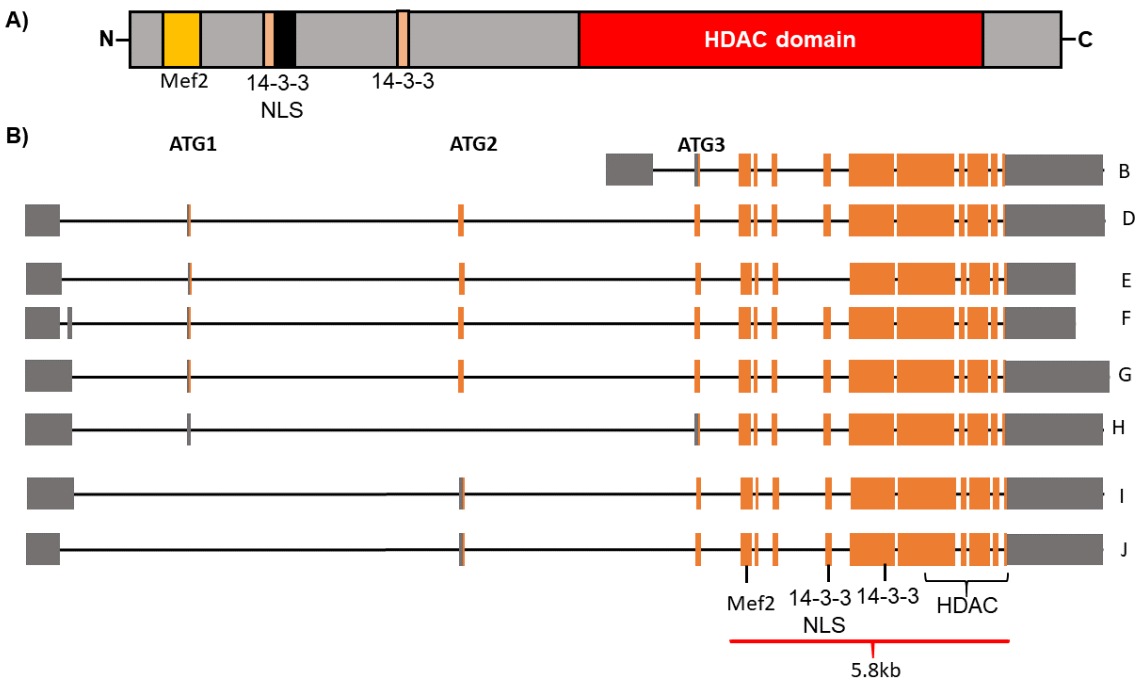


Figure 4. 1 HDAC4 protein structure and locus organisation.
(A) schematic of the HDAC4 protein with relative positions of functionally characterised domains including the Mef2-binding site, nuclear localisation sequence (NLS), 14-3-3 protein binding sites, and the HDAC domain. (B) Different characterised *HDAC4* isoforms (B-J), annotated on flybase. The three annotated START codons (ATG1-3), and regions encoding the functional domains described in **A** are indicated. 5.8kb = size of genomic region encoding functionally relevant domains. Non-coding and coding exons are in grey and orange, respectively.

This information regarding the *HDAC4* locus was significant when considering an appropriate design to generate an insertion-ready deletion allele for *HDAC4.* This is because it was desired to generate a *HDAC4* null allele first, and then have the potential to rescue with a variety of different genetic elements. The more of the gene region that can be deleted, the preferable, given this would provide more certainty of generating a null allele, while also maximising rescue potential. Also, the removal of the start codon would allow rescue with *HDAC4* cDNA, as well as genomic DNA fragments. However, these desirable characteristics must be balanced with the size of the original deletion, as experience suggests increasing deletion size reduces overall targeting efficiency (Poernbacher et al. 2019). Subsequently, because of the size of the *HDAC4* locus, I decided not to attempt to remove the entire locus. Instead, I designed a deletion of approximately 5.8kb, which would still delete approximately 90% of *HDAC4* coding sequence (**Fig. 4.1, 4.2A,B**). Importantly, all the characterised, functionally relevant domains for HDAC4 function would also be removed. This has a dual function. Firstly, it increased the likelihood of generating a null *HDAC4* allele; and secondly, increasing the versatility of the resulting allele by allowing the possible manipulation of multiple domains in the *HDAC4* locus. However, this design would clearly restrict rescue to gDNA-fragment based approaches, since rescue would rely upon 'fusion' between untargeted *HDAC4* gene sequence, and rescue DNA sequence recombined back into the attP landing site.

To target the *HDAC4* gene region and facilitate the generation of the aforementioned deletion, suitable CRISPR sites were identified. For the 5' CRISPR site, a site within intron 4 (of isoform D) was identified approximately 185bp away from the splice accepter site for that intron (**Fig. 4.2B,C**). This was chosen so that on rescue, the residual attP/B scar would hopefully not disrupt normal pre-mRNA splicing. The 3' CRISPR site chosen was just downstream of the endogenous stop codon (**Fig. 4B,D**). Before targeting, the integrity of these CRISPR sites were verified in the injection line by PCR, followed by DNA sequencing (**Fig. 4.2C-D**). This was because the *Drosophila* genome reference sequence does not represent the exact sequence for all fly lines, and even small variations between gRNA and target sequence would be sufficient to dramatically reduce targeting efficiency (Gratz et al. 2014; Ren et al. 2014). The homology arms were designed to be adjacent to the Cas9-induced DSB's, which occurs between base-pairs 17-18 of the protospacer sequence (Ran et al. 2013) (**Fig.**

4.2C-D). The homology arm positioning should define the boundaries of the attP-*Pax-cherry* cassette insertion in the resulting allele following homology-directed repair. Interestingly, for some unknown reason the *HDAC4* locus could not be amplified using Q5 2X master mix, which was the standard high fidelity polymerase used to amplify genomic DNA fragments for cloning purposes. Instead, an alternative high-fidelity enzyme mix, VeriFi (PCR biosystems) was used.

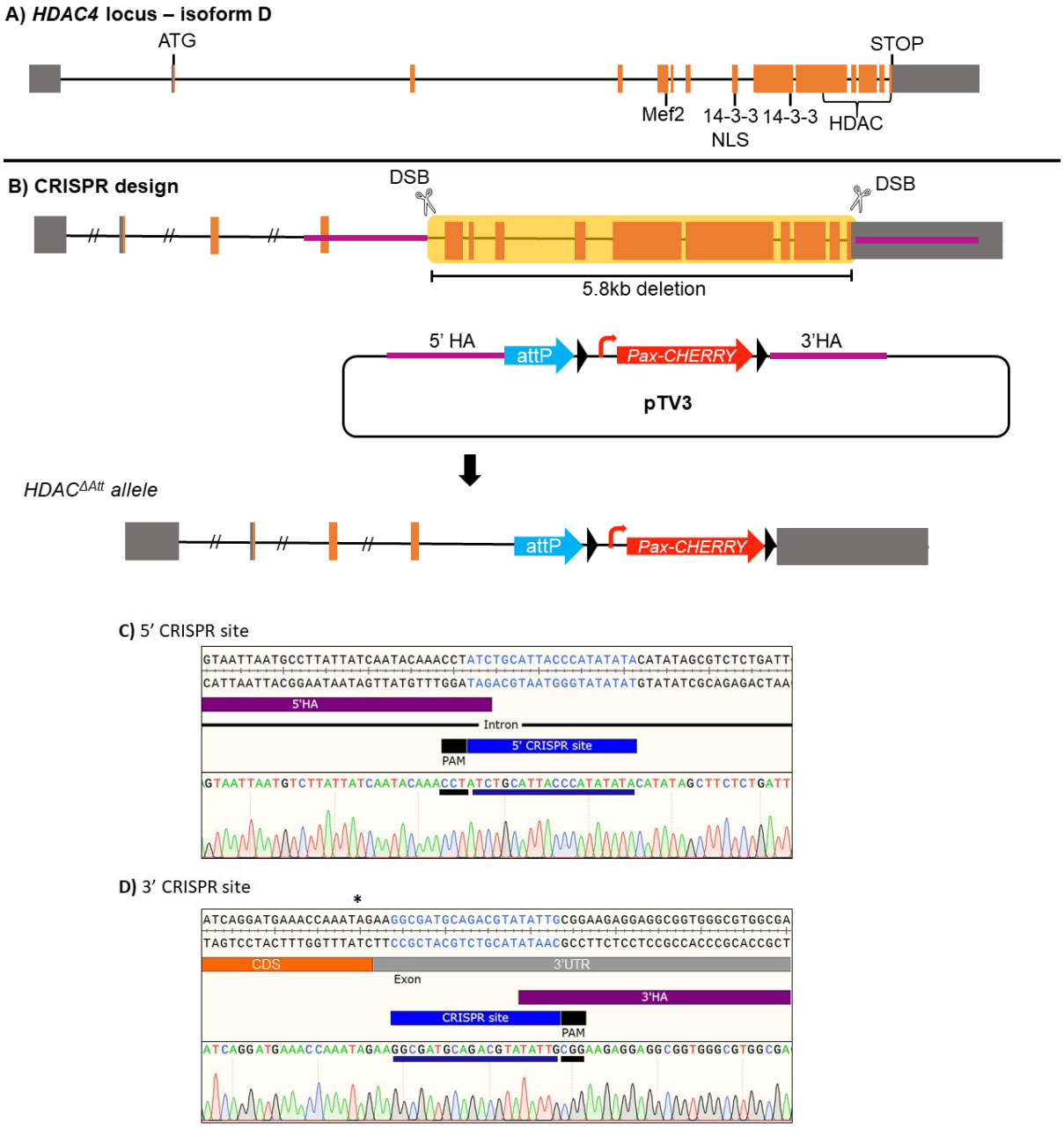


Figure 4. 2. Targeting the HDAC4 locus to generate an insertion-ready deletion allele. (**A**) Schematic of the *HDAC4* locus, isoform D. Relative positions of functionally characterised domains including the Mef2-binding site, nuclear localisation sequence (NLS), 14-3-3 protein

binding sites, and the HDAC domain are indicated. Non-coding and coding exons are in grey and orange, respectively. (**B**) Targeting design to generate the *HDAC4* insertion-ready deletion allele.The targeting vector, pTV3 contains an attP landing site followed by a *Pax-cherry* marker cassette flanked by LoxP sites (black). Homology arms cloned into pTV3 vector equate to those genomic regions either side of the double-strand break (DSB) points induced by Cas9. Deletion location and size is indicated by yellow shading. Structure of designed resulting alleles from these two targeting approaches are indicated.(**C-D**)) DNA sequence trances showing CRISPR site verification in injection lines for 5' (**C),** and 3' (**D**) CRISPR sites. Homology arm positioning (purple) relative to CRISPR sites are indicated. Stop codon is indicated by a (*).

4.2 *HDAC4* can be successfully engineered to generate a novel insertion-ready deletion allele, *HDAC4*$^{\Delta Att}$

Following injection and the generation of seven independent, balanced fly lines (by BestGene), the next aim was to verify whether the targeting of the *HDAC4* locus had been successful. To do so, flies were screened for the expression of the *pax-cherry* marker cassette. The *pax-cherry* cassette drives expression of the mCherry fluorophore in the developing larval CNS, and in the photoreceptors of the adult fly (Horn et al. 2002; Baena-Lopez et al. 2013). In *Oregon-R* control flies, there is no cherry expression in either the wandering L3 larva (**Fig. 4.3A**) or the adult fly eye (**Fig. 4.3B**). By contrast, flies harbouring the assumed targeted *HDAC4* allele display the characteristic cherry expression in both L3 wandering larvae (**Fig. 4.3C)**, and the adult fly eye (**Fig. 4.3D**). Interestingly, all lines were found to not homozygose and remained balanced over FM7. Firstly, this explains the irregular shape of the fly eye shown in **Figure 4.3D**, as this is the 'bar' eye phenotypic marker associated with the balancer chromosome; but also suggests *HDAC4* may be essential for viability.

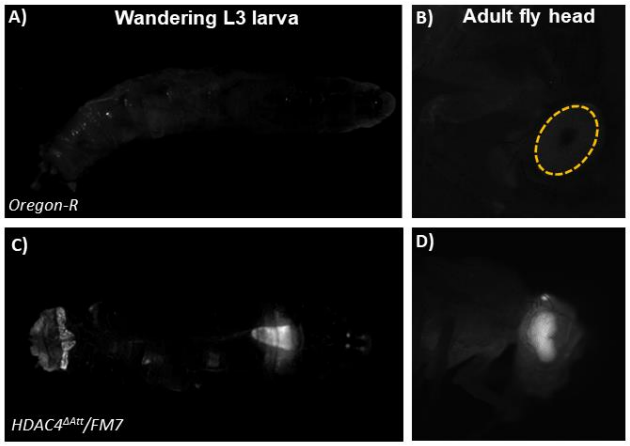


Figure 4. 3. *Pax-cherry* marker expression in *HDAC4*$^{\Delta att}$ engineered alleles
Expression of *Pax-cherry* was visualised in the developing larval CNS **(A,C)** and in the adult fly eye **(B,D).** *Oregon-R* control flies do not display any cherry expression **(A,B),** whereas flies harbouring the Mef2$^{\Delta CDS}$ (**C-D**) or Mef2$^{\Delta ATG}$ (**E-F**) allele display cherry visible cherry expression in both tissues.

Next, this novel allele was verified at the molecular level by PCR and DNA sequencing, to confirm that the *HDAC4* locus that had been targeted as designed. To do so, PCR primers that span the predicted engineered region, were used (**Fig. 4.4A**). Using these primers, the predicted PCR amplicon size would be approximately 2.2kb. By contrast, amplification of the wild-type *HDAC4* locus would yield a 6.8kb fragment. In theory, given that gDNA was isolated from heterozygous females, one could expect two PCR products from the reaction. However, the extension time was limited such that amplification of a 6.8kb fragment would not be possible. As shown in **Figure 4.4B,** PCR amplification yielded products of approximately 2.2kb, suggesting successful engineering of the *HDAC4* locus in all seven independent transgenic lines. These PCR products were purified and subsequently sequenced to map the engineered alleles. Sequencing showed the successful engineering of the *HDAC4* locus precisely as designed, with the insertion point of the attP-*Pax-cherry* cassette being defined by the original placement of the homology arms at both the 5' and 3' cleavage sites (**Fig. 4.4C-D**).

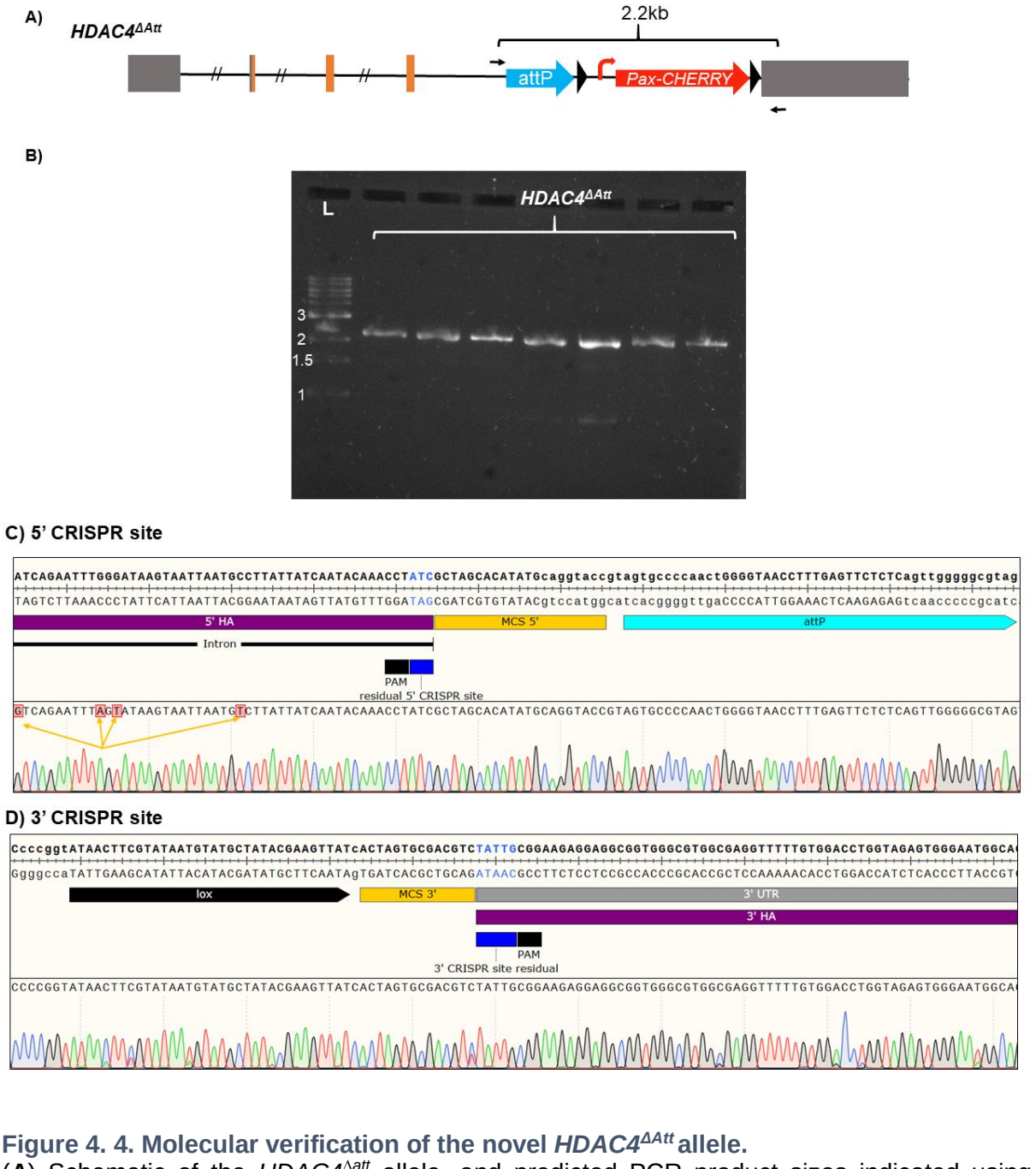


Figure 4. 4. Molecular verification of the novel *HDAC4$^{\Delta Att}$* allele.
(**A**) Schematic of the *HDAC4$^{\Delta att}$* allele, and predicted PCR product sizes indicated using primers either side of the engineered regions (black arrows). (**B**) Agarose gel shown PCR products generated from amplification (using primers described in (**A**) from gDNA extracted from heterozygous engineered flies. Seven independent lines harbouring the *HDAC4$^{\Delta att}$* allele were generated by Bestgene, and PCR amplification of each shows amplification of the predicted 2.2kb fragment. (**C-D**) DNA-sequence verification of purified PCR products (gel shown in **B**) in relation to positioning the of 5' CRISPR site **(C),** and the 3' CRISPR site (**D**) for the designed *HDAC4$^{\Delta att}$* allele. Top is the predicted DNA sequence following targeting, below is a DNA sequence chromatogram of the sequenced flies. This shows the boundaries of the insertion of the attP-*pax-cherry* cassette were accurately defined by the original positioning of the homology arms cloned into the pTV3 homology-directed repair template. SNPs between the reference sequence, and the sequenced line are shown with red boxes and orange arrows.

DNA sequencing of the *HDAC4*$^{\Delta att}$ highlights the sequence variation between the injection line, and the reference genome sequence. For example, multiple SNPs within the 5' homology arm region were identified (**Fig. 4.4C**). These SNPs were first validated to not be PCR-induced mutations by verifying their sequence through identifying their presence in the product from two independent PCR reactions (during the construction of the initial HDR vector), but also through the analysis of annotated DGRP (*Drosophila* genetic reference panel) variants on flybase (**Fig. 4.5**) (MacKay et al. 2012). The *HDAC4* gene region is highly variable, with many naturally occurring variants documented. These include both SNPs, as well as INDELs and are present in both coding and non-coding DNA (**Fig. 4.5A**). To further verify variants seen in homology arms, they were also compared and matched to naturally occurring variants annotated on flybase (**Fig. 4.5B**). This also emphasises the point already made about the importance of validating CRISPR sites in the injection line.

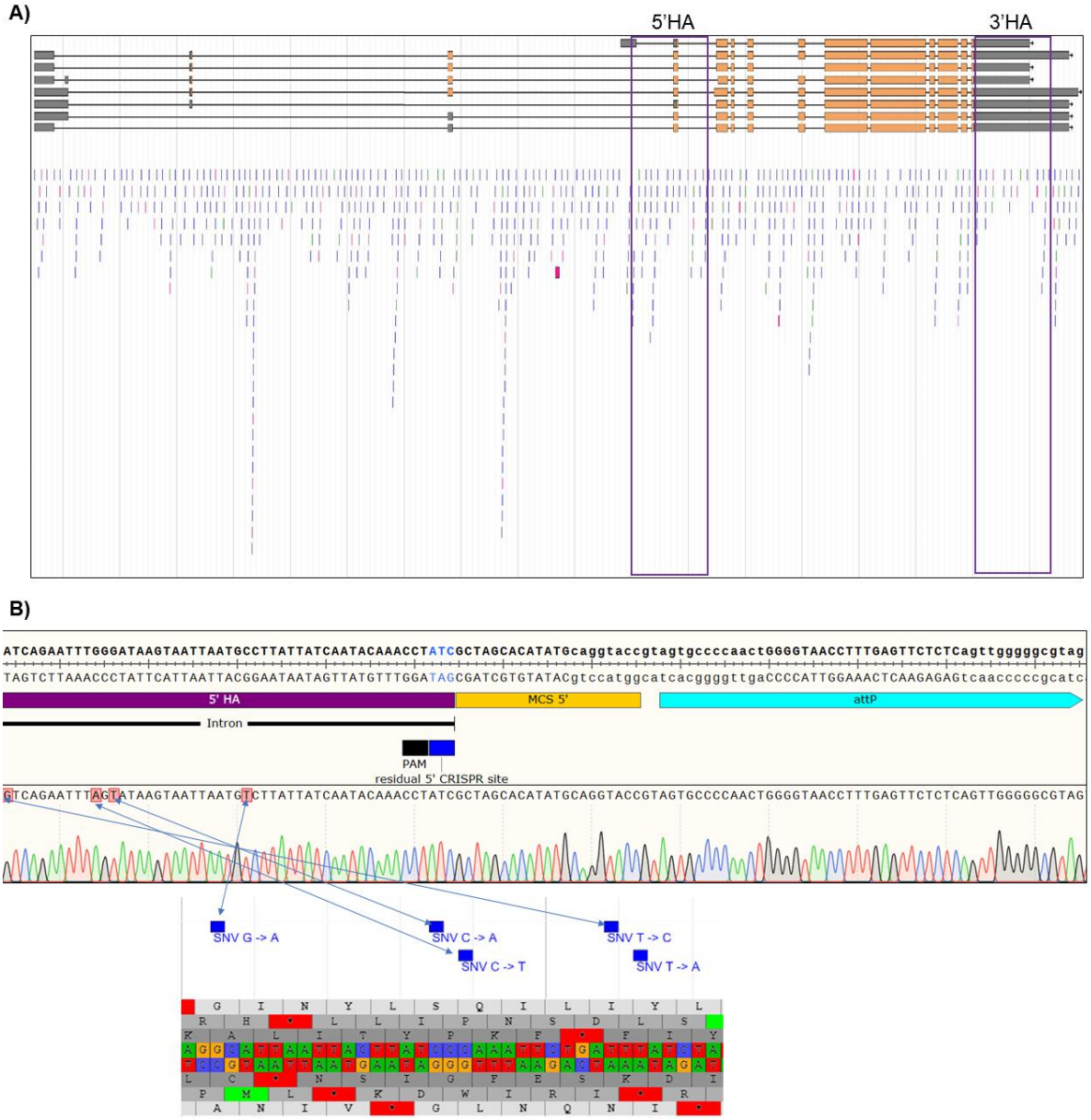


Figure 4. 5. The *HDAC4* locus is highly variable.
(**A**) Screenshot of the *HDAC4* locus on JBrowse, showing all 7 isoforms (top). (Below) track shown is the DGRP annotated variants. Each line indicates a naturally occurring variant. Blue, SNP; green, insertion, pink, deletion. Positioning of the homology arms used to generate the *HDAC4*$^{\Delta att}$ are indicated by purple boxes. (**B**) (top) schematic of the DNA sequence of the verified *HDAC4*$^{\Delta att}$ allele, nucleotides which differ from the reference sequence are indicated. (bottom) all of these SNPs map to naturally occurring annotated DGRP variants. Note these are oriented in opposite directions (bottom =reverse complement).

4.3 Design and verification of a genomic DNA fragment to rescue *HDAC4*$^{\Delta Att}$

Given the successful targeting of *HDAC4* to generate *HDAC4*$^{\Delta Att}$, I next designed a rescue construct to reinsert *HDAC4* genomic DNA deleted during the initial targeting,

with the aim of restoring gene function. A schematic for this design is shown in **Figure 4.6A**. A gDNA fragment was cloned into the RIV^{White} vector, which was subsequently inserted into the *HDAC4*$^{\Delta Att}$ allele through site-specific recombination between the attP and attB sites. The 5' end of the gDNA rescue construct began with the first nucleotides that were deleted during initial targeting (i.e the first nucleotide after the end of the 5' homology arm), such that all DNA removed during the generation of *HDAC4*$^{\Delta att}$ would be restored (**Fig. 4.6B**). At the 3' end, I included the entire annotated 3' UTR sequence for *HDAC4* (**Fig. 4.6C**). This was because, although it would result in partial duplication of the *HDAC4* 3'UTR (as much of it had not been deleted during *HDAC4*$^{\Delta att}$ generation), this design would hopefully mean that transcription of the rescued *HDAC4* allele would not be interrupted by the *Pax-cherry* cassette, which was not removed prior to rescue. Given the molecular verification of the *HDAC4*$^{\Delta att}$, only a single line was used for the rescue.

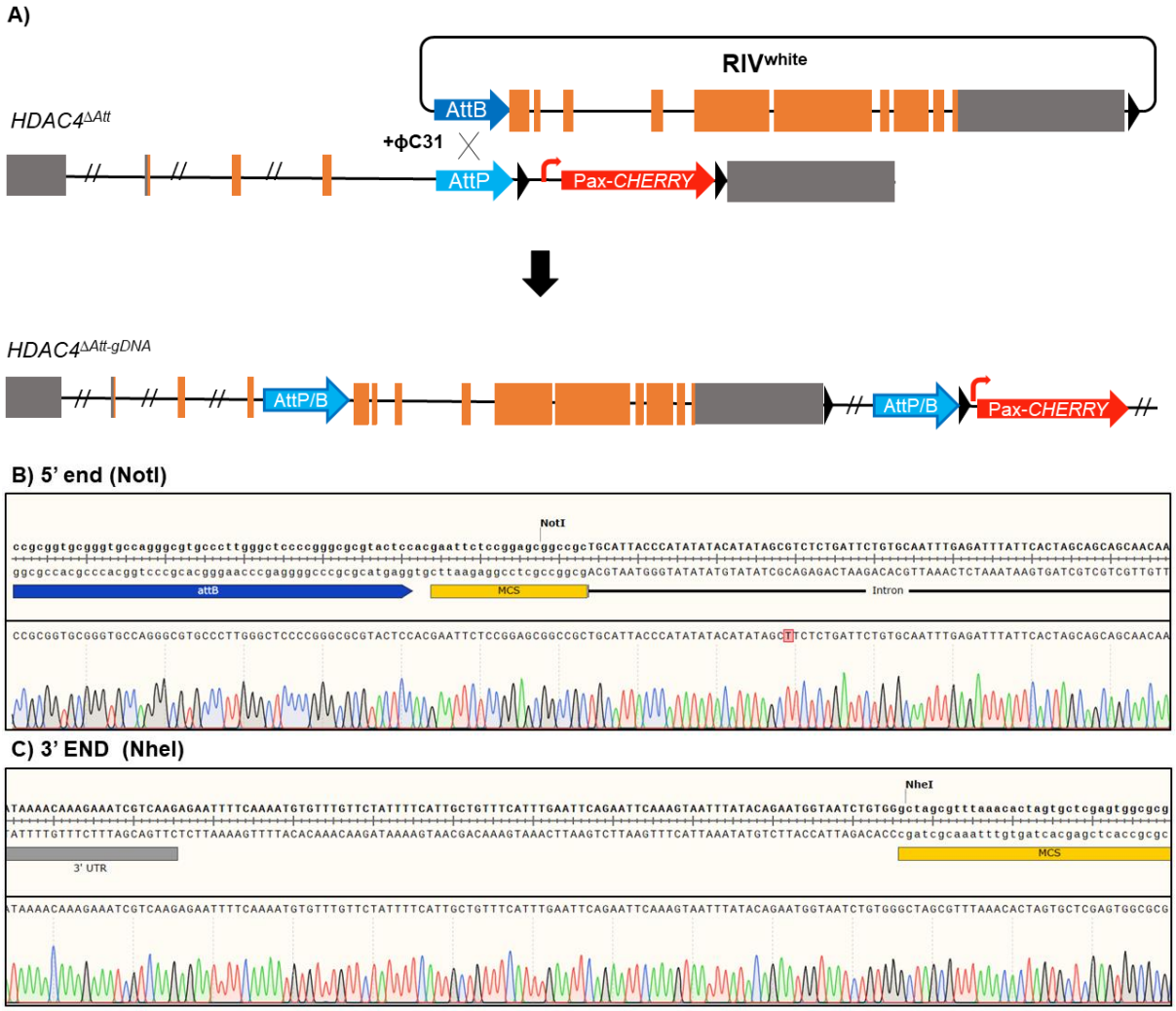


Figure 4. 6. Design of gDNA-based rescue construct to restore function of *HDAC4$^{\Delta Att}$*

(**A**) Schematic of a gDNA rescue to restore function of the *HDAC4$^{\Delta Att}$* allele. (**B-C**) DNA sequence map, and DNA sequence chromatogram of 5' and 3' ends respectively of rescue construct inserted into RIV white. NotI and NheI enzymes are indicated as those used to clone in the rescue fragment. (**B**) schematic (top) and DNA sequence chromatogram (bottom) of 5' end of rescue construct, showing the first bases of the rescue fragment following the NotI site ("TGCATT") is the boundary of what was initially deleted during gene targeting. (**C**) Schematic (top) and DNA sequence chromatogram (bottom) of 3' end of rescue construct. As with the 5' boundary, the amount of gDNA cloned into the rescue construct was defined by the boundary of the initially deletion.

Flies harbouring the rescued *HDAC4$^{\Delta att}$* allele to generate *HDAC4$^{\Delta att\text{-}gDNA}$* were selected for by the presence red-eyed flies. This was due to the *white* transgene being present in the RIV^{white} vector. Unlike in the initial generation of *HDAC4$^{\Delta att}$*, where seven independent transgenic lines were received, only one line for *HDAC4$^{\Delta att\text{-}gDNA}$* was generated. This may be because these lines were generated by the co-injection of mRNA encoding phi-C31 integrase rather than through stable germline expression.

This technique is less efficient than when transgenics are generated using germline expression of phi-C31 (Bischof et al. 2007). Nevertheless, by contrast to *HDAC4*$^{\Delta att}$, which as previously mentioned does not homozygose, the single *HDAC4*$^{\Delta att\text{-}gDNA}$ line was found to be homozygous viable and fertile. Given the mapping of the initial *HDAC4*$^{\Delta att}$ allele, this indicated that restoration of the *HDAC4* locus through integration of the gDNA fragment is sufficient to restore viability. Nevertheless, to also verify the insertion of the rescue DNA fragment molecularly into the landing site within *HDAC4*$^{\Delta att}$, gDNA was extracted from the line, followed by PCR amplification of the locus using three different primer pairs which, should the desired insertion be present, yield PCR products of 2.2kb (1), 3.3kb (2), and 3.8kb (3) (**Fig. 4.7A**). PCR amplification followed by agarose gel electrophoresis produced products of the expected sizes for each of the primer pairs used, therefore confirming integration of the rescue construct into the desired genomic location, to generate the *HDAC4*$^{\Delta Att\text{-}gDNA}$ allele (**Fig. 4.7B**).

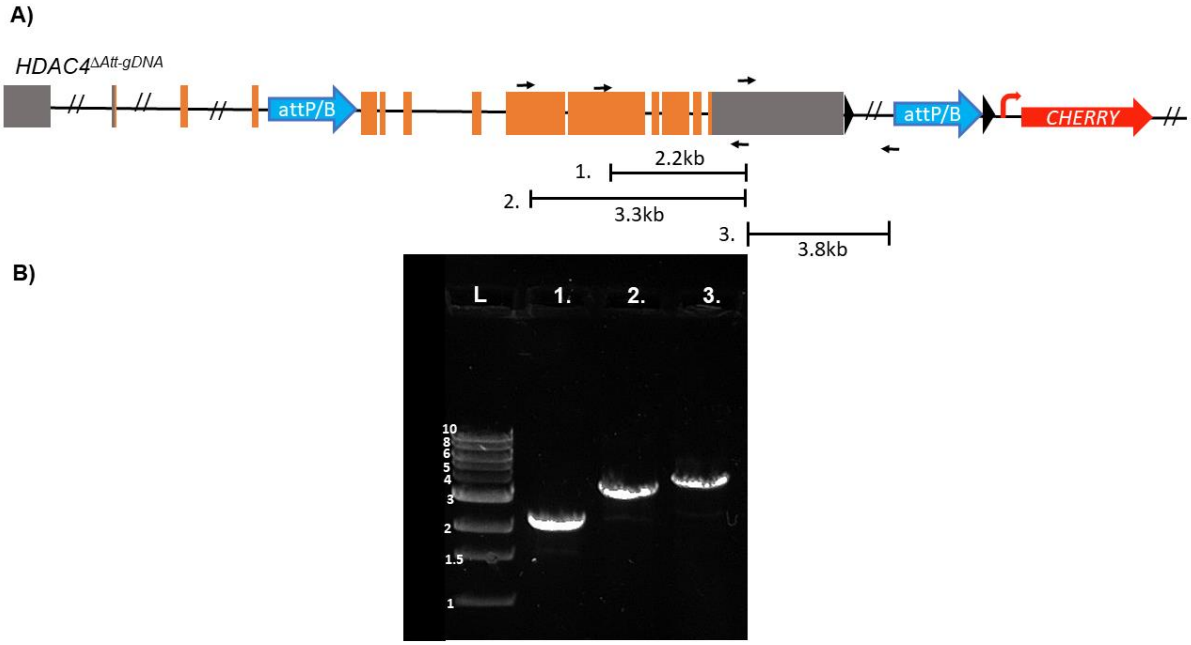


Figure 4. 7. Molecular verification of *HDAC4$^{\Delta Att}$*

(A) Schematic showing organisation of the *HDAC4$^{\Delta Att\text{-}gDNA}$* locus. Non-coding and coding exons are in grey and orange, respectively. Arrowheads represent positions of different primers, and subsequent predicted PCR product sizes following PCR amplification from the *HDAC4$^{\Delta Att\text{-}gDNA}$* locus using 3 different primer pairs (1.,2.,3.). (**B**) Agarose gel showing PCR products generated following amplification of *HDAC4$^{\Delta Att\text{-}gDNA}$* using primers described in **A**.

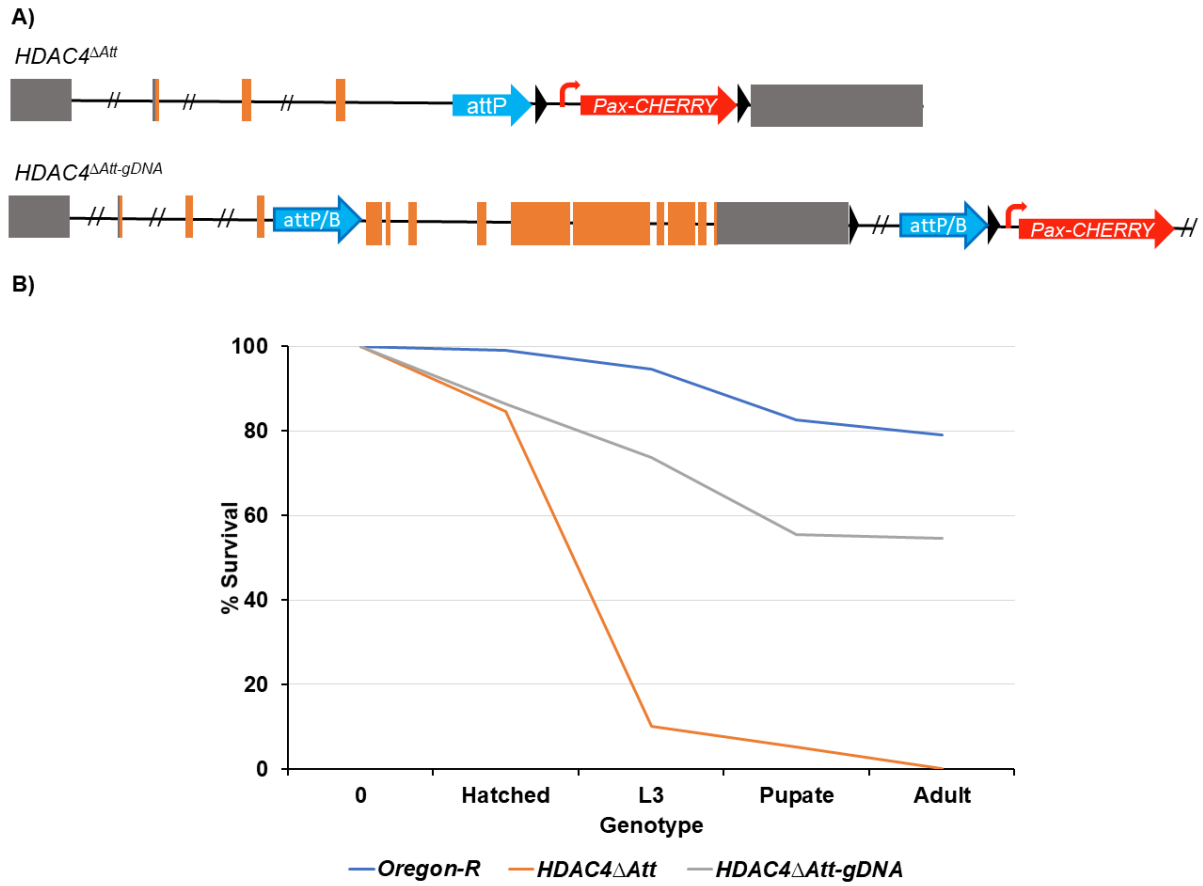
4.4 *HDAC4* loss-of-function mutants display progressive lethality and growth defects.

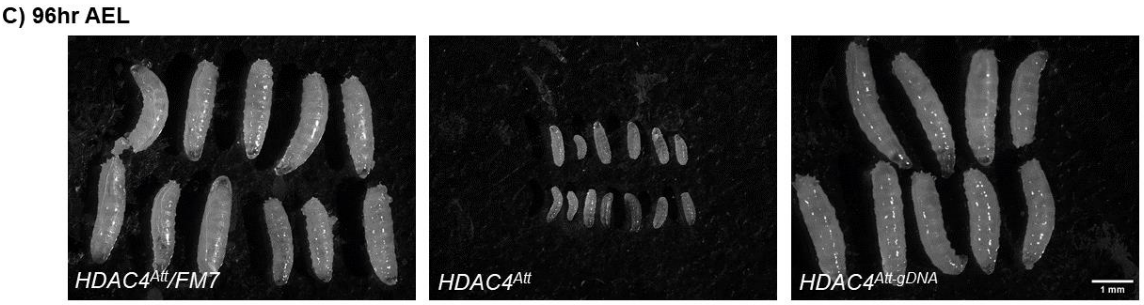
Given that *HDAC4$^{\Delta att}$* was not homozygous viable, whereas *HDAC4$^{\Delta att\text{-}gDNA}$* was, I wanted to define the stage of lethality associated with *HDAC4* loss-of-function. To investigate this, a hatching and survival assay was undertaken, using *Oregon-R* as a wild-type control. Also included was *HDAC4$^{\Delta att\text{-}gDNA}$*, to establish whether the rescue stock's viability was restored to that of wild-type. In this assay, survival was measured at four time points following embryo collection: immediately following hatching, L3 larval stage, white pupa, and eclosed adults. Survival at each stage was presented as a percentage of the original embryo population collected (**Fig. 4.8**). Wild-type *Oregon-R* were found to be relatively healthy: 99% of embryos hatched into L1 larvae; 95% reached L3 stage; 83% pupation; and 79% reached adulthood. By contrast, *HDAC4$^{\Delta att}$* mutants displayed progressive lethality throughout *Drosophila* development, with the majority of individuals dying during larval life. Approximately 85% of fertilised embryos

hatched into L1 larvae, however the majority of these larvae die during larval stages and fail to reach metamorphosis. only 10% of the original fertilised embryos reached the L3 larval stage, and only 5% pupated. None of the flies that did undergo pupation eclosed, although some (not quantified here) did reach the dark pupal stage determined by visible wing darkening (Bainbridge and Bownes 1981). *HDAC4*$^{\Delta att\text{-}gDNA}$, as previously mentioned is homozygous viable, and thus the *HDAC4* loss-of-function induced lethality is rescued in this allele. However, *HDAC4*$^{\Delta att\text{-}gDNA}$ was found to be less healthy than *Oregon-R* controls. 86% of embryos hatched, 74% made L3, 55% pupated, and 55% made it to adulthood. *HDAC4*$^{\Delta att\text{-}gDNA}$ could be less healthy than *Oregon-R* controls for a variety of reasons. Firstly, it could be genetic background, and whether the injection stock itself is less healthy than wild-type controls. Or, secondly it could incomplete rescue of the *HDAC4* locus, possibly caused by the fact of this not a scarless technique.

Interestingly, a growth defect was also observed in *HDAC4*$^{\Delta att}$ mutants compared to heterozygous controls. At four days AEL, where development should have progressed to the mid L3 stage (Hales et al. 2015), surviving *HDAC4*$^{\Delta att}$ mutants were visibly smaller than *HDAC4*$^{\Delta Att}$ heterozygous controls (**Fig. 4.8C**). Interestingly however, these larvae have still developed to L3, determined by mouthpart morphology, and the darkening, orange colour of the posterior spiracles (Koyama and Mirth 2021). By contrast, this phenotype appears to be rescued in *HDAC4*$^{\Delta att\text{-}gDNA}$, with larvae at the same stage resembling that of heterozygous controls (**Fig. 4.8C**). Furthermore, I also visualised surviving wandering *HDAC4*$^{\Delta att}$ L3 larvae and found these were also visually smaller than both the heterozygous control, and the *HDAC4*$^{\Delta att\text{-}gDNA}$ rescue (**Fig. 4.8D**). Furthermore, male *HDAC4*$^{\Delta att}$ mutants which did make it to the dark pupal stage were also significantly smaller than both male controls, and *HDAC4*$^{\Delta att\text{-}gDNA}$ rescue pupae (**Fig. 4.8E-F**). Because pupal measurements are easier due to their immobilised state, compared to larvae which constrict during larval muscle contractions, pupal size was measured. *FM7/Y* were also smaller than *HDAC4*$^{\Delta att\text{-}gDNA}$ rescue male pupae, although not significantly. However, this likely be due to the small sample size for *HDAC4*$^{\Delta att\text{-}gDNA}$ (n=3). This size difference could be due to an FM7 balancer phenotype, because normal pupal length ranges from 2.8-3.9mm (Reeves and Tautz 2017). These data

suggest *HDAC4* is essential for both normal growth during development, as well as viability.





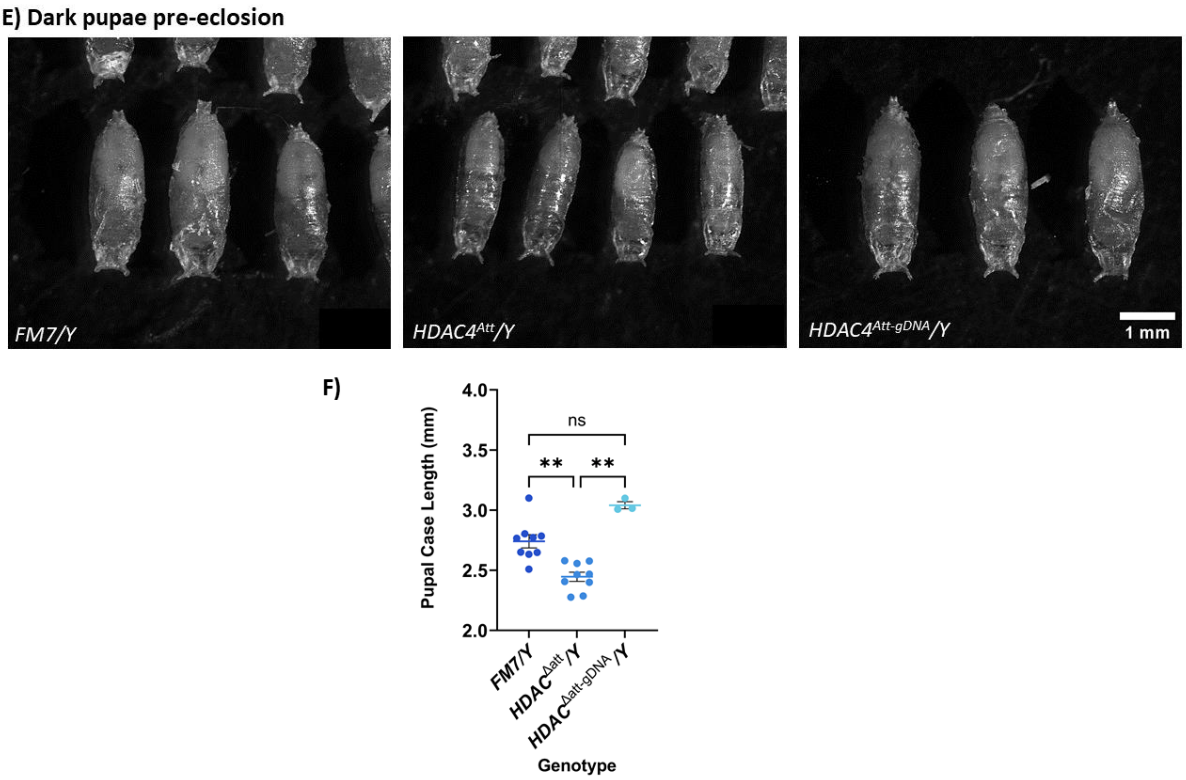


Figure 4. 8. *HDAC4$^{\Delta Att}$* mutants are not viable and display growth defects.
(**A**) schematics of the *HDAC4$^{\Delta Att}$* and *HDAC4$^{\Delta Att\text{-}gDNA}$* loci. (**B**) Hatching and survival data for *Oregon-R* controls, *HDAC4$^{\Delta Att}$*, and *HDAC4$^{\Delta Att\text{-}gDNA}$*. Survival was scored as the percentage of an initial 100 fertilised embryos selected which were alive at each of four developmental stages. (**C-D**) Larvae of each of the three genotypes either 4 days after egg laying (AEL) (**C**) or at the wandering L3 stage (**D**). (**E-F**) Male dark pupae of either *FM7/Y* controls, *HDAC4$^{\Delta Att}$* and *HDAC4$^{\Delta Att\text{-}gDNA}$* (**E**) and measurements of pupal case length of each genotype (**F**). ** = P<0.01 determined from Kruskal-Wallis test followed by Dunn's multiple comparison test. Each coloured circle represents length of individual pupa. Mean and S.E.M are indicated.

4.5 Somatic musculature appears unaffected in *HDAC4$^{\Delta Att}$* mutant embryos

Given the emphasis of this project being on HDAC4 function during muscle development, I next aimed to investigate the effect of *HDAC4* loss-of-function on muscle formation. To do so, the embryonic somatic musculature was first analysed in late-stage *HDAC4$^{\Delta Att}$* mutant embryos, using an antibody against Mhc (**Fig. 4.9**). Interestingly, the musculature of *HDAC4$^{\Delta Att}$* mutant embryos appeared unaffected, with the regular muscle pattern forming normally, and no obvious abnormalities being detected when compared to control embryos (**Fig. 4.9**).

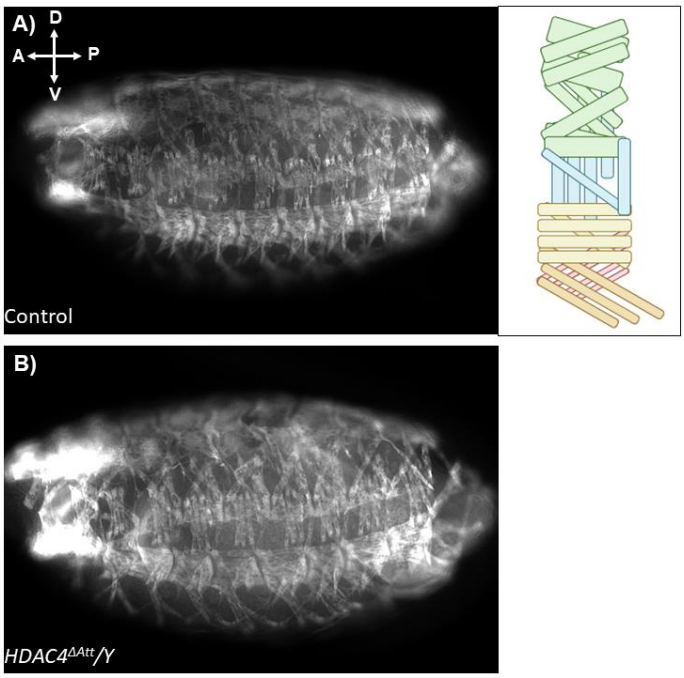


Figure 4. 9. Gross muscle architecture in *HDAC4*$^{\Delta Att}$ mutants appears normal
Late-stage control (**A**), or *HDAC4*$^{\Delta Att}$ mutant (**B**) embryos stained with an antibody against myosin heavy chain (Mhc) to visualise the somatic muscle pattern. No obvious muscle phenotype can be observed in *HDAC4*$^{\Delta Att}$ mutants compared to controls.

4.6 Investigating the effect of *HDAC4* loss-of-function on muscle differentiation: using an Mhc-GFP transgene for live imaging of the larval somatic musculature

Given the hatching and survival data for *HDAC4*$^{\Delta att}$, where mutants primarily die during larval life (**Fig. 4.**), it is possible that the somatic musculature forms normally during embryogenesis, however then deteriorates. Furthermore, imaging of embryos by compound microscopy rather than confocal may not provide the resolution necessary to detect subtle muscle phenotypes. To image the somatic musculature of *HDAC4*$^{\Delta att}$ mutant larvae, I adopted a technique developed by Balakrishnan et al. (2021), whereby expression of a GFP-tagged protein expressed in the somatic musculature can be used to visualise the structure of the larval somatic muscles. This technique would, in theory, have benefits over fixation and staining protocols to visualise muscle phenotypes in larvae. This is firstly due to the ability to keep samples in-tact, which could enable the imaging of a single larva throughout larval life. Secondly, keeping samples in-tact prevents damaging the tissue and thus reduces the risk of incorrectly identifying muscle phenotypes induced through sample preparation. Finally, it could

also streamline the technique, as the method is significantly less time consuming than antibody staining protocol.

For this approach, I used the Mhc-GFP stock from the ftrg library, which I have used throughout the project to detect premature differentiation of L3 AMPs upon Mef2 overexpression (**Chapter 3**). This line had not been used previously to analyse larval muscle structure, and is in contrast to the published technique whereby tropomyosin-GFP, Zasp66-GFP, or α-actinin-GFP were used (Balakrishnan et al. 2021).

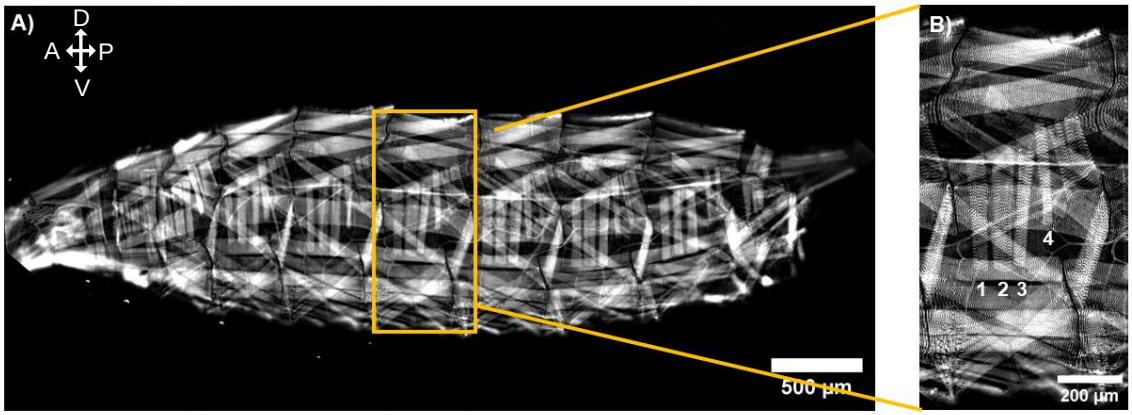


Figure 4. 10. Live imaging of larval somatic muscles using Mhc-GFP
(A) Live wandering L3 larva expressing Mhc-GFP (myosin heavy chain) transgene, which is expressed in the larval somatic muscles. (**B**) magnified view showing larval muscle structure and organisation in one abdominal segment. Numbered are the four lateral-transverse (LT) muscles.

I first used this stock to see whether the somatic musculature could be visualised in wandering L3 larvae. Indeed, the somatic muscle pattern could be easily visualised using this transgene (**Fig. 4.10A**). In particular, I noted the ability to detect Mhc striations within the somatic muscles, possibly suggesting this live imaging technique may also be used to analyse muscle ultrastructure (**Fig. 4.10B**). While the overall muscle pattern was clearly identifiable with this technique, scoring of individual muscles was found to be challenging due to the overlapping nature of the somatic muscles, especially those in the dorsal and ventral region of each hemisegment (**Fig. 4.10**). However, the lateral transverse (LT) muscles were clearly visible and distinguishable, which made them ideal for the identification of subtle muscle phenotypes associated with HDAC4 loss-of-function.

To investigate whether HDAC4 loss of function disrupts the larval somatic musculature, I generated a stock with both the *HDAC4$^{\Delta att}$* allele, and the *Mhc-GFP* transgene. Given the requirement to be able to select for *HDAC4$^{\Delta Att}$* mutants, male L3 larvae were selected based on the enlarged gonads present only in males (Chyb and Gompel 2013), and the presence of the *HDAC4$^{\Delta att}$* allele was selected by screening for *pax-cherry* expression. Although these selection criteria were initially used, I realised the *Mhc-GFP* transgene also possesses a *pax-cherry* cassette. However, to first investigate whether there were any distinguishable differences in the *cherry* expression pattern between the two insertions, I first analysed larvae from both the *Mhc-GFP* stock alone, and the stock containing both *Mhc-GFP* and *HDAC4$^{\Delta att}$* (**Fig. 4.11**). Interestingly, while characteristic cherry expression was visualised in the developing CNS of both lines, the *HDAC4$^{\Delta att}$* stock displayed additional cherry expression in a structure that appeared to be the larval gut (**Fig. 4.11B**). This was not detected in the *Mhc-GFP* stock alone (**Fig. 4.11A).** Therefore, this differential cherry expression was used during larval selection to select *HDAC4$^{\Delta att}$* mutants.

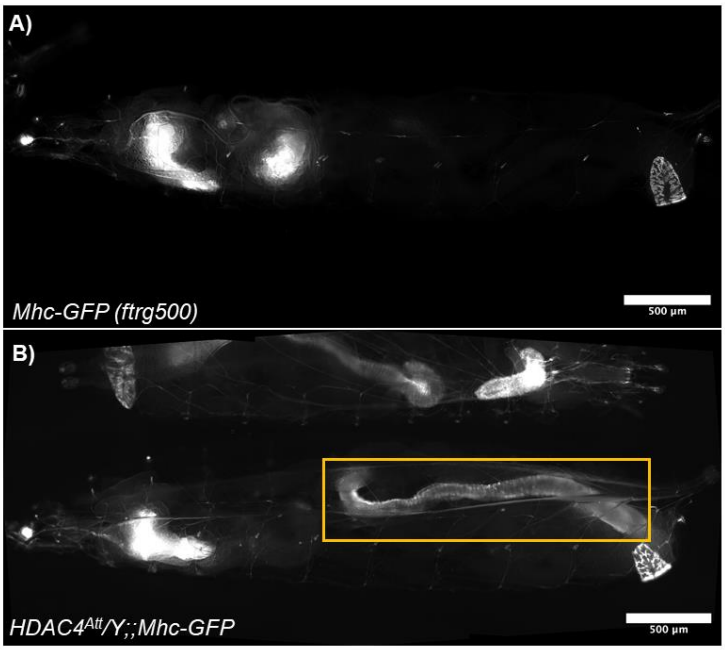


Figure 4. 11. Differential Pax-cherry expression in transgenic larvae
Pax-cherry expression in wandering L3 larva of either the *Mhc-GFP* (**A**) or *HDAC4$^{\Delta Att}$;;Mhc-GFP stocks* (**B**). Cherry expression is detected in the larval gut in *HDAC4$^{\Delta Att}$;;Mhc-GFP* (orange box), which is not detected in the *Mhc-GFP* stock alone.

4.7 *HDAC4*$^{\Delta att}$ mutants have a larval muscle phenotype

Following the development of this approach to select for *HDAC4*$^{\Delta att}$ mutants, I mounted surviving L3 larvae and visualised the larval muscles to assess their structure compared to controls. I initially observed visible patterning defects in the number of LT muscles, which primarily affected LT1-3. Therefore, because of time constraints, this analysis was limited to the patterning of LT1-3. In some hemisegments, mutant larvae displayed LT muscle duplications, while others were lacking LT muscles (**Fig. 4. 12A**). However, although these phenotypes were present, most hemisegments within each larva were unaffected, suggesting only a subtle effect on somatic muscle patterning is caused by *HDAC4* LOF. Nevertheless, to quantify this phenotype, I first counted the proportion of hemi-segments that displayed WT LT patterning, and compared this to *Mhc-GFP* only control larvae. To do so, for each larva I counted hemi-segments A2-A7, which each display the same somatic muscle pattern. On average, 94% of hemisegments within each control larva displayed WT LT muscle number. By contrast, in *HDAC4*$^{\Delta att}$ mutants, this was significantly reduced to 70% (**Fig. 4.12C,D**).

To see whether this phenotype was rescuable in the *HDAC4*$^{\Delta att\text{-}gDNA}$, homozygous virgin females were crossed to males of the *Mhc-GFP* stock, and the progeny's larval somatic musculature was analysed. Indeed, in this genotype, the larval somatic musculature was restored to that observed in control larvae, with on average 95% of hemisegments displaying WT LT number **(Fig. 4.12C,D**). Together, these data suggest *HDAC4* LOF causes a small, albeit measurable effect on the LT muscles of the larval somatic musculature. Moreover, importantly this phenotype is also rescued in *HDAC4*$^{\Delta att\text{-}gDNA}$, emphasising the potential power of this allele to probe HDAC4 function *in vivo*.

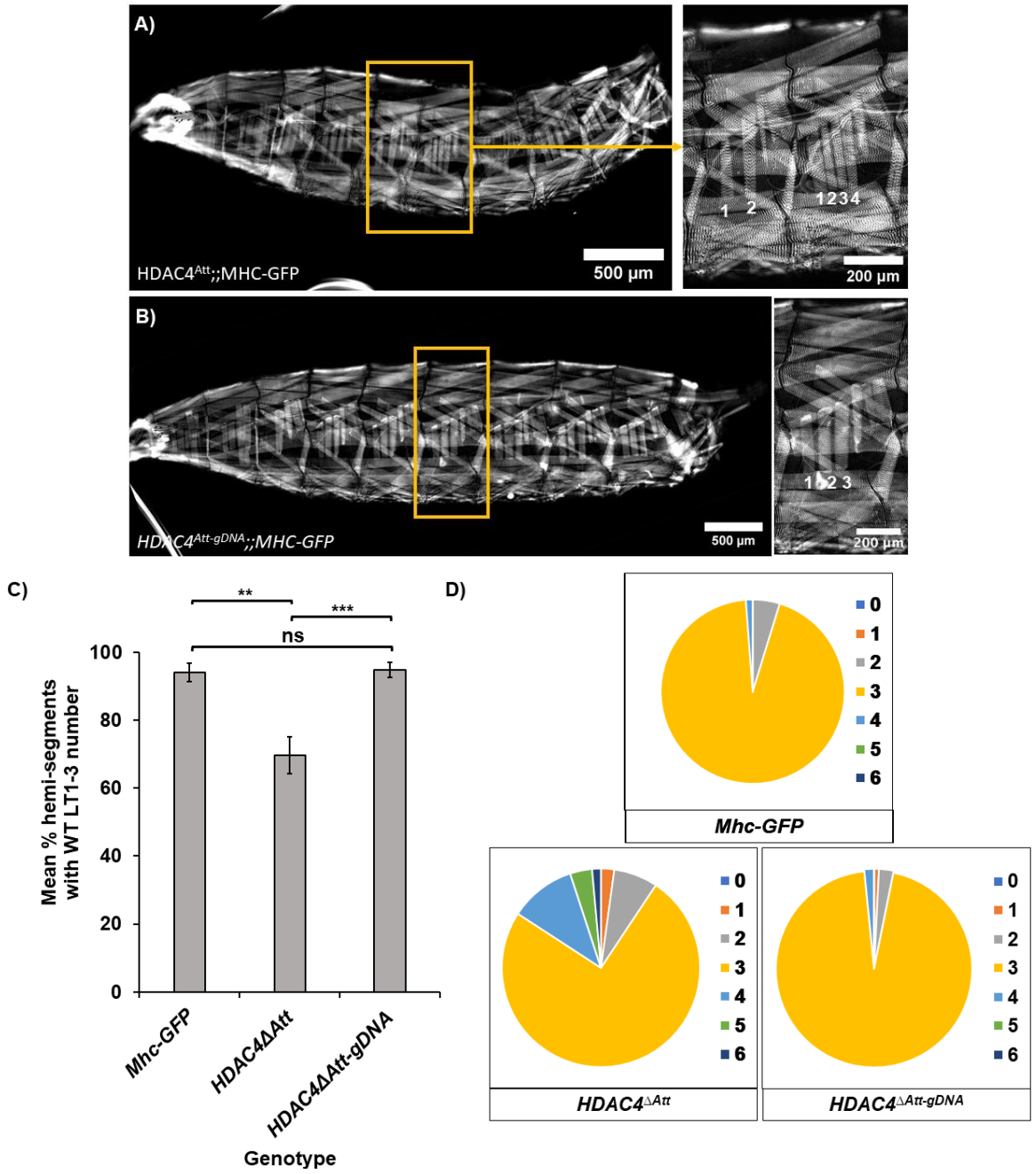


Figure 4. 12. *HDAC4*$^{\Delta att}$ mutants have a larval muscle phenotype

(**A-B**) *HDAC4*$^{\Delta att}$ LOF mutants (**A**) or *HDAC4*$^{\Delta att\text{-}gDNA}$ rescue (**B**) L3 larva expressing Mhc-GFP transgene to visualise larval somatic musculature. (**A**) *HDAC4*$^{\Delta att}$ LOF mutants display patterning defects in the lateral transverse 1-3 (LT1-3) muscles, which include LT muscle loss, or additional muscles. (**B**) LT muscle patterning is restored to wild-type in *HDAC4*$^{\Delta att\text{-}gDNA}$. (**C**) phenotype severity scored by mean percentage of hemisegments with WT LT muscle patterning. Error bars represent S.E.M. *** = p<0.001, ** = p<0.01, calculated from Kruskal-wallis test followed by Dunn's multiple comparisons test. (**D**) Pie chart distribution of number of LT muscle fibres (LT 1-3 scored) per hemithorax in each genotype.

4.8 *Mef2* overexpression may cause a larval muscle phenotype

Because of the overarching interest in HDAC4 function in the context of Mef2 regulation *in vivo*, it could be that this *HDAC4* LOF phenotype may phenocopy the effect of overexpressing Mef2. Thus, I aimed to overexpress Mef2 in the larval somatic musculature, to see whether this also disrupted LT muscle patterning. To do so, I used a Mef2-Gal4 line which also possess the *Mhc-GFP* transgene, to overexpress Mef2 in the developing larval somatic musculature. Mef2-Gal4 also expresses in the myoblasts of the developing somatic musculature during embryogenesis, and *Mef2* overexpression has also been shown to disrupt normal patterning of the embryonic somatic muscles (Gunthorpe et al. 1999). It was unknown whether this would cause premature lethality and prevent the analysis of L3 larvae with this genotype. However, I found that this genotype was not lethal, and larvae do survive to L3 (although relative survival was not quantified). Interestingly, *Mef2* overexpression also disrupted the normal patterning of the LT muscles, similar to the phenotype observed in $HDAC4^{\Delta att}$ mutants (**Fig. 4.13A**). On average, only 80% of hemisegments within a larva possessed WT LT1-3 muscle patterning, which was less severe than the 70% observed in $HDAC4^{\Delta att}$ mutant larvae, however was still much lower than the 94% average observed in *Mhc-GFP* only controls (**Fig. 4.13B,C**). Interestingly however, this phenotype was not found to be statistically significant ($p=0.12$), possibly indicating that sample sizes may not have been large enough to derive full statistical confidence from these results.

Nevertheless, these data show that both Mef2 overexpression, and *HDAC4* LOF are associated with larval muscle patterning defects. However, this does not also rule out the possibility of other muscle phenotypes for which there was no time to investigate.

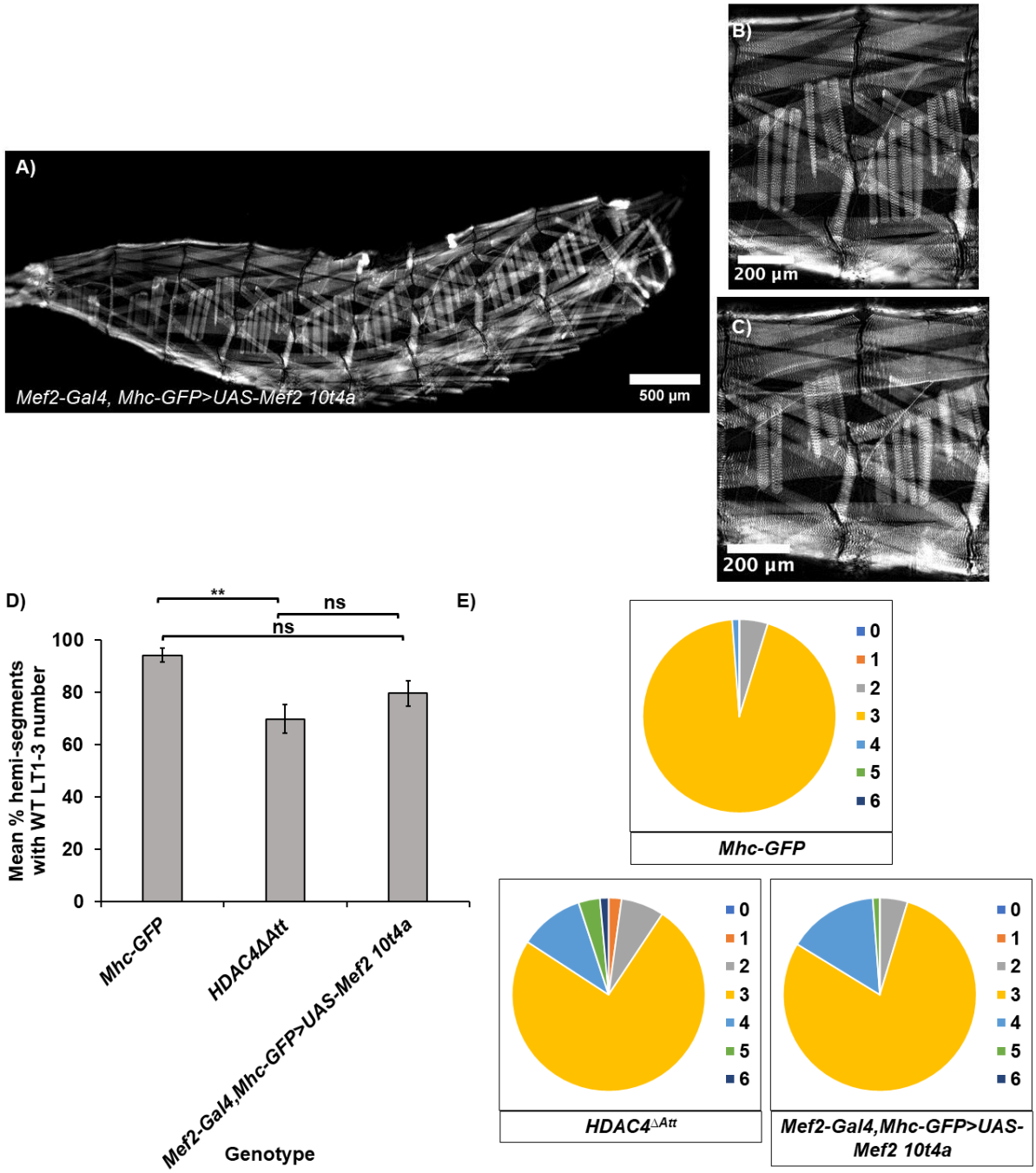


Figure 4. 13. *Mef2* overexpression may cause a larval muscle phenotype

(**A**) Live L3 larva overexpressing Mef2 under the control of Mef2-Gal4, and also harbouring an Mhc-GFP transgene (under endogenous regulation) to visualise larval somatic muscle pattern. (**B-C**) representative hemisegments from larvae overexpressing Mef2, highlighting the defects in patterning of the lateral-transverse 1-3 (LT1-3) muscles. Mef2 overexpression causes both the presence of additional (**B**) and missing (**C**) LT muscles. (**D**) phenotype severity scored by mean percentage of hemisegments with WT LT muscle patterning. Error bars represent S.E.M. Significance ** represents $p<0.01$ from Kruskal-wallis test followed by Dunn's multiple comparisons test. (**E**) Pie chart distribution of number of LT muscle fibres (LT 1-3 scored) per hemithorax in each genotype.

4.9 *HDAC4*$^{\Delta att}$ mutants have a DLM muscle phenotype

I next asked whether, in *HDAC4*$^{\Delta att}$ mutants that do survive to late pupal stage, there is also an adult muscle phenotype associated with *HDAC4* LOF. To do so, the DLM fibres were investigated, given that *HDAC4* overexpression has already been shown to inhibit their formation (**Chapter 3**). For this experiment, transverse thoracic cross-sections were stained with phalloidin, a toxin which binds to F-actin. Similar to hematoxylin, this technique can also be used to stain and visualise DLM fibre morphology in the adult thorax. Interestingly, in *HDAC4*$^{\Delta att}$ mutants there appeared to be a DLM muscle phenotype (**Fig. 4.14**): In control samples, the DLMs show a highly organised, symmetrical pattern previously described (**Fig. 4.14A**). By contrast, *HDAC4*$^{\Delta att}$ mutants did not show WT DLM fibre morphology, however unlike with *HDAC4* overexpression, I did not observe a general inhibition to the development of the fibres. While in some cases, the DLM fibres do appear to form normally, with 6 pairs of symmetrical fibres can be seen either side of the thoracic midline (**Fig. 4.14B**), in others, although the DLM fibres are present within the adult thorax, their organisation appears to be disrupted: In some cases, fibres can be observed in the incorrect position (**Fig. 4.14C**); in others, some fibres appear misshapen and irregularly sized relative to other fibres within the thorax (**Fig. 4.14D**); and in more severely affected samples, there appears to be a more global effect on DLM organisation, with irregularly shaped fibres being present throughout the thorax (**Fig. 4.14E**). I then compared this phenotype to *HDAC4*$^{\Delta att\text{-}gDNA}$, and found this a rescue of this phenotype and restoration to WT DLM morphology, with 6 pairs of symmetrical, regularly sized DLM fibres being correctly positioned throughout the thorax of the adult fly (**Fig. 4.14F-I**).

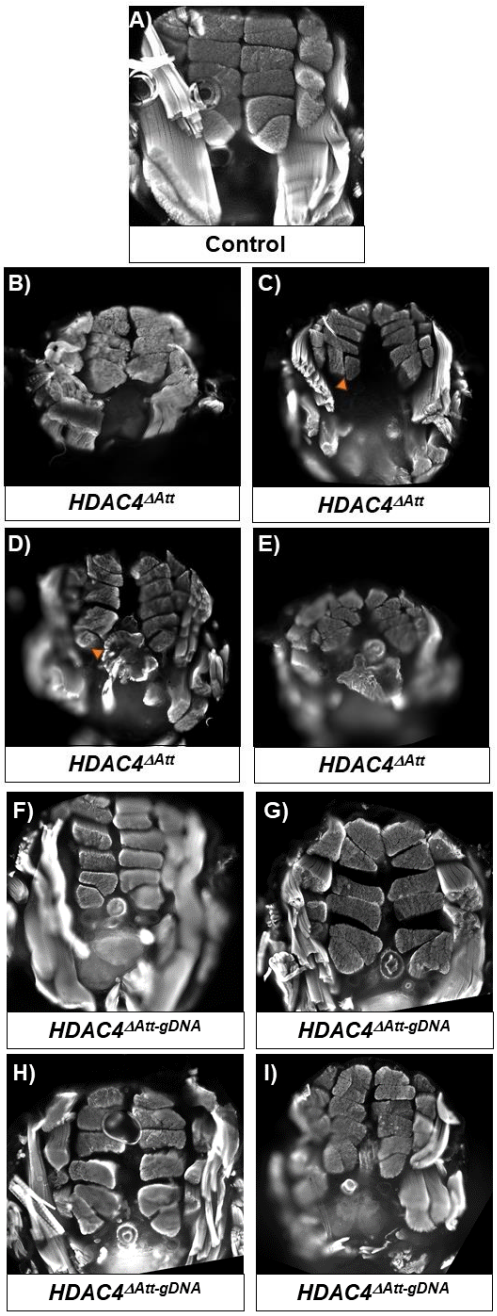


Figure 4. 14 *HDAC4* loss of function mutants display a DLM phenotype
Phalloidin-stained thoracic cross-sections showing organisation of the Dorso-longitudinal flight muscles (DLMs) in control (**A**), *HDAC4*$^{\Delta att}$ (**B-E**), or *HDAC4*$^{\Delta att\text{-}gDNA}$ (**F-I**). Orange arrows indicate visible defects in DLM fibre morphology.

4.10 HDAC4 knockdown does not phenocopy *HDAC4*$^{\Delta Att}$ LOF

The next question to answer would be whether the muscle phenotypes associated with *HDAC4* LOF are muscle-specific, as opposed to a more general affect caused by global *HDAC4* LOF. To do so, I focused on DLM formation, given that previously, *HDAC4* knockdown in the developing larval somatic musculature during

embryogenesis did not produce a phenotype (Schnorrer et al. 2010). I used RNAi to knock down *HDAC4* in developing DLM muscles using either 1151-Gal4, or Act88F-Gal4. The first line I used was GD20522, a line originating from the VDRC library which has been shown to induce knockdown of approximately 50% when expressed in the fly brain (Fitzsimons et al. 2013). However, knockdown with this line using either 1151--, or Act88F-Gal4 had no effect on DLM formation, and the fibres had a morphology and organisation indistinguishable from controls (**Fig. 4.15A,G**). However, different RNAi transgenes may yield different knockdown efficiency (Heigwer et al. 2018), and given that *HDAC4*$^{\Delta att}$ mutants displayed only a relatively subtle DLM phenotype (**Fig. 4.14**), it was possible RNAi with this line did not induce sufficient *HDAC4* knockdown to present with a phenotype. Thus, I tested three additional RNAi lines: VSH330055, a new line from the VDRC short-hairpin library; and two RNAi lines from the Harvard TRiP RNAi project (BL28549&BL34774). Only the VSH330055 line was tested using both the 1151- and Act88F-Gal4 drivers, whereas the TRiP lines were tested using only the 1151-Gal4 driver. However, *HDAC4* knockdown using any of these lines also yielded a WT phenotype indistinguishable from controls (**Fig. 4.15**). Moreover, in an attempt to boost knockdown, I also raised the temperature at which the crosses were raised to 29°C, however this also resulted in a WT DLM phenotype. Finally, in future, one could use an alternative muscle driver, such as Mef2-Gal4, which drives expression throughout muscle development, rather than 1151- or Act88F-Gal4 which display more temporally restricted patterns of expression.

These data may suggest muscle specific knockdown of *HDAC4* may not affect DLM development in the same way as *HDAC4*$^{\Delta att}$ global LOF mutants. However, this may not be conclusive given I did not have time to do qPCR analysis to quantify *HDAC4* knockdown efficiency in developing muscle using these RNAi lines. It is entirely plausible that knockdown was simply not successful, or sufficient to replicate the *HDAC4*$^{\Delta Att}$ phenotype. To attempt to give some insight into this possibility, as a final experiment I used a daughterless-Gal4, a driver thought relatively ubiquitously expressed throughout development, although of course not necessarily at the highest level in all developing tissues (Wodarz et al. 1995). This was to see whether *HDAC4* knockdown could phenocopy the lethality observed in *HDAC4*$^{\Delta att}$ mutants. However, *HDAC4* knockdown with this driver, using either GD20522, VSH330055, or BL28549

appeared to have no obvious effect on the progeny of the cross. Therefore, the effect of *HDAC4* knockdown on DLM development remains inconclusive.

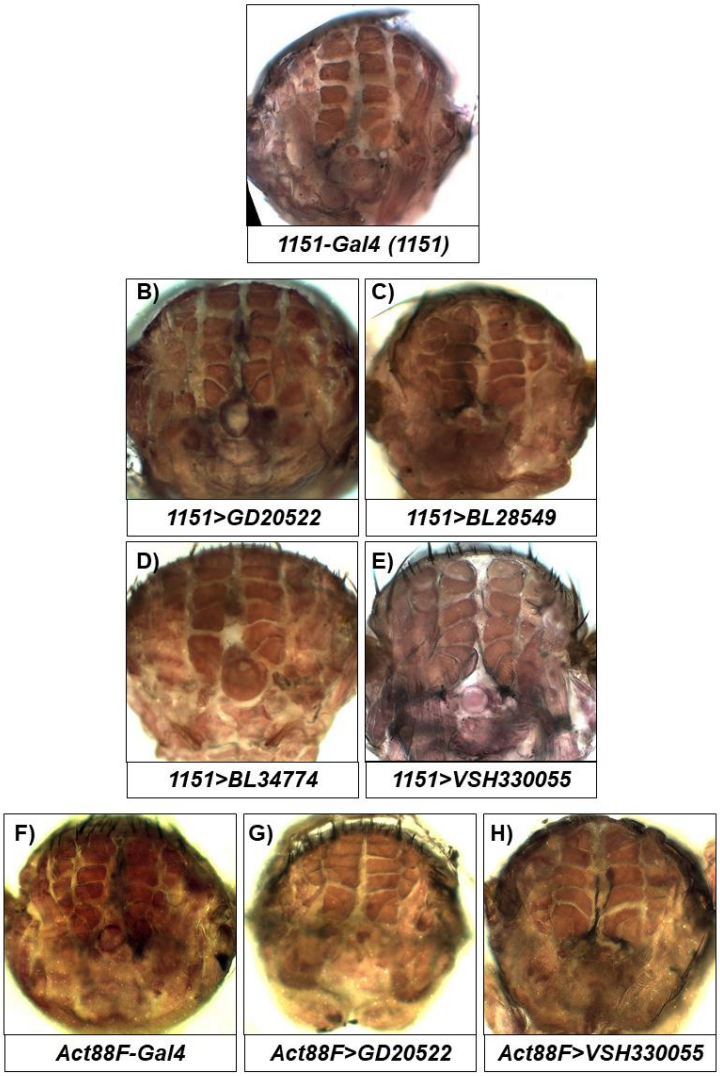


Figure 4. 15. *HDAC4* knockdown does not induce a larval muscle phenotype.
Hematoxylin-stained thoracic cross sections of either *1151-Gal4* driven **(A-E),** or *Act88F-Gal4* driven (**F-H**) knockdown of HDAC4. *HDAC4* knockdown does not affect DLM formation.

Results summary

In this chapter, I have successfully generated a novel *HDAC4* insertion-ready deletion allele, and utilised it to carry out some initial experiments investigating its potential role during muscle development. I firstly show that *HDAC4* loss-of-function is lethal. Secondly, *HDAC4* LOF mutants also present with muscle phenotypes, particularly in the larval somatic musculature, possibly indicating a role In *Drosophila* muscle

development *in vivo*. Importantly, these phenotypes are rescuable by recombination of a gDNA rescue fragment into the edited locus, thereby reconstituting all DNA sequences lost during the initial gene-targeting. This not only emphasises the potential power and significance of this genetic tool for the study of *HDAC4* function, but also reassures users of this genetic tool that any phenotypes observed are directly due to *HDAC4* LOF, rather than uncertainties surrounding potential off-target effects.. These data do indicate a possible role for *HDAC4* in *Drosophila* muscle development *in vivo.* However, reflecting on the data in chapter 3, thus far the data does not conclusively show a function for endogenous HDAC4 in negatively regulating Mef2 transcriptional activity *in vivo*. Nevertheless, the versatility of this novel allele not only shows the consequences *HDAC4* LOF on *Drosophila* survival, but also facilitates more intricate future genetic experiments to probe endogenous HDAC4 function. In particular, in this context this will be in the context of Mef2 regulation during muscle differentiation, but also in any other biological context where HDAC4 may function.

Chapter 5: Generating insertion-ready deletion alleles for the study of Mef2 *in vivo*, and novel tagged alleles to study HDAC4/Mef2 localisation and expression.

Introduction

Much progress has been made into understanding Mef2 function *in vivo* using *Drosophila* as a model system. However, with the advent of CRISPR-Cas9, introduced in **1.6.2** and **1.6.3**, and implemented in this project in **chapter 4**, there is now greater opportunity to manipulate the *Mef2* locus for more targeted functional analysis. As described in **chapter 4**, the use of CRISPR-Cas9 to generate a rescuable *HDAC4* null mutant demonstrates the ability to use the CRISPR-Cas9 mediated insertion-ready deletion allele technology to generate novel, versatile alleles which can be used to investigate endogenous gene function in *Drosophila in vivo*. Furthermore, because Mef2 has many functionally important domains throughout its length, insertion-ready deletion alleles could provide a platform for the systematic analysis of different region of Mef2.

Secondly, a key part of the *in vitro* model for Class IIa HDAC function is how its subcellular localisation may affect its ability to regulate Mef2 function *in vivo*. Indeed, while there is a Mef2 antibody, none exist for HDAC4, hampering the use of antibodies for the study of HDAC4/Mef2 function. Moreover, while antibody staining can be useful, they are limited to the use of fixed tissue and cannot be used to analyse protein dynamics in living cells. However, CRISPR-Cas9 technology also enables researches to tag endogenous loci, which would mitigate the aforementioned challenges associated with antibody staining, particularly in the context of Mef2 and HDAC4 biology. Indeed currently, there is no existing tagged *Mef2* allele, while an existing *HDAC4* protein trap allele is limited by only tagging a subset of *HDAC4* isoforms. Therefore, novel tagged alleles of *HDAC4* and *Mef2* would be of significant benefit to researchers. Thus. In this chapter I implement CRISPR-Cas9 to address the below research aims:

5.1 Targeting the Mef2 locus for the generation of novel insertion-ready deletion alleles

Given the fundamental interest in how Mef2 is regulated *in vivo*, I aimed to generate insertion-ready deletion alleles of the *Mef2* locus, such that a versatile platform for the study of endogenous Mef2 *in vivo*, could be generated. To do so, the same general approach was followed as for the generation of *HDAC4* insertion-ready deletion alleles. Unlike *HDAC4*, the *Mef2* locus located on chromosome II (46C4-46C7) (2R) is smaller, with the entire coding sequence only spanning approximately 5kb. However,

the total gene region for the majority of isoforms (A,B,C,D,I, J,K,L) including UTR-exons, spans a region of approximately 16kb (**Fig. 5.1**).

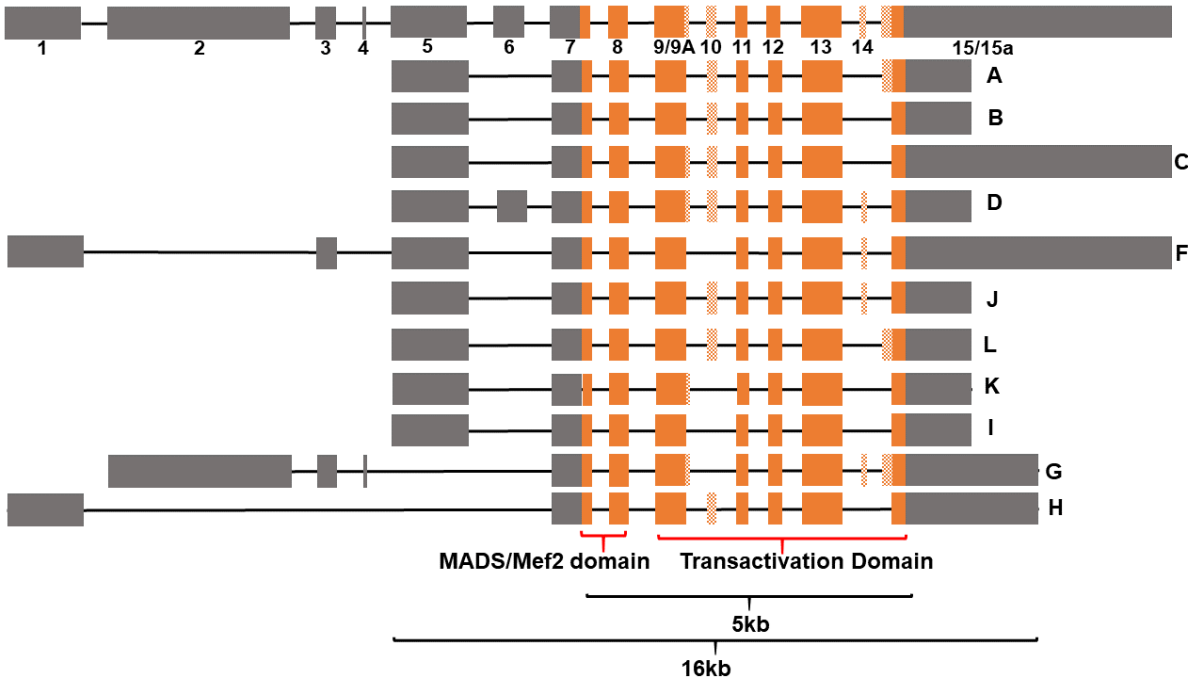


Figure 5. 1. Mef2 gene structure and annotated splice variants.
(Top) schematic of all possible exons found in different combinations in different Mef2 genomic locus. (below) Existing annotated Mef2 isoforms, adapted from flybase. Regions of the Mef2 sequence encoding the MADS/Mef2 and transactivation domains are indicated. Non-coding exons are in grey, coding exons are in orange. Alternatively spliced coding exons are displayed as an orange/white chequered pattern.

Because increasing deletion size reduces efficiency of HDR-mediated repair following Cas9-cleavage (Poernbacher et al. 2019), I aimed to remove as little DNA as possible to maximise the potential for generating the desired transgenic. However, while this is necessary, I also wanted to ensure this gene targeting would generate a *Mef2* null allele. Thus, two different approaches to engineer the *Mef2* locus were designed: In the first, from here on in referred to as $Mef2^{\Delta CDS}$, the aim was to delete all coding exons by utilising a 5' CRISPR site within the 5' UTR, and the 3' CRISPR just downstream of the stop codon (**Fig. 5. 2B**). In total this would induce a 5.3kb deletion. In the second, referred to as $Mef2^{\Delta ATG}$, a minimal deletion of only 0.25kb was designed, utilising the same 5' CRISPR site as for $Mef2^{\Delta CDS}$, but instead choosing a 3' CRISPR site in the second intron (of Mef2 isoform C) (**Fig. 5. 2C**). I hypothesised that, although this deletion would remove only 54bp of CDS, the equates to the sequence that encodes the first 18 amino acids of the MADS/Mef2 domain. Thus, as previously described,

because the MADS/Mef2 domain is highly conserved, vital for Mef2 function, and is sensitive to mutation, this deletion may be sufficient to also generate a *Mef2* null. Before targeting, CRISPR sites in the *yw;;nos-Cas9* injection stock were verified by PCR, followed by sequencing (**Fig. 5.2D-F**). 5' and 3' homology arm ends were designed to be adjacent to the DSB location, in between base-pairs 17-18 of the protospacer sequence (**Fig. 5.2D-F**).

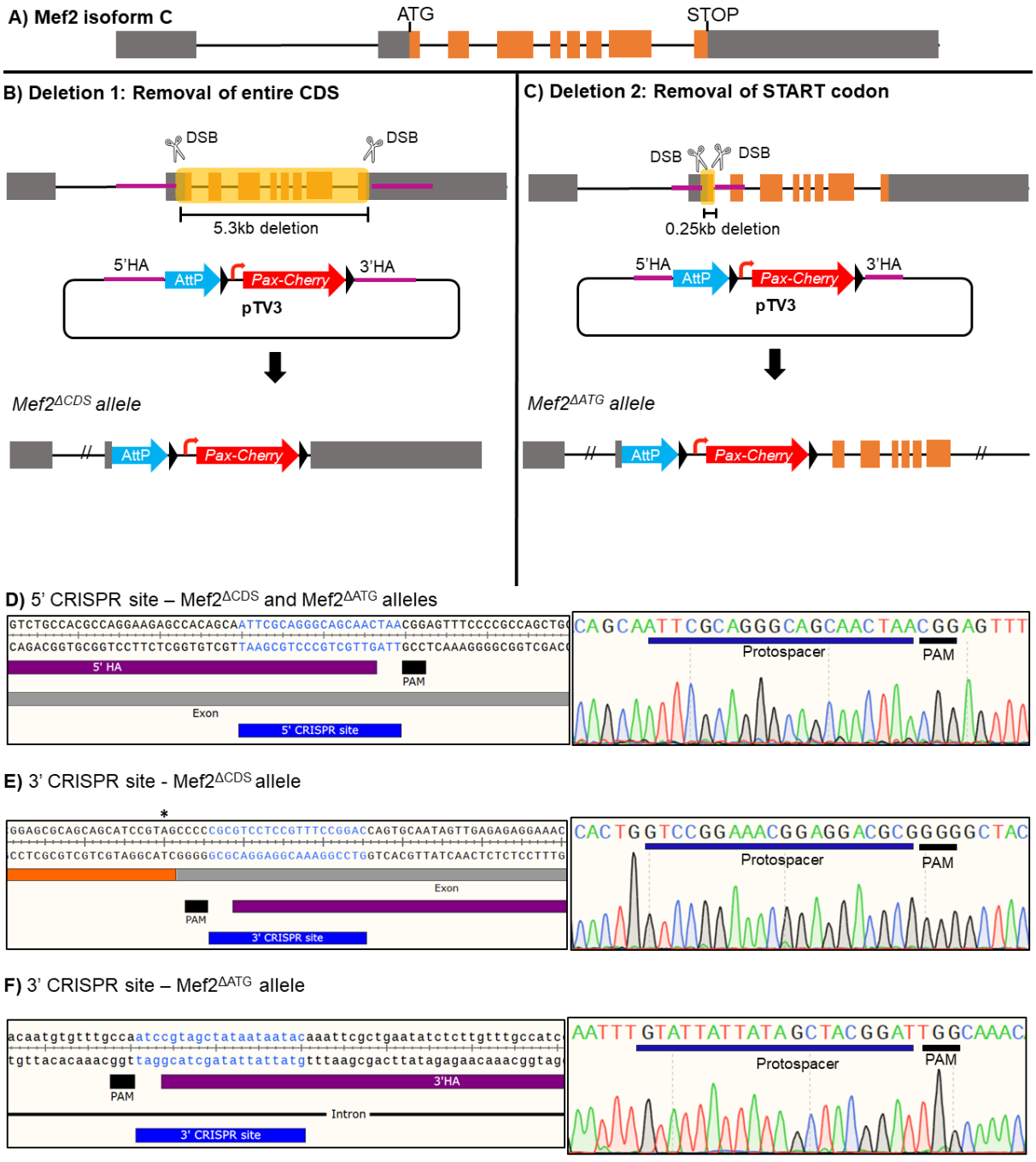


Figure 5. 2. Targeting of the Mef2 locus to generate insertion-ready deletion alleles.
A) Schematic of Mef2 intron-exon structure, isoform C. **(B)** Design of a deletion to generate the allele Mef2$^{\Delta CDS}$. **(C)** Design of deletion to generate the allele Mef2$^{\Delta ATG}$. The targeting

vector, pTV3 contains an attP landing site followed by a *Pax-cherry* marker cassette flanked by LoxP sites (black arrows). Homology arms cloned into pTV3 vector equate to those genomic regions either side of the double-strand break (DSB) points induced by Cas9. Deletion size for each design is indicated in yellow shading. Structure of designed resulting alleles from these two targeting approaches are indicated. **(D,F,H)** CRISPR site and relative homology arm positioning. Mef2$^{\Delta CDS}$ and Mef2$^{\Delta ATG}$ share a 5' CRISPR site **(D),** whereas the 3' CRISPR sites are different for Mef2$^{\Delta CDS}$ **(F)** and Mef2$^{\Delta ATG}$ **(H).** (**E,G,I**) DNA sequence trances showing CRISPR site verification in injection lines for 5' CRISPR site (**E),** and 3' CRISPR sites for Mef2$^{\Delta CDS}$ **(F)** and Mef2$^{\Delta ATG}$ **(G).**

5.2 *Mef2* can be successfully targeted to generate insertion-ready deletion alleles

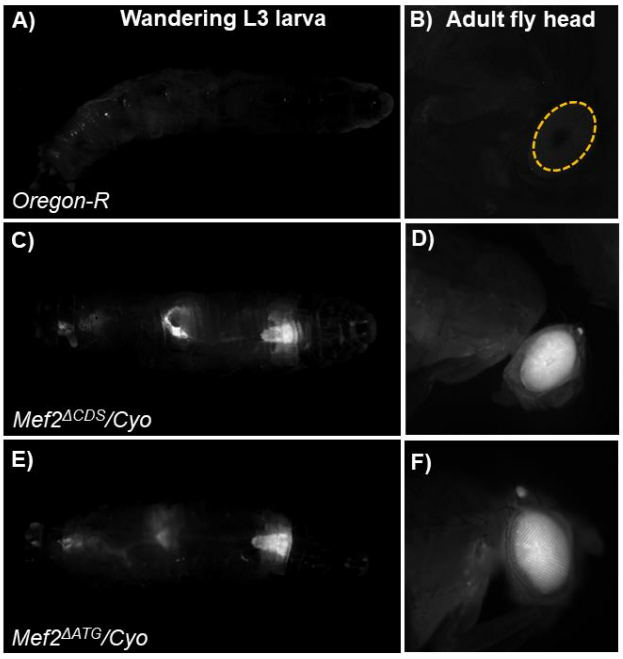


Figure 5. 3. Verification of Mef2 insertion-ready deletion alleles through Pax-cherry marker detection.
Expression of *Pax-cherry* was visualised in the developing larval CNS **(A,C,E)** and in the adult fly eye **(B,D,F).** *Oregon-R* control flies do not display any cherry expression **(A,B),** whereas flies harbouring the Mef2$^{\Delta CDS}$ (**C-D**) or Mef2$^{\Delta ATG}$ (**E-F**) allele display visible cherry expression in both tissues.

Following microinjection, transgenic flies harbouring the engineered chromosome were initially screened by analysing *Pax-cherry* marker cassette expression in larvae and adult flies, similar to *HDAC4$^{\Delta Att}$* (**Fig. 5.3**). Again compared to *Oregon-R*, which do not express cherry in either the L3 wandering larvae, or the adult fly eye (**Fig. 5.3**),

both *Mef2*$^{\Delta CDS}$ and *Mef2*$^{\Delta ATG}$ display characteristic cherry expression in both the developing larval CNS, and the adult fly eye (**Fig. 5.3**). Next, I aimed to genotype these transgenics to further define the changes induced by both alleles. To do so, PCR was used to amplify the engineered region (**Fig. 5.4A-B**). Seven and six independent transgenic lines were received for *Mef2*$^{\Delta CDS}$ and *Mef2*$^{\Delta ATG}$ respectively, and all lines were genotyped by PCR. Because *Mef2*$^{\Delta CDS}$ and *Mef2*$^{\Delta ATG}$ did not homozygose and remained over a balancer chromosome, this firstly indicated these novel alleles were *Mef2* loss-of-function. However, this also meant that there were two possible amplicons from the PCR: The first would be from the engineered alleles, 1.7kb for *Mef2*$^{\Delta CDS}$ and 1.6kb for *Mef2*$^{\Delta ATG}$ **(Fig. 5.4A**); and the second from the balancer chromosome which harbours a wild-type *Mef2* locus. Successful amplification from this chromosome would yield products of 5.8kb (*Mef2*$^{\Delta CDS}$) and 528bp (*Mef2*$^{\Delta ATG}$). However, because the PCR cycling conditions were set for amplification of the engineered locus, a 5.8kb band would not be amplified. PCR amplification of both engineered alleles generated amplicons of the predicted sizes, in all independent transgenic lines received for both *Mef2*$^{\Delta CDS}$ (7 lines), and *Mef2*$^{\Delta ATG}$ (6 lines) (**Fig. 5.4B**). For the *Mef2*$^{\Delta ATG}$ allele, the band sat slightly lower relative to the ladder than initially predicted. Furthermore, we also see a 528bp amplicon reflective of the amplification of the *Mef2* locus from the balancer chromosome. Amplicons were then also purified and sequenced to accurately define these new alleles (**Fig. 5.4C-E**). This DNA sequencing further verified the successful engineering of the *Mef2* locus for both *Mef2*$^{\Delta CDS}$ and *Mef2*$^{\Delta ATG}$: The boundaries of the engineered regions for both alleles were accurately defined by the homology arm placement, confirming again the accuracy of the CRISPR-Cas9 system. This also suggests the location of the *Mef2*$^{\Delta ATG}$ amplicon, appearing below the 1.5kb line, may be inaccurate.

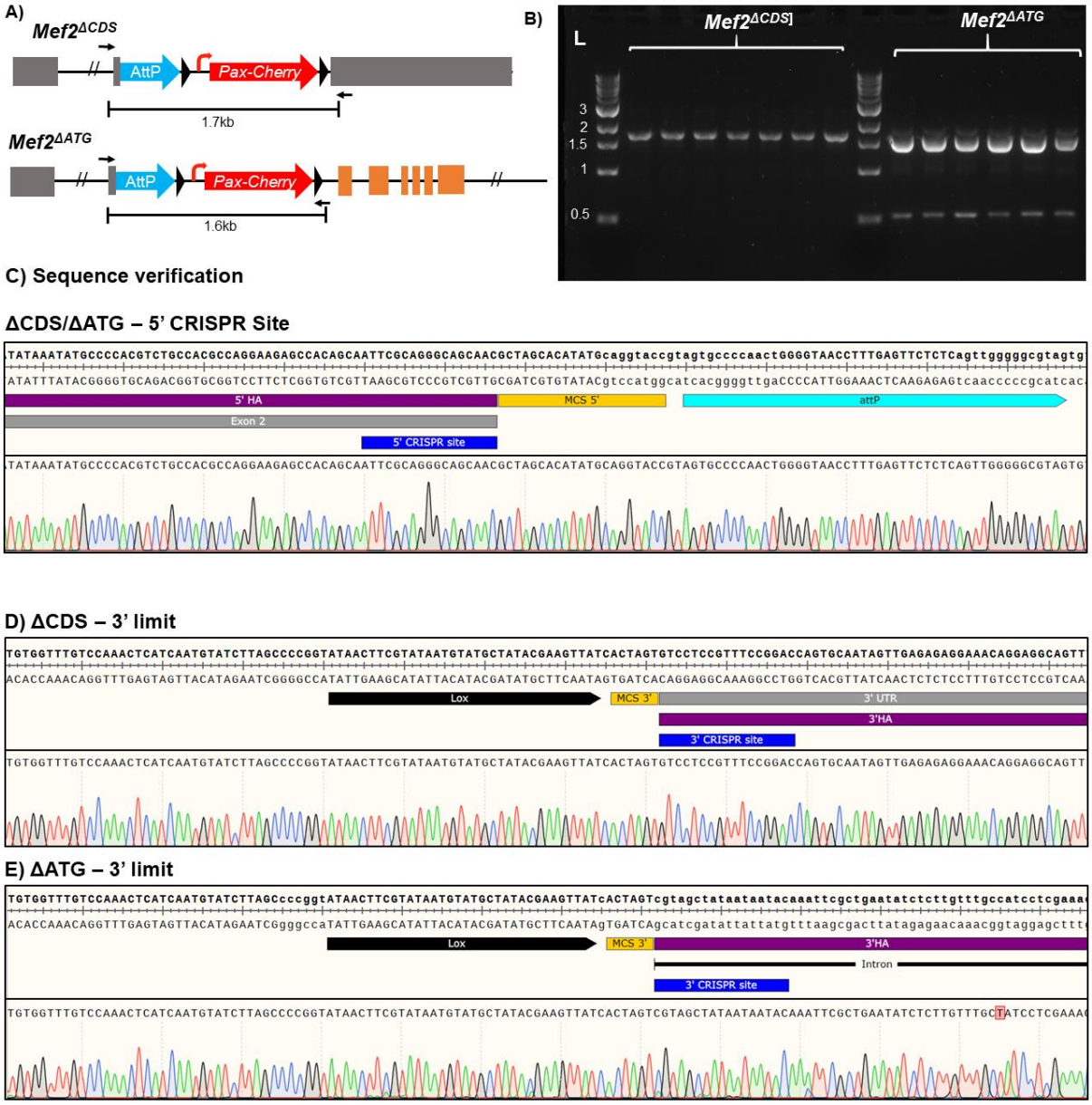
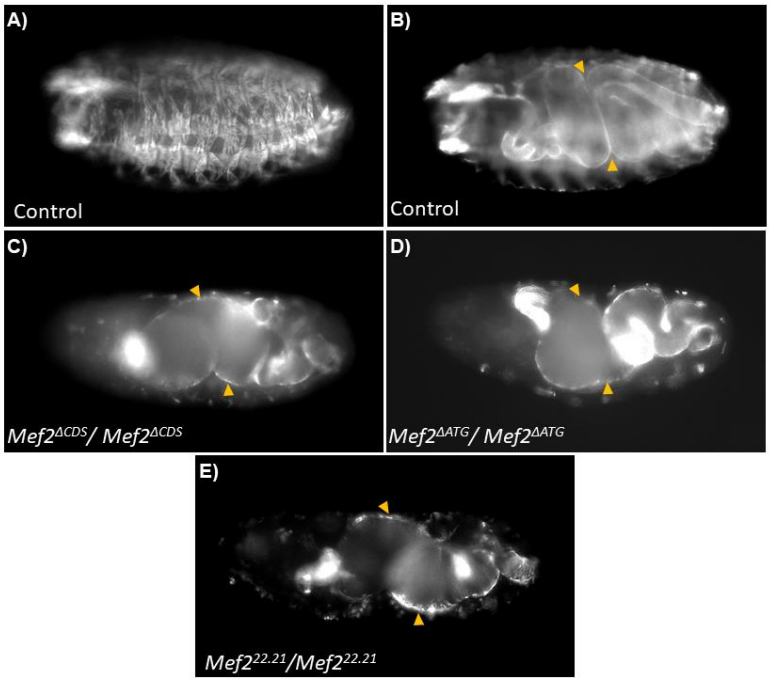


Figure 5. 4. Molecular verification of Mef2 insertion-ready deletion alleles.
(A) Schematic of the Mef2$^{\Delta CDS}$ and Mef2$^{\Delta ATG}$ alleles, and predicted PCR product sizes indicated using primers either side of the engineered regions (black arrows). **(B)** Agarose gel shown PCR products generated from amplification (using primers described in (A) from gDNA extracted from heterozygous engineered flies. 7 and 6 independent transgenic lines were obtained for Mef2$^{\Delta CDS}$ and Mef2$^{\Delta ATG}$ designs respectively. ~500bp product present in Mef2$^{\Delta ATG}$ lanes indicate amplification of smaller product from wild-type *Mef2* copy on balancer chromosome. **(C-E)** DNA-sequencing verification of ends of purified PCR products in relation to positioning the of 5' CRISPR site **(C),** and the 3' CRISPR sites for Mef2$^{\Delta CDS}$ **(D)**and Mef2$^{\Delta ATG}$ (**E**), showing successful engineering of the Mef2 locus for both alleles. Sequence trace for 5' CRISPR site shows Mef2 CDS only.

5.3 *Mef2*$^{\Delta CDS}$ and *Mef2*$^{\Delta ATG}$ are null alleles

Following the initial verification of the *Mef2*$^{\Delta CDS}$ and *Mef2*$^{\Delta ATG}$ alleles, we next sought to investigate the phenotypic consequences of these modifications. Given the entire CDS was removed in the generation of the *Mef2*$^{\Delta CDS}$, it was almost certain this would be a Mef2 null allele. However, given only 250bp was deleted in the *Mef2*$^{\Delta ATG}$ allele, the functional consequences were perhaps less certain. As previously mentioned, because these alleles are not homozygous viable, this was the first line of evidence suggesting that these were Mef2 loss-of-function alleles, given that Mef2 function is essential for viability (Bour et al. 1995; Lilly et al. 1995; Ranganayakulu et al. 1995; Nguyen et al. 2002). Nevertheless, the embryonic musculature in *Mef2* mutant embryos were analysed compared to *Mef2*$^{22.21}$, a known *Mef2* null (Bour et al. 1995). To do this, I used an antibody against Mhc (**Fig. 5.5**). Given that all lines were verified by PCR and sequencing, there was no reason to expect different phenotypes between different independent mutant lines for each design. Therefore, for this analysis I used a single transgenic line for *Mef2*$^{\Delta CDS}$and *Mef2*$^{\Delta ATG}$.

Embryos homozygous for *Mef2*$^{\Delta CDS}$ (**Fig. 5.5C)** and *Mef2*$^{\Delta ATG}$ (**Fig. 5.5D**) phenocopy *Mef2*$^{22.21}$ (**Fig. 5.5E)**, displaying a complete lack of differentiated somatic muscle. Furthermore, akin to *Mef2*$^{22.21}$ I also saw characteristic disruption to gut morphology in *Mef2*$^{\Delta CDS}$ and *Mef2*$^{\Delta ATG}$ homozygous embryos. This is an established phenotype associated with Mef2 loss-of-function, caused by disruption to the expression of *α-PS2*, an integrin required for normal visceral muscle formation, and whose expression is regulated by Mef2 (Ranganayakulu et al. 1995). To confirm embryonic lethality of *Mef2*$^{\Delta CDS}$ and *Mef2*$^{\Delta ATG}$ homozygous mutants, I also counted hatched larvae from laying pots of each *Mef2*$^{\Delta CDS}$ and *Mef2*$^{\Delta ATG}$ balanced, heterozygous stocks. Homozygous mutant larvae for either genotype were not observed, confirming the embryonic lethality of *Mef2*$^{\Delta CDS}$ and *Mef2*$^{\Delta ATG}$ homozygous mutants (**Fig. 5.5F**). Finally, I also crossed *Mef2*$^{\Delta CDS}$, *Mef2*$^{\Delta ATG}$, and *Mef2*$^{22.21}$ to each other and found that trans-heterozygotes were also homozygous lethal, since all flies which emerged from these crosses harboured the balancer chromosome (Cyo) (**Fig. 5.5G)**. These data support that both *Mef2*$^{\Delta CDS}$ and Mef2$^{\Delta ATG}$ are novel *Mef2* null alleles, and that notably, deletion of only 54bp of Mef2 CDS (*Mef2*$^{\Delta ATG}$) is also sufficient to completely abrogate Mef2 function *in vivo*.



F)

Genotype	Expected ratio	Observed ratio
Mef2ΔATG/CyoAct5CGFP	0.5	1.0
Mef2ΔATG/Mef2ΔATG	0.25	0
CyoAct5CGFP/CyoAct5CGFP	0.25	0
Mef2ΔCDS/CyoAct5CGFP	0.5	1.0
Mef2ΔCDS/Mef2ΔCDS	0.25	0
CyoAct5CGFP/CyoAct5CGFP	0.25	0

G)

	$Mef2^{\Delta ATG}$	$Mef2^{\Delta CDS}$	$Mef2^{22.21}$
$Mef2^{\Delta ATG}$		X	X
$Mef2^{\Delta CDS}$	X		X
$Mef2^{22.21}$	X	X	

Figure 5. 5. *Mef2* insertion-ready deletion alleles are null.
Somatic muscle pattern of late-stage embryos, stained with an antibody against myosin heavy chain (*Mhc*). (**A**) Somatic muscle pattern of wild-type control. (**B**) Gut morphology of wild-type control embryos. Embryos homozygous for $Mef2^{\Delta CDS}$ (**C**) or $Mef2^{\Delta ATG}$ (**D**) display a lack of complete lack of somatic muscle in late stage embryos, and this phenocopies that of an existing *Mef2* null mutant, $Mef2^{22.21}$ (**E**). This is also accompanied with disruption to midgut morphology (indicated by orange arrows). (**F**) Homozygous Mef2 mutants are embryonic lethal. Ratio of genotypes of L1 larvae screened for homozygous mutants, based on the absence of fluorescent balancer chromosome. n=150 for each genotype. (**G**) Mef2 trans-heterozygote mutants are lethal. Flies of each genotype were crossed to each other, with homozygous viability scored by absence of balancer chromosome. X = non-viable.

5.4 Designing constructs to rescue *Mef2*$^{\Delta CDS}$ and *Mef2*$^{\Delta ATG}$

Given the successful generation of novel *Mef2* insertion-ready deletion alleles, I next designed constructs aimed to rescue the *Mef2* loss-of-function phenotype. This is because, successful rescue of this phenotype would be due to the initial Mef2 engineering and not an unknown off-target effect. Because, and unlike *HDAC4*^{ATT}, the start codon was removed during the targeting of the *Mef2* locus, there was also a possibility of rescuing with a cDNA construct as well as gDNA.

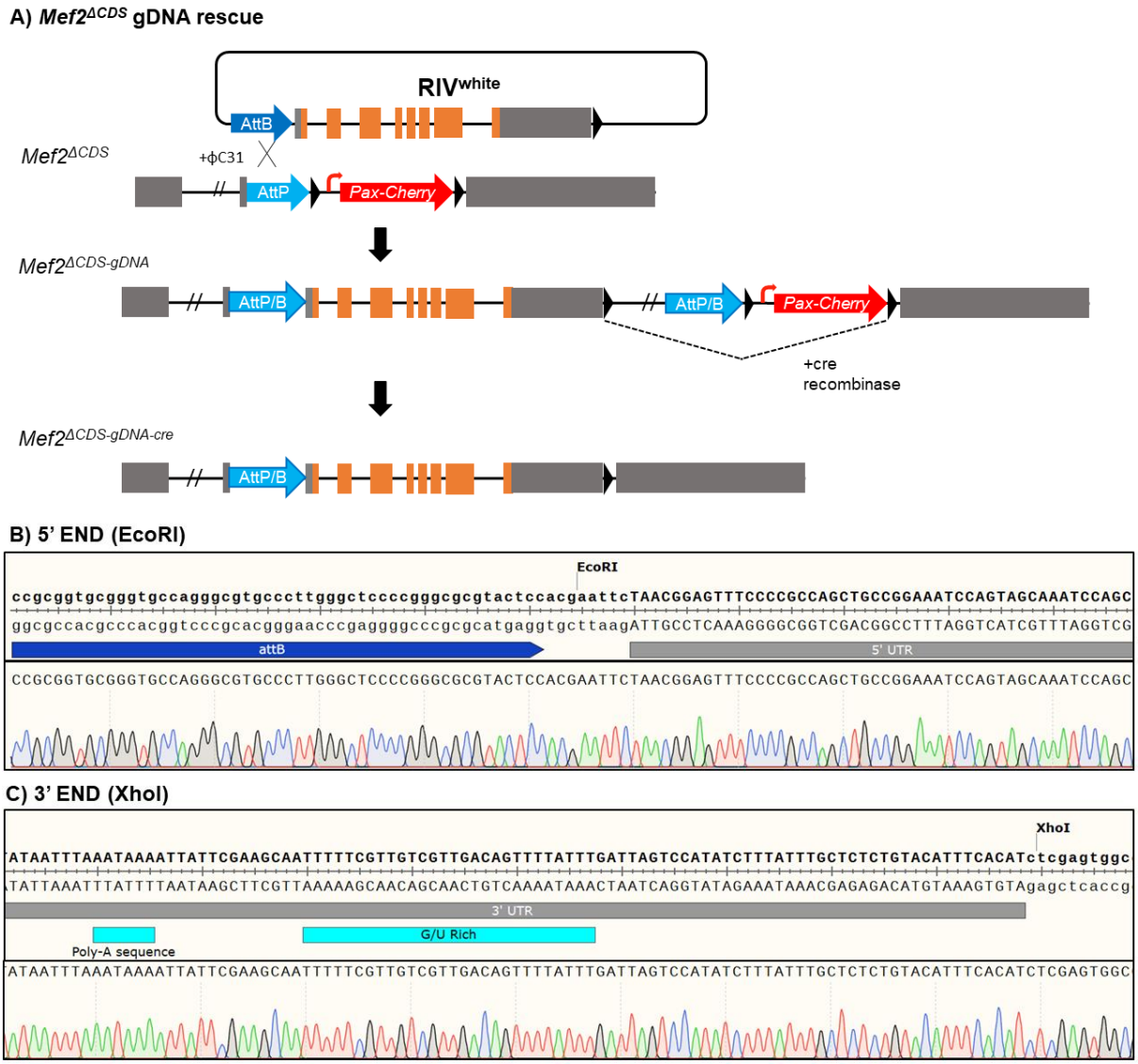


Figure 5. 6. Design of gDNA-based rescue construct to restore function of Mef2ΔCDS. Schematic of a gDNA rescue to restore function of the *Mef2*$^{\Delta CDS}$ allele. **(B-C)** DNA sequence map, and DNA sequence chromatogram of 5' and 3' ends respectively of rescue construct inserted into RIV white. EcoRI and XhoI enzymes are indicated as they were used to clone in the rescue fragment. **(B)** schematic (top) and DNA sequence chromatogram (bottom) of 5'

end of rescue construct, showing the first bases of the rescue fragment following the GAATTC EcoRI site (TAA) is the boundary of what was initially deleted during *Mef2* gene targeting. **(C)** Schematic (top) and DNA sequence chromatogram (bottom) of 3' end of rescue construct. The amount of 3' UTR incorporated in the rescue construct equated to where a transcription termination sequence was identified in the *Mef2* 3'UTR. Poly-A and G/U rich sequences in blue indicate sequence identifiable transcription termination sequence within the *Mef2* 3'UTR.

For *Mef2* $^{\Delta CDS}$, I designed a gDNA construct, with the 5' end beginning at the deletion point of the initial targeting, such that all DNA removed during the targeting would be re-inserted (**Fig. 5.6A-B**). At the 3' end, we included 1.3kb of 3' UTR, despite the 3' breakpoint leaving much of the endogenous 3' UTR intact. This is because the *Pax-Cherry* marker cassette would interrupt the endogenous 3'UTR if we relied upon 'fusion' between the inserted gDNA rescue construct, and the endogenous *Mef2* sequence that had not been initially engineered (**Fig. 5.6A**). Thus, the rescue construct was designed to include 3' UTR up until the sequence-identifiable transcription termination sequence within the *Mef2* 3'UTR (**Fig. 5.6C**). By contrast, for the gDNA rescue of *Mef2*$^{\Delta ATG}$ (**Fig. 5.7A**), the same 5' end was incorporated, given it shares the same 5' boundary as the *Mef2*$^{\Delta CDS}$ allele (**Fig. 5.7B**), however because the majority of the *Mef2* coding sequence remained intact, I only reinserted the same 250bp that had initially been deleted, with the 3' boundary of the inserted rescue construct matching the region initially deleted during gene targeting (**Fig. 5.7C**). To restore *Mef2* gene function with this construct, it would then require a successful 'hybrid' between endogenous sequence not targeted, and the sequence that is inserted as part of the genomic rescue.

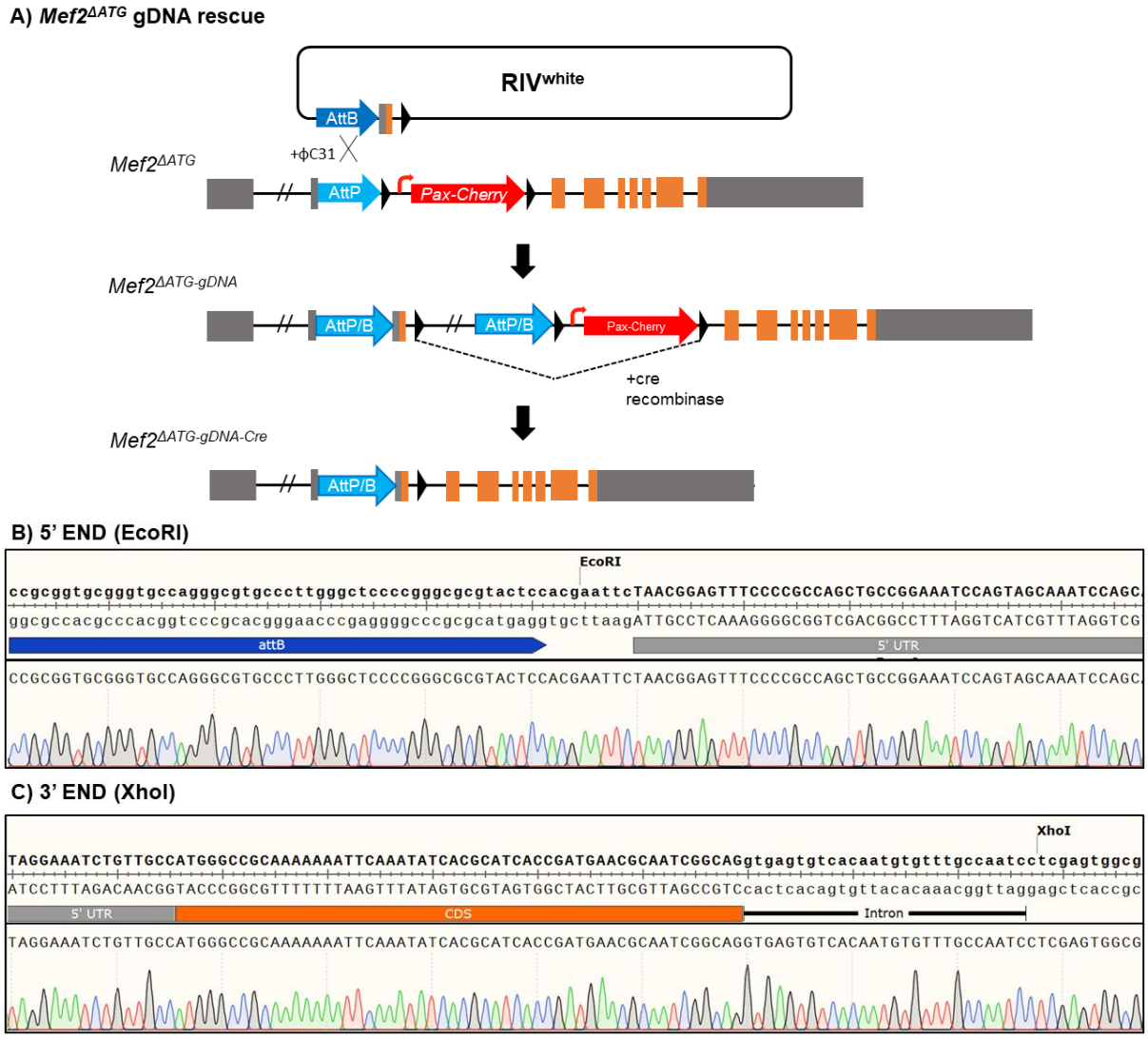


Figure 5. 7. Design of gDNA-based rescue construct to restore function of Mef2ΔATG. **(A)**Schematic of a gDNA rescue to restore function of the Mef2$^{\Delta ATG}$ allele. **(B-C)** DNA sequence map, and DNA sequence chromatogram of 5' and 3' ends respectively of rescue construct inserted into RIV white. EcoRI and XhoI enzymes are indicated as those used to clone in the rescue fragment. **(B)** schematic (top) and DNA sequence chromatogram (bottom) of 5' end of rescue construct, showing the first bases of the rescue fragment following the GAATTC EcoRI site (TAA) are is the boundary of what was initially deleted during *Mef2* gene targeting. **(C)** Schematic (top) and DNA sequence chromatogram (bottom) of 3' end of rescue construct. As with the 5' boundary, the amount of gDNA cloned into the rescue construct was defined by the boundary of the initially deletion.

For the cDNA rescue, I designed the same construct to rescue both *Mef2*$^{\Delta CDS}$ and *Mef2*$^{\Delta ATG}$ (**Fig. 5.8A**). Because there were no prior assumptions with regards to which *Mef2* isoform to rescue with, I chose the coding sequence for isoform C, for which we have used throughout the project in the generation of Mef2 overexpression constructs, and was the cDNA available to at the time. However, the cDNA clone did not possess

all the 3’ UTR itself which, akin to the *Mef2*$^{\Delta CDS}$ gDNA rescue, I wanted to recapitulate as much as possible in the rescue construct. Therefore, this fragment was cloned sequentially from both a cDNA (for the coding sequence), and then gDNA (for residual 3’ UTR). As for both gDNA rescue constructs, the same sequence for rescue from the 5’ breakpoint was used (**Fig. 5.8B**). In contrast to *Mef2*$^{\Delta CDS}$ gDNA rescue however, the amount of 3' UTR incorporated was extended to that of the entire isoform C annotated 3’ UTR (**Fig. 5.8C**) (Gramates et al. 2022), meaning the 3’ UTR in the cDNA construct was approximately 3.2kb in length, as opposed to 1.3kb in the *Mef2*$^{\Delta CDS}$ gDNA rescue.

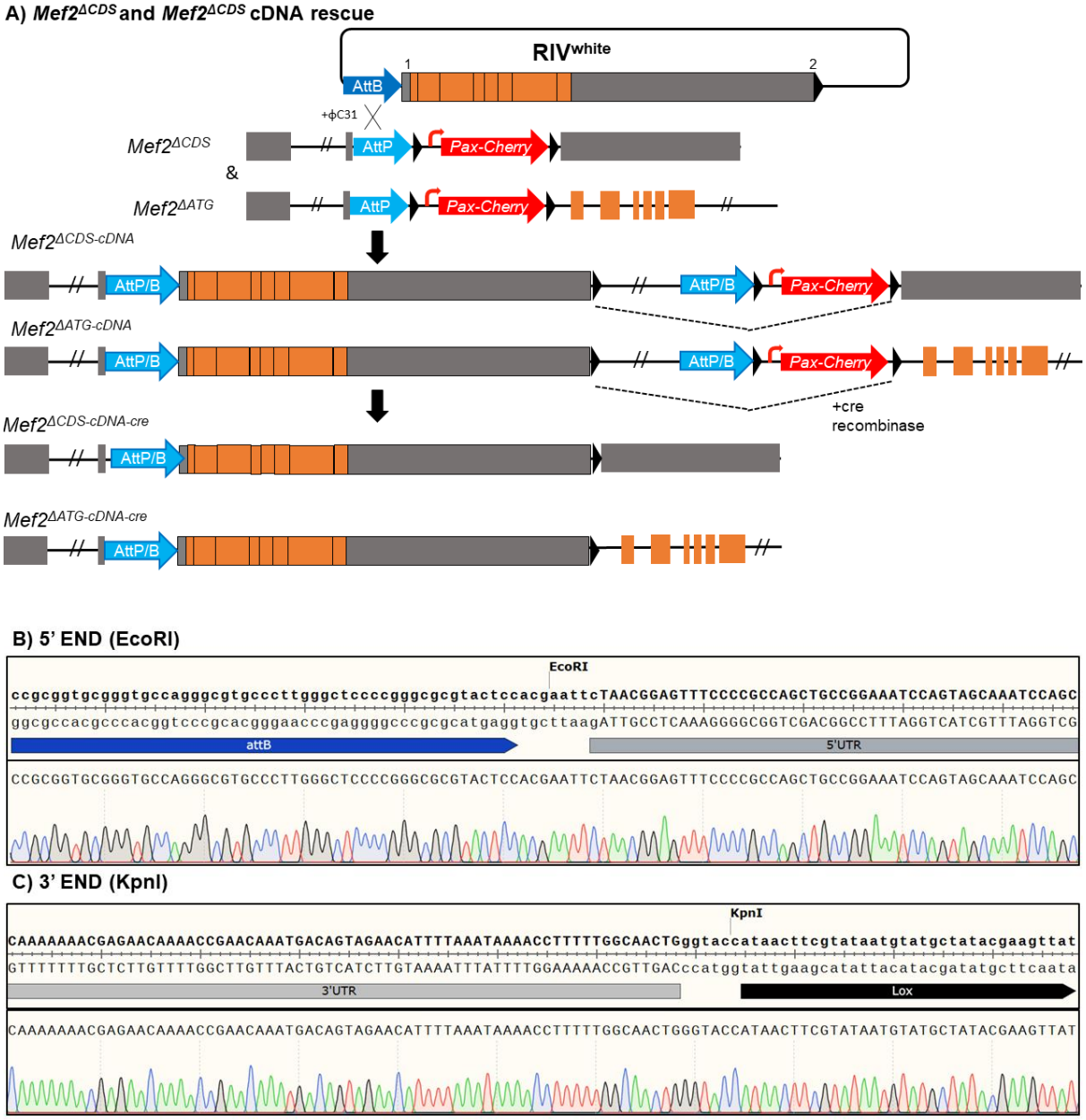


Figure 5. 8. Design of a cDNA based rescue construct to restore function of Mef2ΔATG and Mef2ΔCDS.

(A) Schematic of a cDNA rescue. cDNA from Isoform C was used. **(B-C)** DNA sequence map, and DNA sequence chromatogram of 5' and 3' ends respectively of rescue construct inserted into RIV white. EcoRI and XhoI enzymes are indicated as those used to clone in the rescue fragment. **(B)** schematic (top) and DNA sequence chromatogram (bottom) of 5' end of rescue construct, showing the first bases of the rescue fragment following the EcoRI site (TAA) is the boundary of what was initially deleted during *Mef2* gene targeting. **(C)** Schematic (top) and DNA sequence chromatogram (bottom) of 3' end of rescue construct.

5.5 Differential ability to rescue *Mef2$^{\Delta CDS}$* and *Mef2$^{\Delta ATG}$* alleles

For rescue, I inserted the rescue constructs into one line of each of the *Mef2$^{\Delta CDS}$* and *Mef2$^{\Delta ATG}$* alleles. The injected lines were the same ones as used for the phenotypic analysis. Following the site-specific recombination of rescue constructs into the *Mef2$^{\Delta CDS}$* and *Mef2$^{\Delta ATG}$* alleles, the question was as to whether Mef2 loss of function could be rescued. Because Mef2 is essential for viability, we hypothesised that a successful rescue could initially be inferred by whether these lines were homozygous viable or not. Three independent lines of each of these rescue lines were received: gDNA or cDNA rescue of *Mef2$^{\Delta CDS}$* (*Mef2$^{\Delta CDS\text{-}gDNA}$*/Mef2$^{\Delta CDS\text{-}cDNA}$); gDNA or cDNA rescue of *Mef2$^{\Delta ATG}$* (Mef2$^{\Delta ATG\text{-}gDNA}$/ Mef2$^{\Delta ATG\text{-}cDNA}$). Interestingly, only one rescue line, that of *Mef2$^{\Delta ATG\text{-}cDNA}$* was homozygous viable, and then only in two out of three lines (**Fig. 5.9A**), whereas all other rescue constructs remained over a balancer chromosome. This remained the case over multiple generations from when these fly stocks were initially received. It is possible that the residual *Pax-cherry* cassette in these alleles disrupted *Mef2* function even following rescue construct integration, in particular *Mef2$^{\Delta ATG\text{-}gDNA}$*, where endogenous *Mef2* sequence downstream of the excess integrated vector DNA and *Pax-Cherry* cassette would be required to restore *Mef2* function. To remove this excess DNA in fly lines which did not appear to be homozygous viable, the Cre-lox system was used to excise the DNA present between loxP sites (**Fig. 5.6-5.8**). Flies were screened for cre-mediated excision having taken place by the loss of *Pax-cherry* expression in the adult fly (**Fig. 5.9B**). However, following this, homozygous viability in any of *Mef2$^{\Delta CDS\text{-}gDNA}$*, Mef2$^{\Delta CDS\text{-}cDNA}$ or *Mef2$^{\Delta ATG\text{-}gDNA}$* was still not observed. This suggests that there may be some reasons for why *Mef2* expression and/or Mef2 function may be perturbed in these lines. However they remain unknown.

A)

Genotype	Cre-Lox?	Homozygous viable (/number of independent lines)?
Mef2$^{\Delta CDS\text{-}gDNA}$	N	N (0/3)
Mef2$^{\Delta CDS\text{-}gDNA}$	Y	N (0/3)
Mef2$^{\Delta CDS\text{-}cDNA}$	N	N (0/3)
Mef2$^{\Delta CDS\text{-}cDNA}$	Y	N (0/3)
Mef2$^{\Delta ATG\text{-}gDNA}$	N	N (0/3)
Mef2$^{\Delta ATG\text{-}gDNA}$	Y	N (0/3)
Mef2$^{\Delta ATG\text{-}cDNA}$	N	Y (2/3)
Mef2$^{\Delta ATG\text{-}cDNA}$	N/A	N/A

B)

Figure 5. 9. Verification of rescue of Mef2 insertion-ready deletion alleles by determining homozygous viability.

(A) Table describing homozygous viability of each of the lines generated to rescue Mef2 function in either *Mef2*$^{\Delta CDS}$ or *Mef2*$^{\Delta ATG}$. Viability was assessed before and after cre-mediated excision of excess genomic DNA, apart from in *Mef2*$^{\Delta ATG\text{-}cDNA}$. (**B**) representative example of the fluorescence from transgenic flies containing a rescue DNA fragment, before or after cre-lox. In the absence of cre, *Pax-cherry* expression is still visible, which is lost following cre-mediated excision.

5.6 Mef2 function is restored by *Mef2*$^{\Delta ATG\text{-}cDNA}$ *in vivo*

Despite the variable ability for different rescue constructs to full restore Mef2 function *in vivo,* one line, *Mef2*$^{\Delta ATG\text{-}cDNA}$, was homozygous viable, suggesting this line has restored Mef2 function. To first verify that this line contained the desire insertion, DNA was extracted from one line of the two lines that was homozygous viable, and PCR amplified the genomic region. The same PCR reaction with gDNA template isolated from the heterozygous *Mef2*$^{\Delta ATG}$ line was also run as a control. A schematic of the *Mef2* allelic combination within each line is displayed in **Fig. 5.10A**. Using the same PCR primers, *Mef2*$^{\Delta ATG\text{-}cDNA}$ was predicted to yield a single PCR product of 1.35kb, whereas Mef2$^{\Delta ATG}$ would produce two bands, 4.6kb (from the engineered chromosome) and 3.5kb from the wild-type *Mef2* allele present on the balancer

chromosome. PCR amplification confirmed this, showing the predicted bands present when separated by agarose gel electrophoresis (**Fig. 5.10B**). I also observed a fainter band of approximately 2.1kb present in both samples, however this PCR was not fully optimised and is likely this is an off-target amplification product (**Fig. 5.10B**).

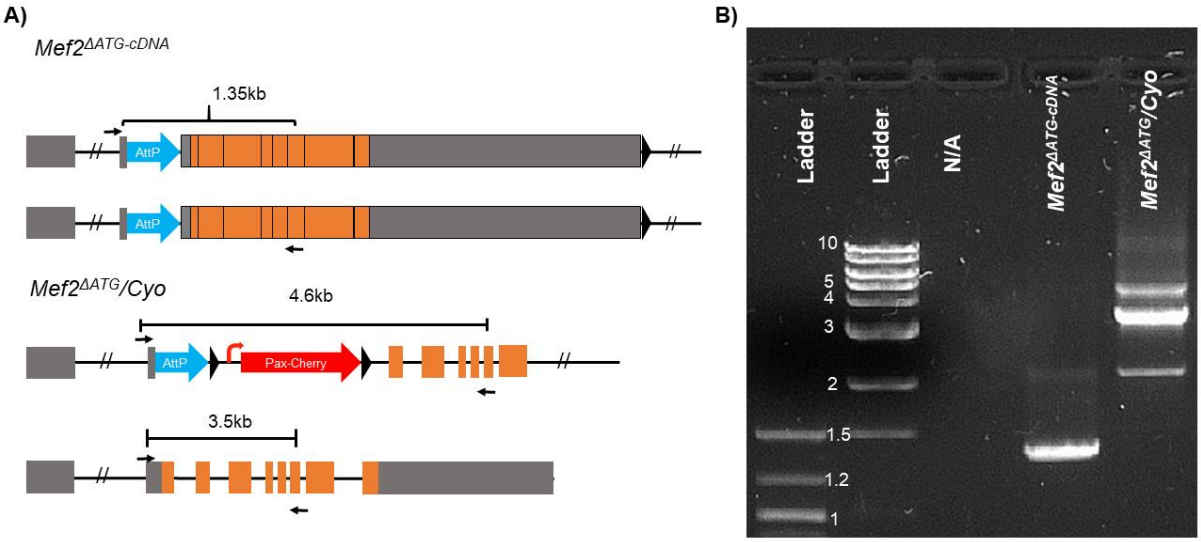


Figure 5. 10.Molecular verification of Mef2ΔATG-cDNA.
(**A**) Schematic showing *Mef2* locus organisation in $Mef2^{\Delta ATG\text{-}cDNA}$ homozygotes, and $Mef2^{\Delta ATG}$ heterozygotes. Using PCR primers to amplify a region of the *Mef2* locus from gDNA extracted from either genotype, expected PCR fragment sizes are indicated. (**B**) Agarose gel electrophoresis of PCR products using primers and gDNA described in (**A**).

Given the PCR verification of this $Mef2^{\Delta ATG\text{-}cDNA}$ allele, I asked whether Mef2 function was restored in this line. To do so, I looked at both the embryonic somatic musculature, and the adult DLMs, which both require proper Mef2 function for their formation (Bour et al. 1995; Lilly et al. 1995; Ranganayakulu et al. 1995; Soler et al. 2012). In an initial qualitative assessment, I observed a restoration of both the embryonic somatic muscle pattern in the late embryo, when compared to $Mef2^{\Delta ATG}$ (**Fig. 5.11**); and the DLM fibres (**Fig. 5.12**). Because *Mef2* null alleles are embryonic lethal, Mef2 knockdown in the developing DLM fibres was used as the control for the impact of Mef2 loss-of-function of DLM development (**Fig. 5.12**). These data suggest that Mef2 function is restored sufficiently to ensure both normal somatic myogenesis and viability in $Mef2^{\Delta ATG\text{-}cDNA}$. Of course, given the single isoform rescue, it remains unknown as to whether this allele is rescued to wild type, and whether any unknown, isoform specific functions of Mef2 exist.

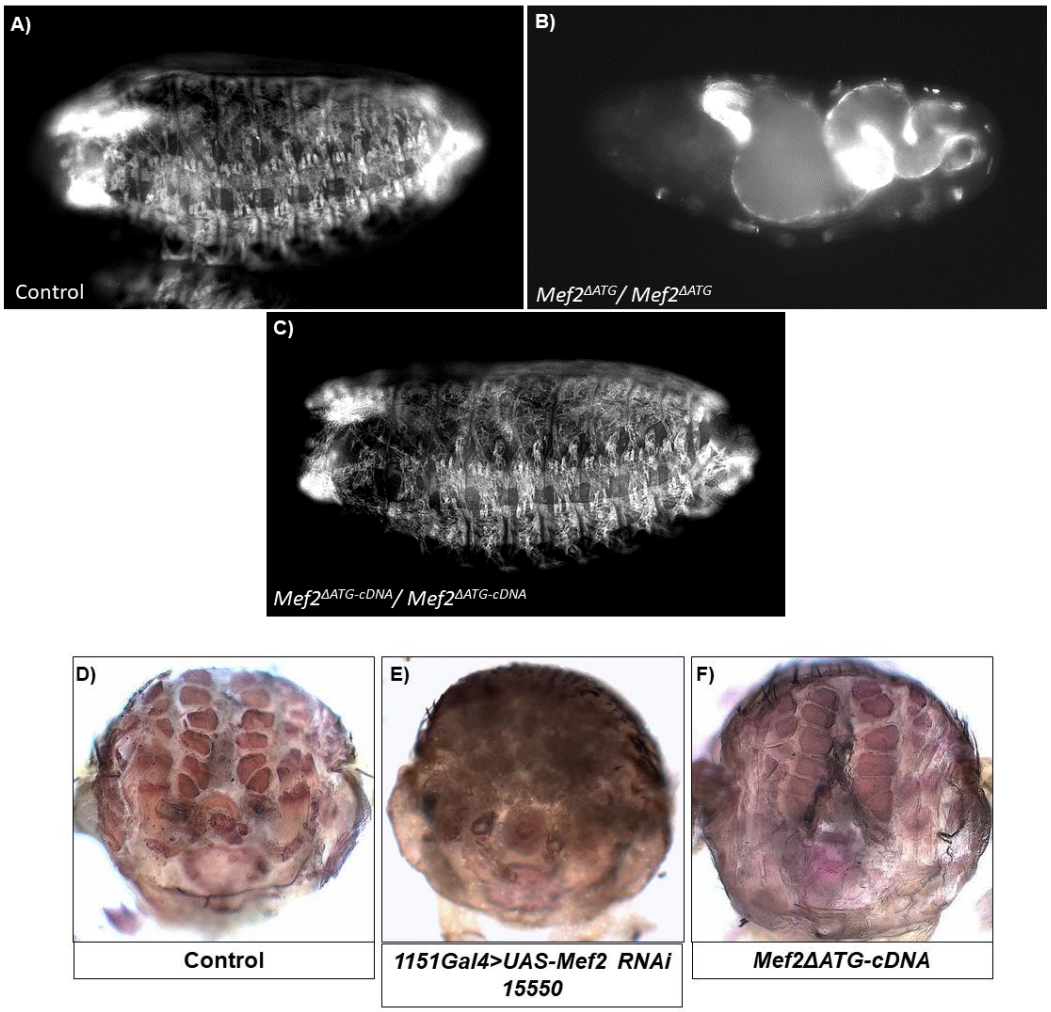


Figure 5. 11.Somatic muscle phenotype is restored in Mef2ΔATG-cDNA.
(A-E) Late-stage wild-type (**A**), *Mef2*$^{\Delta ATG}$ homozygous (**B**), or (**C**) *Mef2*$^{\Delta ATG\text{-}cDNA}$ homozygous embryos stained with an antibody against Myosin heavy chain. **(D-F)** Transverse thoracic cross sections stained with hematoxylin, showing DLM fibre structure in 96hr APF wild type controls (**D**), *Mef2* knockdown (**E**), or *Mef2*$^{\Delta ATG\text{-}cDNA}$ homozygous pupae. Mef2 loss-of-function inhibits formation of both the embryonic somatic musculature, and the adult DLM fibres. In *Mef2*$^{\Delta ATG\text{-}cDNA}$, somatic musculature appears normal.

5.7 Mef2-EGFP: Targeting the Mef2 locus to generate a C-terminal EGFP tag.

The next aim was to generate a novel, fluorescently tagged *Mef2* allele, which would enable furthering the understanding of Mef2 expression, localisation, and function during development *in vivo*. One application of this would be the ability to study Mef2 expression in the context of its possible regulation by HDAC4. The initial aim was to fluorescently tag both Mef2 and HDAC4 using the insertion-ready deletion allele approaches described here. However, for many reasons, it was decided to use an alternative CRISPR-based approach to generate these novel, tagged alleles.

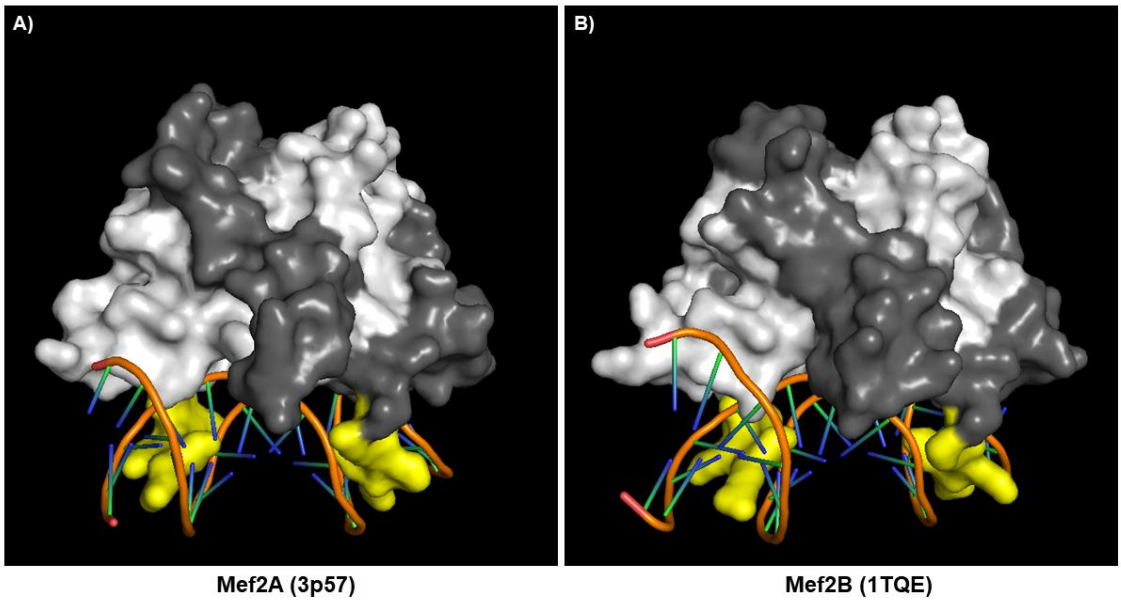


Figure 5. 12.The N-terminus of Mef2 is positioned closely to DNA.
Pymol renders of existing Mef2A (**A**) and Mef2B (**B**) crystal structures, each showing a Mef2 dimer (light grey, dark grey), bound to DNA. Structure reveals that the N-terminus (yellow) of Mef2 lies in very close proximity to DNA when bound. Highlighted N-terminus indicates amino acid residues 1-5 in the Mef2 sequence. PDB files 3P57 (Mef2A) and 1TQE (Mef2B).

For Mef2, the initial considerations were centred on at which terminus to add the tag: Mef2 has multiple annotated transcripts, which vary due to three annotated transcription start sites (TSS's) and alternative splicing (**Fig. 5.1A**) (Lin and Baines 2019; Gramates et al. 2022), although the specific functions of different Mef2 isoforms remains relatively poorly understood (Taylor and Hughes 2017). All isoforms share the same start codon, and the first two coding exons encode the MADS/Mef2 domain, which unsurprisingly, given its critical role in known *Mef2* function, is present in all

isoforms (**Fig. 5.1A**). Moreover, the same stop codon is also used by all transcripts apart from isoform H. This isoform is predicted to utilise a stop codon downstream of the canonical termination codon, however this is a prediction based on flybase annotation and has not been experimentally verified (Gramates et al. 2022). Thus, based on this evidence, either the N- or C-terminus could be adopted to tag all possible Mef2 isoforms. However, I decided to tag the C-terminus for two main reasons: Firstly, given the importance of the N-terminal region encoding the MADS/Mef2 domain, and crystal structures implying the N-terminus lies in close proximity to DNA, it is possible that adding a bulky fluorescent tag here could negatively impact Mef2 function (Han et al. 2005; He et al. 2011) (**Fig. 5.12**). Secondly, previously published Mef2 transgenes (UAS-lines) harbour a C-terminal 3xHA epitope tag that does not affect Mef2 function (Sandmann et al. 2006; Clark et al. 2013).

To tag the endogenous Mef2 locus, I used the pCFD3 gRNA-expressing vector and a HDR repair template, including homology arms flanking a cassette. This cassette contained the coding sequence for the EGFP fluorophore, and the same *Pax-Cherry* cassette used as a marker previously in the generation of insertion-ready deletion alleles (**Fig. 5.13A**). Similarly, this cassette is flanked by loxP sites and thus cre-mediated excision is possible if required. I first identified a suitable CRISPR site in the vicinity the stop codon, and verified its integrity in the *yw;;nos-Cas9* injection stock by PCR, followed by sequencing (**Fig. 5.13B**). Importantly, this CRISPR site was chosen to be as close as possible to the desired insertion point. This is because it is understood that the efficiency of CRISPR-mediated editing reduces with increasing distances between the Cas9-induced DSB, and the desired insertion location (Cong et al. 2013). Because the tag must be in-frame with the endogenous Mef2 CDS, correct positioning of the 5' homology arm was imperative. Thus, the 5' arm was designed to end just prior to the stop codon, such that the EGFP CDS (excluding the ATG-encoding Methionine) would be inserted in-frame, directly following the final amino-acid encoding codon in the Mef2-sequence (**Fig. 5.13B**). Because the 3' homology arm, starting immediately after the endogenous Mef2 stop codon, included 14/20 nucleotides of the protospacer and PAM, including the most proximal 'seed' sequence most important for defining gRNA specificity (Gratz et al. 2014; Ren et al. 2014), there was a minor concern that this could cause Cas9-induced cleavage of the HDR vector, albeit at a low efficiency. If this were to occur, this would compromise the HDR-repair

template and thus negatively affect the chances of generating the desired transgenic. To compensate for this, I incorporated a mutation in the 3' HA (by introducing a SNP in the primer used to PCR amplify this HA) to abolish the PAM (**Fig. 5.13C**). Prior to injection, the completed HDR vector was sequence verified to ensure correct HA positioning in relation to the desired insertion point of the construct (**Fig. 5.13C**). The 3' homology arm was designed to start immediately after the endogenous stop codon because the *EGFP* contained its own stop codon and thus the endogenous *Mef2* stop could be deleted.

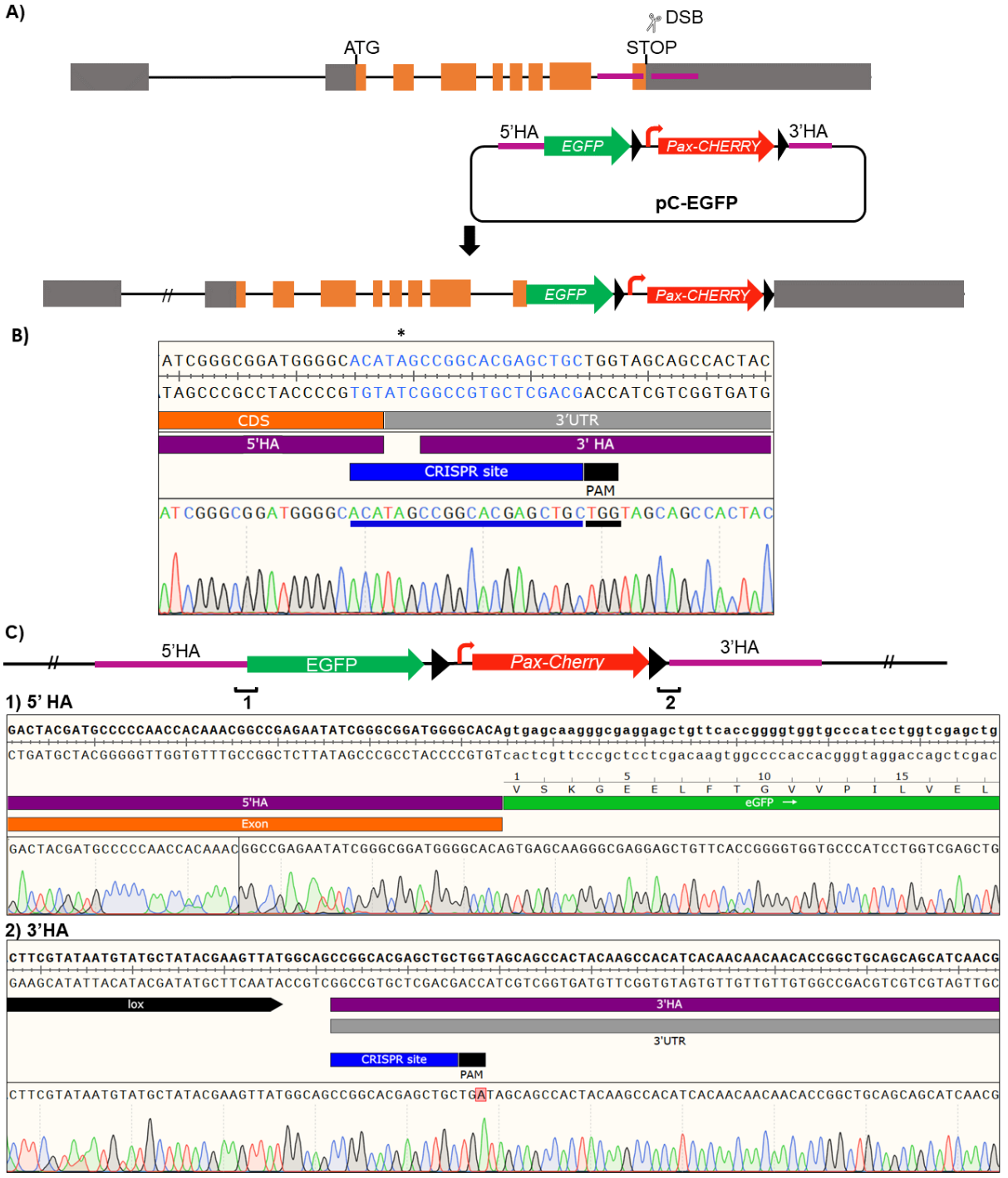


Figure 5. 13. Targeting the Mef2 locus to generate Mef2-EGFP.
(A) Targeting strategy to add a C-terminal EGFP tag to Mef2. The HDR targeting vector contains a cassette containing the coding sequence for EGFP, followed by a *Pax-cherry* marker cassette flanked by loxP sites. The HDR template is guided to the cas9-induced DSB (induced close to the STOP codon) by homology arms either side of the cassette. **(B) (Top)** CRISPR site location, and positioning of homology arms relative to the Mef2 stop codon (asterisk). (bottom) DNA sequence trace verifying the integrity of the CRISPR site in the injection fly line. **(C)** Sequence verification of the homology arm-cassette boundaries in the constructed HDR vector to be used for Mef2-EGFP generation. To avoid Cas9-mediated cleavage of the HDR vector, the 3' homology arm was mutated to remove the PAM.

5.8 GFP-tagging of Mef2 does not impede Mef2 function *in vivo*

Seven independent transgenic lines were received by BestGene. These were verified for the desired insertion through *Pax-Cherry* marker expression, the first important question was whether the tagging of Mef2 impacted its function. There were two possibilities here: Firstly, the EGFP could disrupt Mef2 as it is considerably larger than most epitope tags, and thus has a greater probably of causing a steric effect; or the insertion of the *Pax-cherry* cassette into the 3' UTR could potentially affect *Mef2* mRNA processing. However, I noted that flies harbouring *Mef2-EGFP* are homozygous viable, fertile, and maintain the ability to fly. This is true even for the transgenics still harbouring the *Pax-Cherry* cassette, suggesting disruption of the 3' UTR also is not sufficient to ablate Mef2 function. Because Mef2 function is essential for viability, it was concluded that GFP tagging with Mef2 does not have a significant effect on Mef2 function *in vivo*.

5.9 Testing Mef2-EGFP: Mef2 is detectable by GFP fluorescence throughout development.

Given one major advantage of a tagged Mef2 is the ability to detect Mef2 protein in living tissues without the need for fixation and immunostaining, different tissues were screened for the ability to detect Mef2-EGFP. To do so, I looked at both embryonic and adult muscle tissues, which have both been used throughout this project and are of interest within the field of muscle biology. Firstly, strong GFP signal is seen in mounted, dechorionated embryos, and pattern of GFP matches that of previously described Mef2 expression in the nuclei of differentiated embryonic somatic muscle (**Fig. 5.14A,C,D**), and the cardioblasts of the heart (**Fig. 5.14B)** (Bour et al. 1995). To see if this signal is detectable into larval life, first instar larvae were also mounted, and it was found that GFP signal is still easily detectable in the somatic muscle nuclei, even through the larval cuticle (**Fig. 5.14E**).

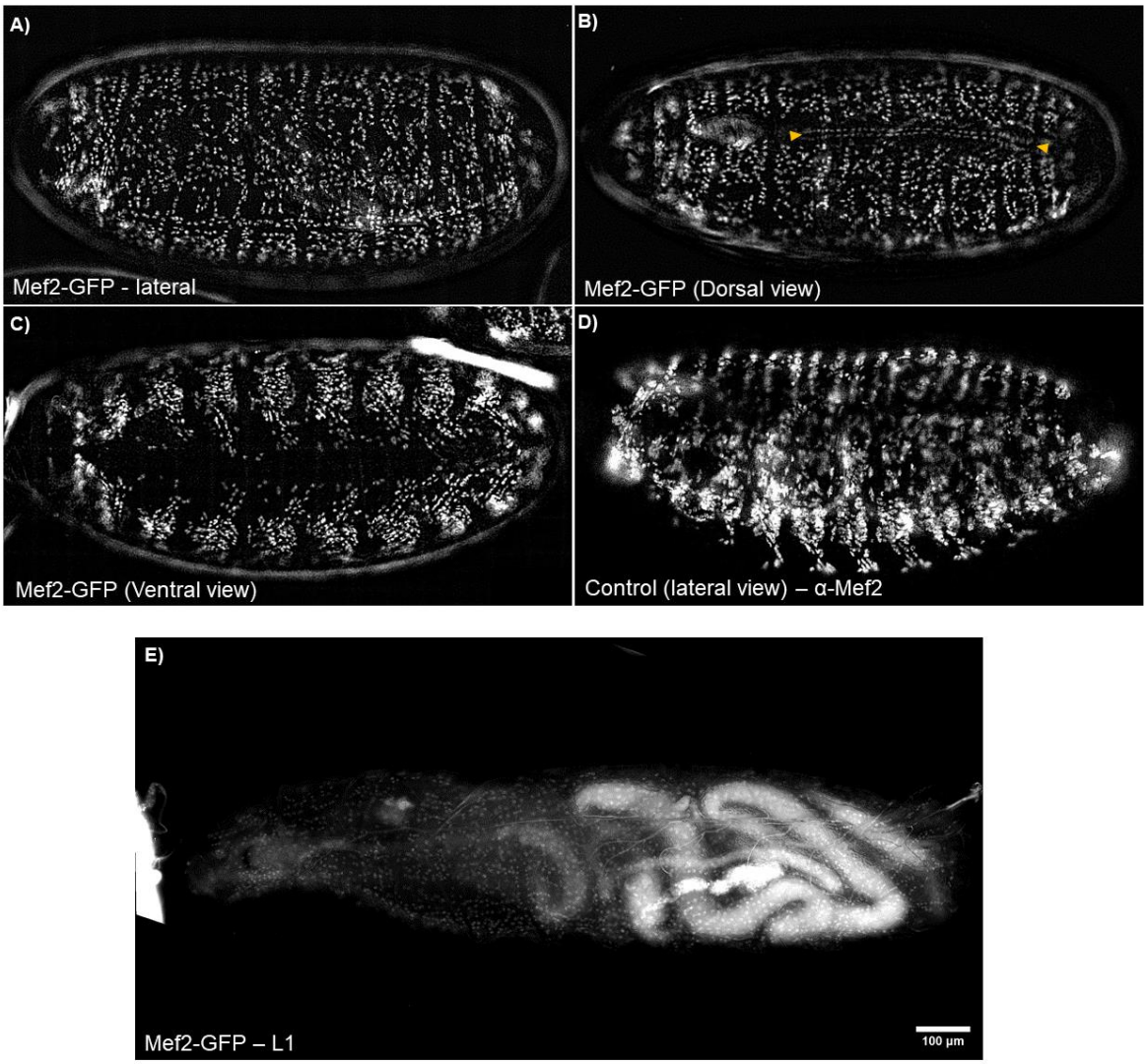


Figure 5. 14.Mef2-EGFP detection in larval somatic musculature.
(**A-C**) GFP signal in the nuclei of somatic muscle, and the cardioblasts of the heart in a late stage, dechorionated embryo. Cardioblasts are located between the orange arrows shown in **B**. **(D)** wild-type, late-stage embryo stained with an antibody against Mef2. (**E**) Live, first-instar larva expressing Mef2-EGFP. GFP signal is still readily detectable in the nuclei of the somatic muscles. Larval gut displays autofluorescence in the green channel.

Following this, I next investigated whether Mef2-EGFP could also be detected during adult myogenesis and, in line with overall project, investigated whether Mef2 could be detected in both undifferentiated L3 AMPs, and differentiated DLM fibres. In a similar fashion to the embryonic and larval musculature, Mef2-EGFP was again detectable in the AMPs of unfixed L3 wing imaginal discs (**Fig. 5.15A**), giving a comparable signal to using an antibody against Mef2 in control samples (**Fig. 5.15B**). In both the embryonic somatic musculature, and the L3 AMPs, it was noted that signal obtained

through Mef2-EGFP appeared more well defined and with less background than with antibody staining against Mef2 protein. Finally, I also attempted to detect Mef2 in differentiated DLM fibres. These dissections were carried out using Mef2-EGFP by Rob Mitchell, another PhD student whom I work closely with. Dissections were technically challenging and it was difficult to obtain in-tact sagittal sections of adult thoraces without fixation. However, following fixing, Mef2-EGFP was readily detectable in the nuclei of mature DLM fibres (**Fig. 5.15C**). Thus, Mef2-EGFP is a highly useful tool, which can be used to study Mef2 across a range of tissues and developmental periods in *Drosophila*.

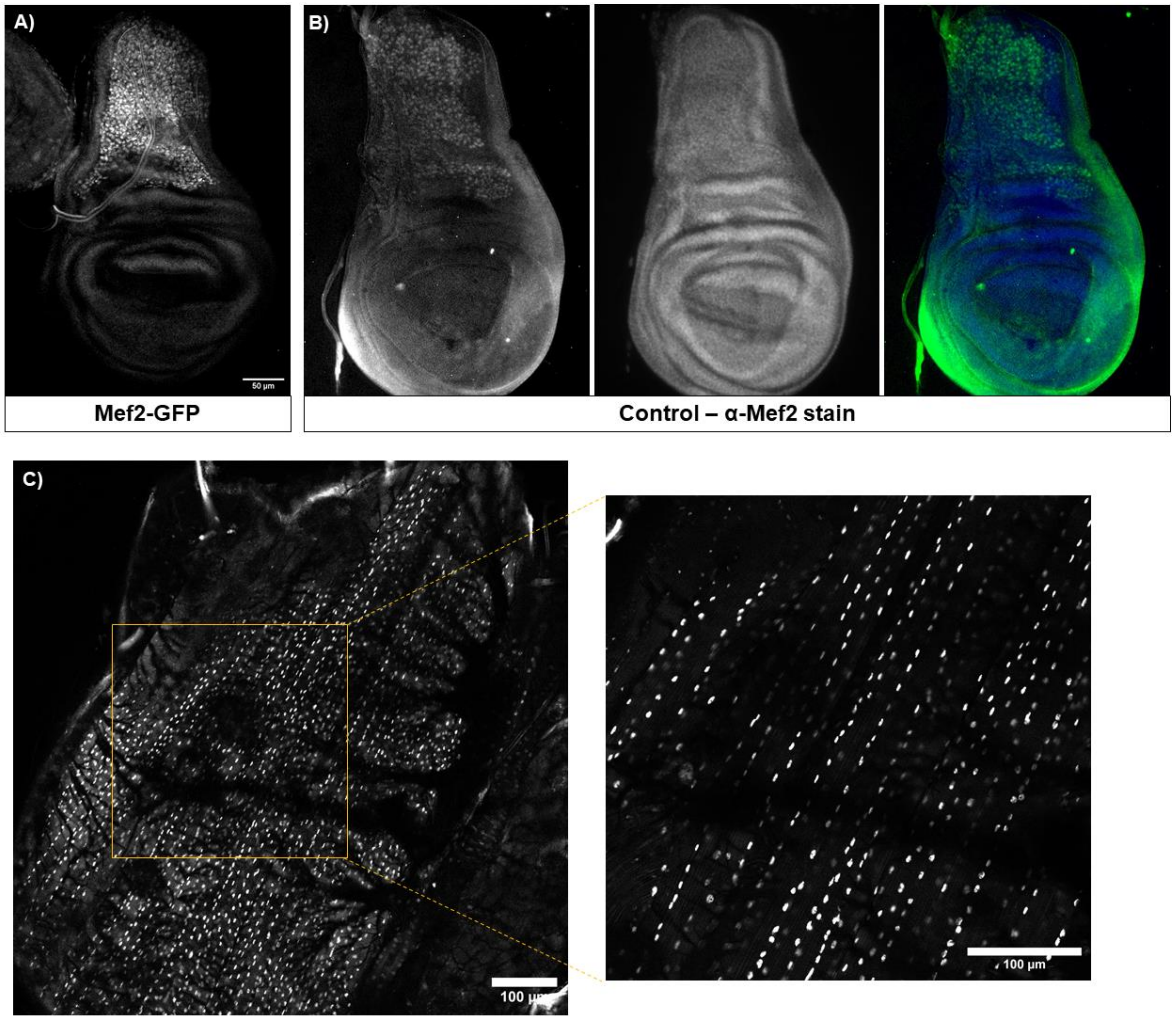


Figure 5. 15.Mef2-EGFP detection during adult muscle differentiation.
(A) Mef2-EGFP signal is detectable in live, unfixed AMPs associated with the wing imaginal disc of wandering L3 larva. **(B)** Antibody staining against Mef2 in the AMPs of control

wandering L3 larva. (**C**) confocal micrograph of a sagittal cross-section of a fixed thorax of a fly expressing Mef2-EGFP. Left = 10x, right =20x. Images for **B** and **C** acquired by Rob Mitchell

5.10 HDAC4-mScarlet: Targeting the *HDAC4* locus to generate an RFP-tagged HDAC4 for expression and localisation analysis

Following the successful generation of Mef2-EGFP, which will be a very useful tool in the future analysis and understanding of Mef2 function *in vivo*, I next wanted to generate tools to facilitate the analysis of HDAC4 expression and localisation *in vivo*. One reason for this is because the *in vitro* characterization of Class IIa HDAC function has implicated nucleo-cytoplasmic shuttling in regulating their function. Therefore, it was of great interest to analyse the subcellular distribution of endogenous HDAC4 during muscle differentiation *in vivo*.

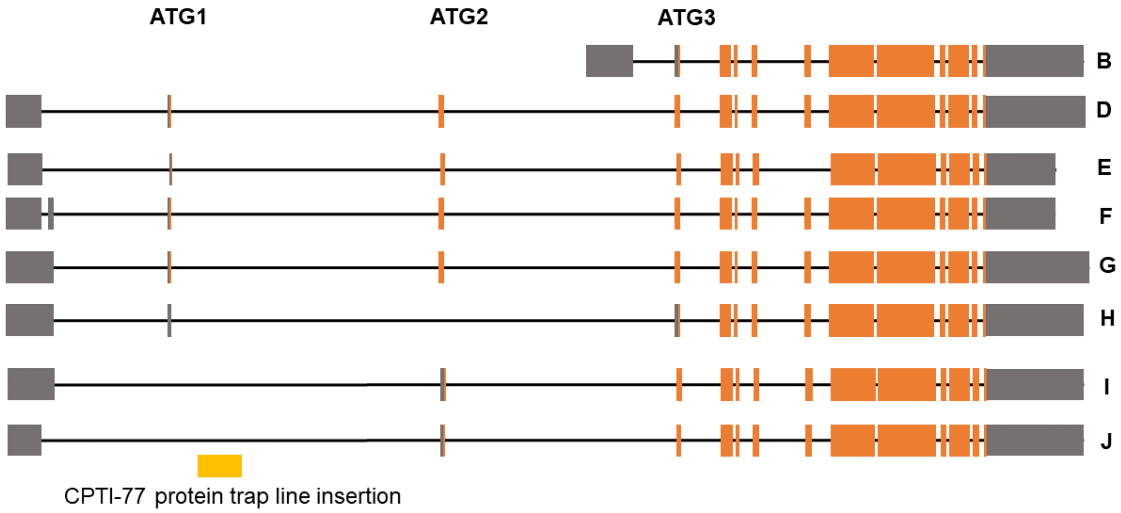


Figure 5. 16. HDAC4 isoforms and location of existing insertion for a tagged HDAC4. Schematic of the 8 annotated HDAC4 isoforms on flybase. Non-coding exonic sequences are in grey, coding sequences are in orange. There are 3 annotated, in-frame start codons used by either isoforms D,E,F,G (ATG1), I,J (ATG2), and B,H (ATG3). Genomic location of the CPTI-protein trap insertion is indicated, suggesting it tags isoforms that utilise ATG1.

The *HDAC4* locus is complex, with eight annotated transcripts and three annotated in-frame start codons (**Fig. 5.16**). Moreover, very little is understood with regardless to the abundance of individual transcripts in different tissues, or what the dominant isoform is. This is one of the considerations when designing these *HDAC4* alleles in terms of wanting to generate a tool that would tag the majority, if not all, expressed isoforms. By contrast to the variable N-terminal region, a common stop codon is utilised by all isoforms, which may suggest a C-terminal protein tag would be suitable

for HDAC4. However, unlike Mef2, there is limited information with regards to whether tagging either terminus of *HDAC4* could negatively affect its function. From this project, I knew that a C-terminal myc tag did not affect the ability for a HDAC4 to inhibit DLM formation (**See chapter 3**), although this tag is significantly smaller than a fluorescent protein. On the other hand, an existing HDAC4-YFP protein trap line, for which the fluorophore is inserted into the an the intron immediately following the first coding exon for isoforms D,E,F and G (**Fig. 5.16**), is homozygous viable (Lowe et al. 2014). However, although this CPTI line suggests N-terminal tagging does not have a severe detrimental impact on HDAC4 function *in vivo*, this line has limitations: Because of the complexity of the *HDAC4* locus, there lack of understanding as to whether it tags the majority of expressed HDAC4, and our knowledge it is not bright enough to be visualised without antibody staining and confocal microscopy either in the *Drosophila* brain (Fitzsimons et al. 2013), or in the wing imaginal disc (our lab, data not shown). Furthermore, given Mef2-EGFP, and our lab's particular interest in studying the functional relationship between HDAC4 and Mef2, the aim was to develop complimentary tools which would allow the analysis of both proteins' simultaneously and ideally with the ability to image in live tissues. Consequentially, this required HDAC4 to be tagged with a fluorophore with fluorescent characteristics suitably distant from GFP.

I opted to use mScarlet, a synthetic, bright RFP derived from mCherry and other RFP proteins (Bindels et al. 2016). This fluorophore was made readily available to us in an appropriate targeting vector with characteristics described in the generation of Mef2-EGFP (**section 6.7**), provided by a collaborator, who has also used it for protein tagging in *Drosophila* (Poernbacher et al. 2019). In addition to their suitability for colocalization analysis, GFP-RFP proteins are complimentary fluorophores for FRET-FLIM microscopy (**F**orster **r**esonance **e**nergy **t**ransfer measured by **f**luorescence **l**ifetime **im**aging). This technique allows the quantification of interactions occurring at <10nm through the non-radiative energy transfer between donor and acceptor fluorophores (Bajar et al. 2016), and in the long-term could be used to probe whether Mef2 and HDAC4 physically interact *in vivo*.

However, the mScarlet sequence in this HDR targeting vector corresponded to mScarlet-H. This is a variant of mScarlet with greater photostability, however lower brightness and longer maturation time than mScarlet. I used site-directed mutagenesis to change the mScarlet-H sequence to another variant, mScarlet-I, which is brighter than both mScarlet-H and wild-type mScarlet in cells (Bindels et al. 2016). This required the mutagenesis of two residues, and was done sequentially. Firstly, histidine-164-methionine (H164M), which restored mScarlet-H back to mScarlet; and threonine-74-isoleucine (T74I), which changed mScarlet into mScarlet-I. Following mutagenesis, the generated vector was fully sequenced to confirm the desired changes (**Fig. 5.17**). Because of the aforementioned reasons in terms of fluorophore placement, I designed and generated two constructs aimed to generate both an N-, and C-terminal HDAC4 fusion protein with the mScarlet-I fluorophore.

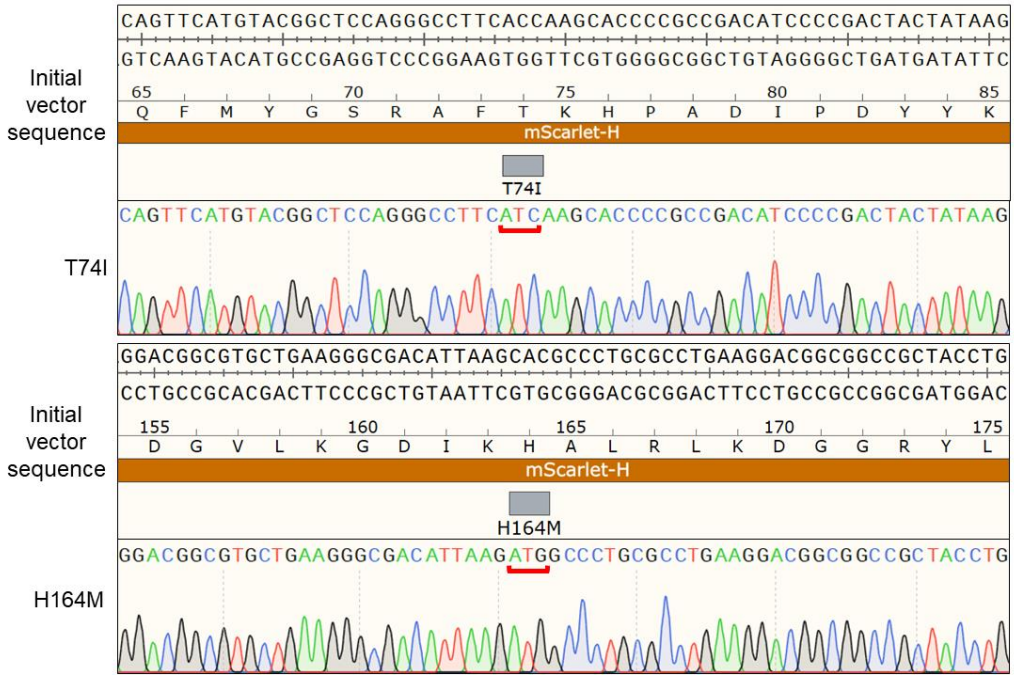


Figure 5. 17. Mutagenesis of mScarlet-H to mScarlet-I.
Initial vector sequence compared to DNA sequence chromatograms following (**A**) threonine-74-isoleucine and (**B**) histidine-164-methionine mutagenesis to convert the mScarlet-H coding sequence to mScarlet-I, within the HDR targeting vector used for direct tagging.

5.11 HDAC4 N-terminal tagging to generate a transcriptional reporter, mScarlet-I-HDAC4

To generate an N-terminally tagged HDAC4, I first targeted the most N-terminal start codon annotated to be used by isoforms D,E,F, and G, given the position of the insertion location for the previously published CPTI line (Fitzsimons et al. 2013; Lowe

et al. 2014) (**Fig. 5.18A**). As with Mef2-EGFP, the HDR-targeting vector shares some similarities: homology arms flanking a cassette containing the fluorophore (mScarlet-I in this case) and a *Pax-cherry* marker cassette flanked by loxP sites to enable cre-recombinase mediated excision (**Fig. 5.18A**). However, unlike the C-terminal targeting vector, there is also a sequence encoding the 3'UTR of the p10 baculovirus in the region bound by loxP sites (**Fig. 5.18A**). Thus, in the resulting transgenic, transcription from the endogenous TSS of the *HDAC4* locus transcribes through the mScarlet-I fluorophore before terminating, and thus interrupting transcription of the rest of the *HDAC4* locus. Because the *Pax-cherry* cassette is expressed independently, it would remain unaffected by this mechanism. Therefore, and unlike C-terminal tagging (where the entire endogenous locus is transcribed prior to the tag and the only disruption being to the 3'UTR), this insertion would first generate a *HDAC4* null allele and a transcriptional reporter (through expression of mScarlet-I). This allele would only then be converted into an in-frame, mScarlet-I-HDAC4 tagged allele following cre-mediated excision (**Fig. 5.18A**). Between the end of the tag, and the second codon of the *HDAC4* CDS is the residual loxP site which forms part of a short glycine-serine linker (**Fig. 5.18C**).

As with all CRISPR-cas9 targeting in this project, I first identified a suitable CRISPR site in the vicinity of the desired insertion site, and verified its integrity in the line to be injected (*yw;;nos-cas9(attP2*)) by PCR, followed by DNA sequencing (**Fig. 5.18B**). Furthermore, as with the importance of homology arm placement in Mef2-EGFP generation, the 5' arm was designed to incorporate the ATG at its end. Firstly, this would ensure that the endogenous ATG would be used in this tagged allele, as the mScarlet-I fluorophore within the HDR targeting vector has the start codon removed. Secondly, it was also ensure the an in-frame insertion of the mScarlet-I CDS (**Fig. 5.18C**) The 3'HA was designed to begin at the start of the second coding exon, such that no *HDAC4* CDS would be deleted through this tagging approach (**Fig. 5.18C**). The location of the LoxP site and G/S linker relative to the positioning of the 3'HA can also be visualised here (**Fig. 6.20C**). Similarly to the design of Mef2-EGFP, the 3' homology arm also included the full protospacer sequence, and the PAM, which would cause Cas9-mediated cleavage of the HDR-targeting vector and thus would impede transgenesis. Therefore, the 3'homology arm was mutated to disrupt the PAM site (**Fig. 5.18C**). It is important to note here that this would likely mutate the resulting

HDAC4 allele, and because this change would affect the *HDAC4 CDS*, it was ensured that this mutation would be silent and not affect the HDAC4 protein sequence: the CCC → CCA codon change would not affect translation into a proline residue.

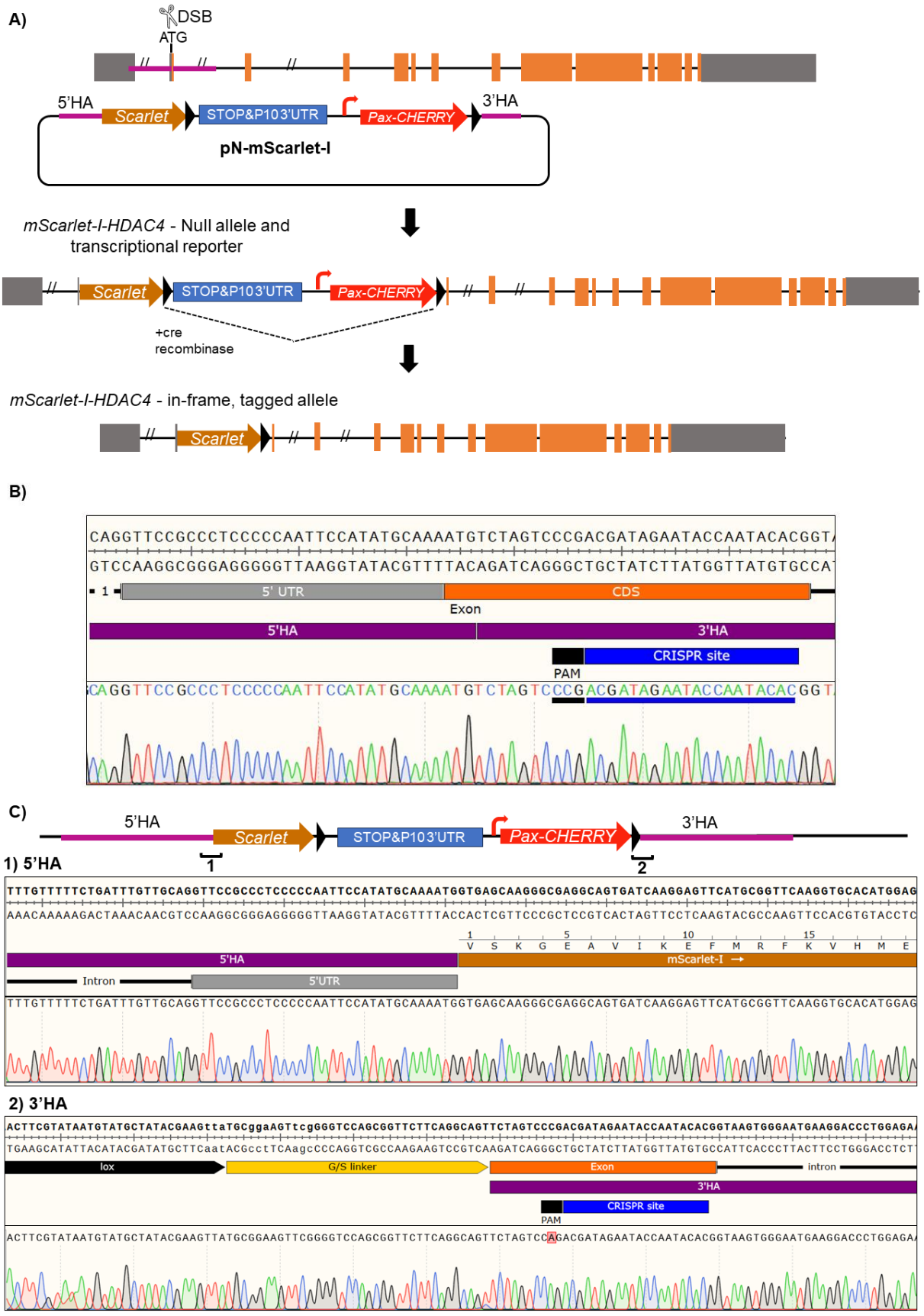


Figure 5. 18. Targeting the HDAC4 locus to generate an N-terminal tag and transcriptional reporter.

(A) Tagging strategy to generate mScarlet-I-HDAC4. isoform RD is shown. The HDR-targeting vector is guided to cas9-induced double-strand break (DSB) homology arms, which flank a cassette containing: the coding sequence for mScarlet-I, the 3' UTR from the p10 baculovirus, and a *Pax-cherry* marker. The p10 3'UTR and *Pax-cherry* is also flanked by loxP sites. Insertion of this cassette generates mScarlet-I tagged HDAC4 at the N-terminus, which should function as a null allele and transcriptional reporter, as transcription of the p10 3'UTR causes premature transcription termination. Cherry marker expression is regulated independently and therefore unaffected. Cre-mediated excision of the p10 3'UTR and *pax-cherry* marker converts the transcriptional reporter into an N-terminally tagged *HDAC4* allele. **(B) (Top)** CRISPR site location, and positioning of homology arms relative to the HDAC4 start codon (asterisk). (bottom) DNA sequence trace verifying the integrity of the CRISPR site in the injection fly line. **(C)** Sequence verification of the homology arm-cassette boundaries in the constructed HDR vector to be used for mScarlet-I-HDAC4 generation. To avoid Cas9-mediated cleavage of the HDR vector, the 3' homology arm was mutated to remove the PAM.

5.12 Targeting the *HDAC4* locus to generate a C-terminally tagged allele, HDAC4-mScarlet-I

In contrast to the generation of an N-terminally tagged *HDAC4* allele, the vector used to generate a HDAC4-mScarlet-I C-terminal fusion protein was identical to that of generating Mef2-EGFP. The only difference in targeting vector was the presence of the CDS for mScarlet-I as opposed to EGFP (**Fig. 5.19A**). However, I needed to synthesise this new vector to enable C-terminal tagging with mScarlet-I. To do so, I amplified the the mScarlet-I CDS by PCR, and ligated it into the C-terminal tagging vector used in the generation of Mef2-EGFP (with EGFP CDS removed). Despite this difference, the overall approach for generating C-terminally tagging HDAC4 with this approach remained the same as Mef2-EGFP (**Fig. 5.19A**): I first selected a suitable CRISPR site in the vicinity of the stop codon (in *yw;;nos-cas9(attP2)*), and verified its integrity by PCR followed by DNA sequencing (**Fig. 5.19B**). As previously described, this stop codon is common to all *HDAC4* isoforms and thus in theory would tag all HDAC4 expressed. The 5'HA and 3'HA were designed to end just before, and immediately after the stop codon respectively, such that the mScarlet-I tag would be inserted in-frame, before the endogenous stop codon (**Fig. 5.19C**). However, unlike Mef2-EGFP, and mScarlet-I-HDAC4, where the 3' HA was mutated to prevent cas9-mediated cleavage of the HDR construct, this was not necessary in this design. This was because the location of the CRISPR site in relation to homology arm positioning meant that the PAM was only followed by 2bp at the end of the 5' arm, which would not be sufficient to induce cas9-mediated cleavage (**Fig. 5.19B**)

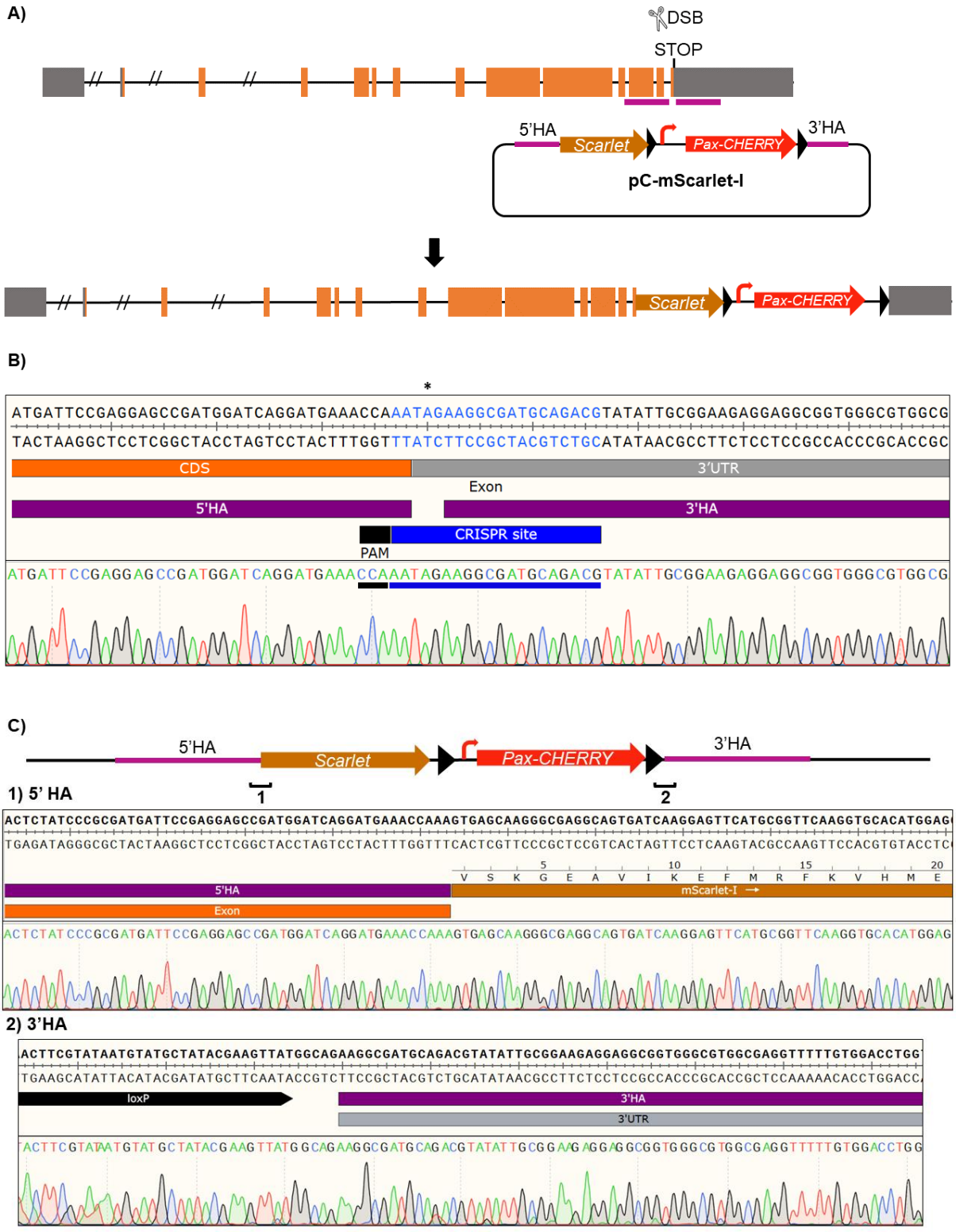
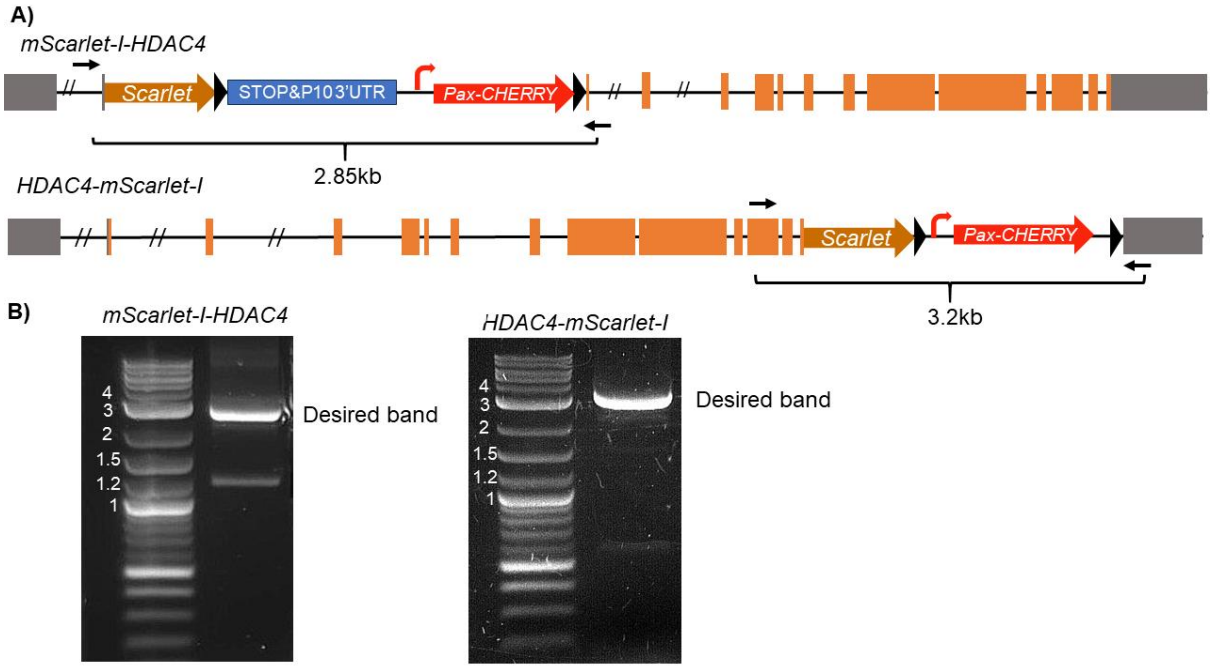


Figure 5. 19. Targeting the HDAC locus to generate a C-terminal fusion protein.
(A) Targeting strategy to add a C-terminal mScarlet-I tag to HDAC4. isoform D is shown. The HDR targeting vector contains a cassette containing the coding sequence for mScarlet-I, followed by a *Pax-cherry* marker cassette flanked by LoxP sites. The HDR template is guided to the cas9-induced DSB (induced close to the STOP codon) by homology arms either side of the cassette. **(B) (Top)** CRISPR site location, and positioning of homology arms relative to the *HDAC4* stop codon (asterisk). (bottom) DNA sequence trace verifying the integrity of the

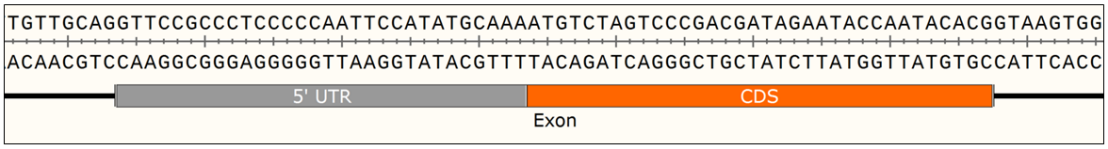
CRISPR site in the injection fly line. **(C)** Sequence verification of the homology arm-cassette boundaries in the constructed HDR vector to be used for HDAC4-mScarlet-I generation. To avoid Cas9-mediated cleavage of the HDR vector, the 3' homology arm was mutated to remove the PAM.

5.13 Verification of the *HDAC4* tagged alleles.

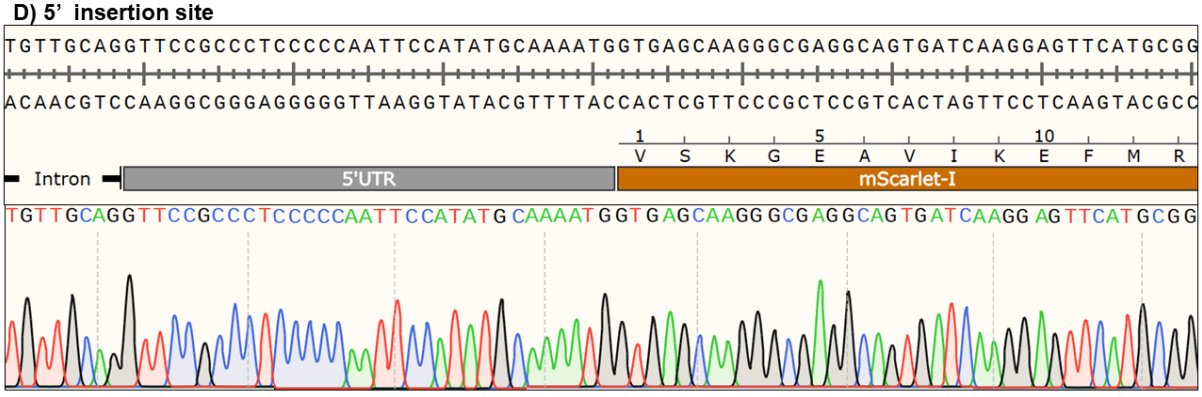
Unlike Mef2, where the expression in some tissues is well-characterised, which made verification possible through visualising it in known Mef2-expressing tissues, little is known of the expression pattern and localisation of HDAC4. Thus, to verify the correct insertion of fluorophores following HDAC4 tagging, I amplified the targeted regions by PCR, using gDNA isolated from transgenic flies (one line per genotype), and primers flanking the desired insertion sites (**Fig. 5.20**). For both mScarlet-I-HDAC4 and HDAC4-mScarlet-I, the expected fragment sizes were approximately 2.85kb, and 3.2kb respectively (**Fig. 5.20A)**. PCR amplification yielded products of approximately the expected sizes for both mScarlet-I-HDAC4 and HDAC4-mScarlet-I (**Fig. 5.20B**), although band separation of the high MW markers was not optimal. This meant band sizes were not accurately determined through visualisation of the gel products alone. In both PCR reactions, some off-target amplification at approximately 1.3kb and 600bp during the mScarlet-I-HDAC4 and HDAC4-mScarlet-I PCR reactions respectively, was observed (**Fig. 5.20B**). On-target PCR products were then sequence verified from either end to further characterise the nature of the novel HDAC4 tagged alleles. Due to the insertion sizes, the entire insertion was not fully sequenced, however for both mScarlet-I-HDAC4 (**Fig. 5.20C-E**) and HDAC4-mScarlet-I (**Fig. 5.20F-H**), it was validated that the HDR cassette was correctly inserted such that the mScarlet-I fluorophore was inserted in-frame with the endogenous *HDAC4* CDS, with the insertion boundaries defined by homology arm positioning.



C) Endogenous locus – N-terminus



D) 5' insertion site



E) 3' insertion site

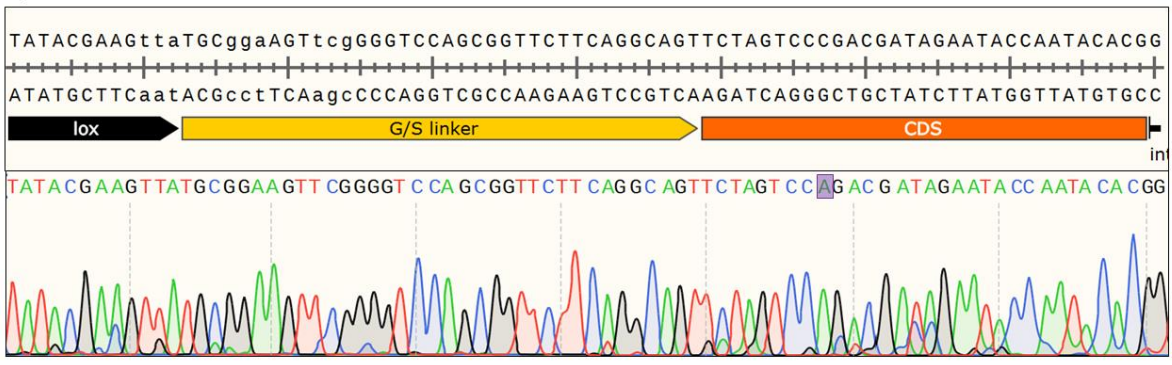


Figure 5. 20. Molecular verification of mScarlet-I-HDAC4 and HDAC4-mScarlet-I.
(A) Schematic of the mScarlet-I-HDAC4 and HDAC4-mScarlet-I alleles and predicted PCR product sizes indicated using primers either side of the engineered regions (black arrows). **(B)** Agarose gel of PCR products generated following amplification of mScarlet-I-HDAC4 (left) and HDAC4-mScarlet-I (right) using primers described in (A). **(C-H)** verification of mScarlet-I-HDAC4 **(C-E)** and HDAC4-mScarlet-I (**F-H**) transgenic lines, through DNA sequencing of PCR-products in **(B).** (**C**) Endogenous DNA sequence for the ATG-containing exon prior to engineering. (**D)** schematic (top) and DNA sequence trace (bottom) of 5' insertion site, showing in-frame insertion of mScarlet-I CDS immediately following the endogenous ATG. **(E)** schematic (top) and DNA sequence trace (bottom) of 3' insertion site, showing the end of the insertion marked by the G/S linker is followed immediately by the second coding exon. Incorporation of silent mutation in homology arm has been integrated into resulting transgenic line, as indicated. (**F**) Endogenous DNA sequence for the final coding exon of *HDAC4,* surrounding the endogenous stop codon. **(G)** schematic (top) and DNA sequence trace (bottom) of 5' insertion site boundary, showing incorporation of the mScarlet-I coding sequence in-frame, before the stop codon. **(H)** schematic (top) and DNA sequence trace of 3' insertion site boundary, showing the end of the HDR cassette is immediately followed by the endogenous 3' UTR, starting at the first base following the endogenous stop codon.

Interestingly however, genetic analysis of the mScarlet-I-HDAC4 allele suggests it is not functioning as initially designed. While the insertion has been mapped to its designed insertion point (**Fig. 5.20C-E**), the mScarlet-I-HDAC4 fly lines are homozygous viable and fertile. Given that my novel, rescuable att CRISPR alleles, show that HDAC4 loss-of-function is lethal (**Chapter 4**), this indicates that this mScarlet-I-HDAC4 is not functioning as a loss-of-function allele. There are multiple explanations for this, however they have not yet been investigated. Given the complexity of the *HDAC4* locus and the presence of multiple possible transcription start sites and in-frame start codons, it could be that the genomic insult caused by insertion of the tagging cassette could be mitigated through utilisation of a different TSS. It could alternatively be that the isoforms utilising this start codon are not the dominant isoforms, or at least the dominant isoforms required for function. This could therefore also mean that signal obtained through analysis of the CPTI line may not give a full picture of HDAC4 expression. Finally, and similar to that observed with Mef2-EGFP, disruption to the endogenous 3'UTR through retention of the *Pax-cherry* cassette did not appear to have a significant negative impact on HDAC4 expression or function *in vivo*, given the line is also homozygous viable and fertile.

5.14 Initial analysis of *HDAC4* in the wing imaginal disc

To next test the ability to detect HDAC4, I chose to investigate its expression and localisation in the wing imaginal disc. *HDAC4* is known to be expressed in the tissue, based on previous RNA-seq datasets (Spletter et al. 2018; Zappia et al. 2020). Furthermore, the presence of Mef2-expressing, undifferentiated AMP cells makes this an interesting location to investigate the possible role of HDAC4 in regulating Mef2 activity during muscle differentiation. Furthermore, due to its accessibility, and the ability to detect Mef2 in live, unfixed samples, it was also a suitable starting point in the aim of detecting endogenous HDAC4. Because of time constraints, I limited this analysis to HDAC4-mScarlet-I alongside the CPTI line due to the initial observations identified that mScarlet-I-HDAC4 is homozygous viable.

To do so, wing discs were first dissected and mounted in the absence of fixation, and analysed to see if HDAC4 could be detected using the same methodology as with Mef2-EGFP. However unfortunately, no signal was detected above background either using the CPTI, or the HDAC4-mScarlet-I lines (Data not shown). This was true even when confocal microscopy was used. Instead, antibody staining was used to detect

HDAC4 protein. For the CPTI and mScarlet-I lines, I used anti-GFP, and anti-RFP antibodies respectively, since there also is no available antibody against *Drosophila* HDAC4. Given the presence of multiple cell types in the wing imaginal disc, I used an antibody against Mef2 alongside either GFP (for CPTI) or RFP (for HDAC4-mScarlet-I) to investigate HDAC4 expression within the AMPs, and its localisation in relation to nuclear Mef2 protein. Unfortunately, because HDAC4 could not be detected in the absence of antibody staining, there was no immediate benefit of generating a fly line containing both tagged HDAC4, and Mef2.

I observed inconsistent staining results when comparing the CPTI line with the HDAC4-mScarlet-I line (**Fig. 5.21**). In the CPTI-line, HDAC4 displayed a honeycomb-like appearance in cells not expressing Mef2, which are the epithelial cells of the wing-disc proper (**Fig. 5.21A-B**). By contrast in the Mef2-expressing AMPs, HDAC4 was more diffuse and did not appear to be either dominantly nuclear or cytoplasmic (**Fig. 5.21A-B**). By contrast when HDAC4-mScarlet-I is analysed, although the characteristic honeycomb pattern in the epithelial cells is seen, the signal detected from the AMP population is more variable much lower signal relative to the epithelial cells is detected (**Fig. 5.21C-D**).

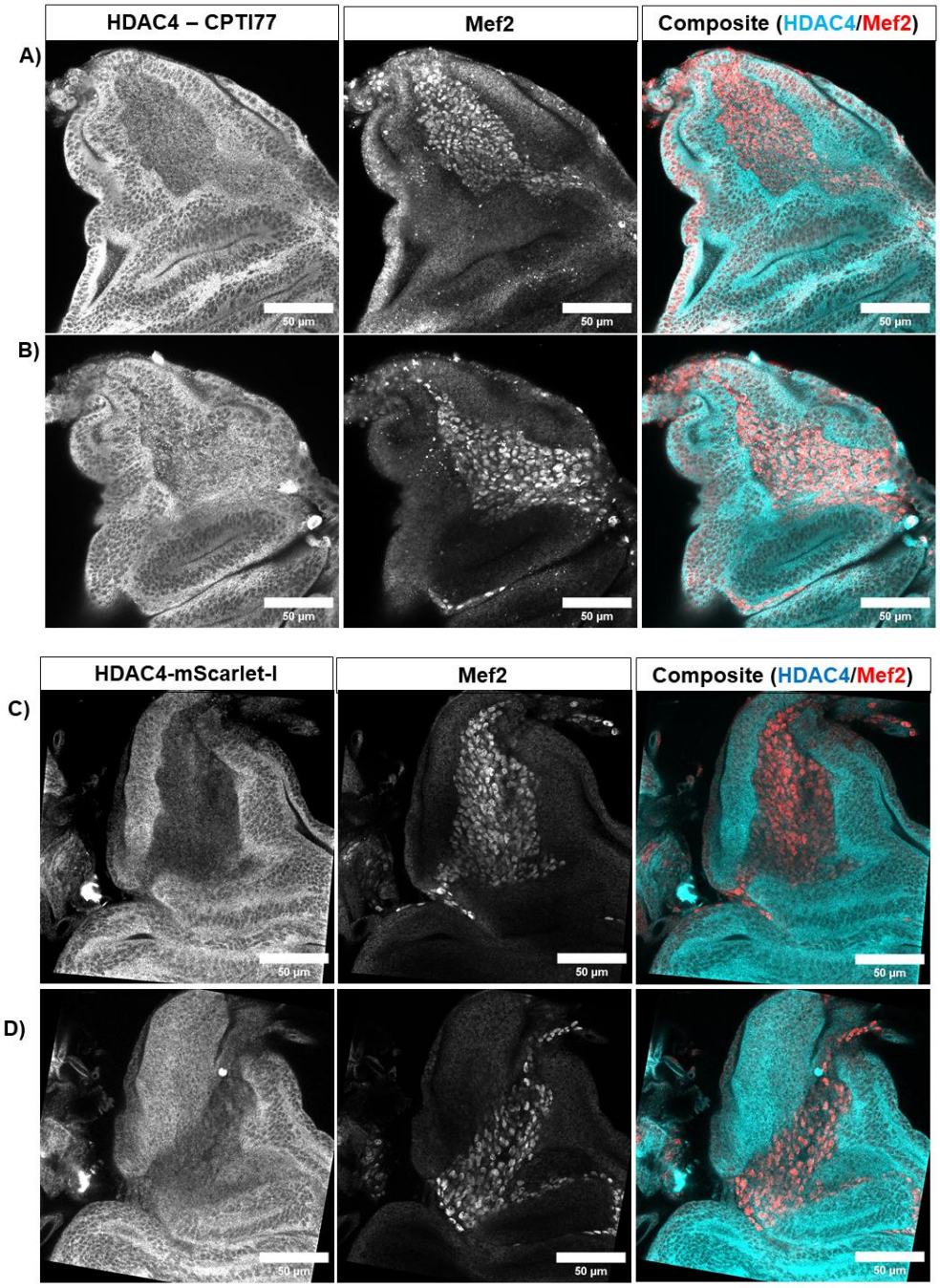


Figure 5. 21 . Co-staining HDAC4 tagged alleles with Mef2

Confocal micrographs of L3 wing imaginal disc (notum is shown) of either the CPTI77 HDAC4 protein trap line (**A-B**), or HDAC4-mScarlet-I (**C-D**). Wing discs are stained against Mef2 and HDAC4 using either a GFP (CPTI77) or RFP (HDAC4-mScarlet-I) antibody. **A-B** and **C-D** are two different Z-planes (1µm) of the same wing disc sample.

However, this signal was also variable depending on the sample (**Fig. 5.22**). For example in some wing discs, HDAC4-mScarlet-I expression was detected as described (**Fig. 5.22A**), in others its expression was more like that observed when analysed using the CPTI-line, displaying a distinctive honeycomb appearance in the epithelial cells, but more pan-cellular in the Mef2-expressing AMP population (**Fig. disc A**). This could possibly be a result of the binding affinity of the RFP antibody used, coupled with the expression level of HDAC4. Firstly, HDAC4 could be relatively lowly expressed in the AMPs especially, which could also explain why it could not be detected in the absence of antibody staining. Secondly, as the RFP antibody is not mScarlet-I specific, it may bind less well than the GFP antibody does to the YFP fluorophore found in the CPTI line. These two factors could lead to variable signal detection using the HDAC4-mScarlet-I line. Nevertheless in any case, the comparable signal detected between C-terminally tagged HDAC4 (which should tag all isoforms) and the CPTI line (restricted to isoforms D,E,F,G), especially in the epithelial cells, suggests the latter could be a reliable tool for detecting HDAC4 protein *in vivo*.

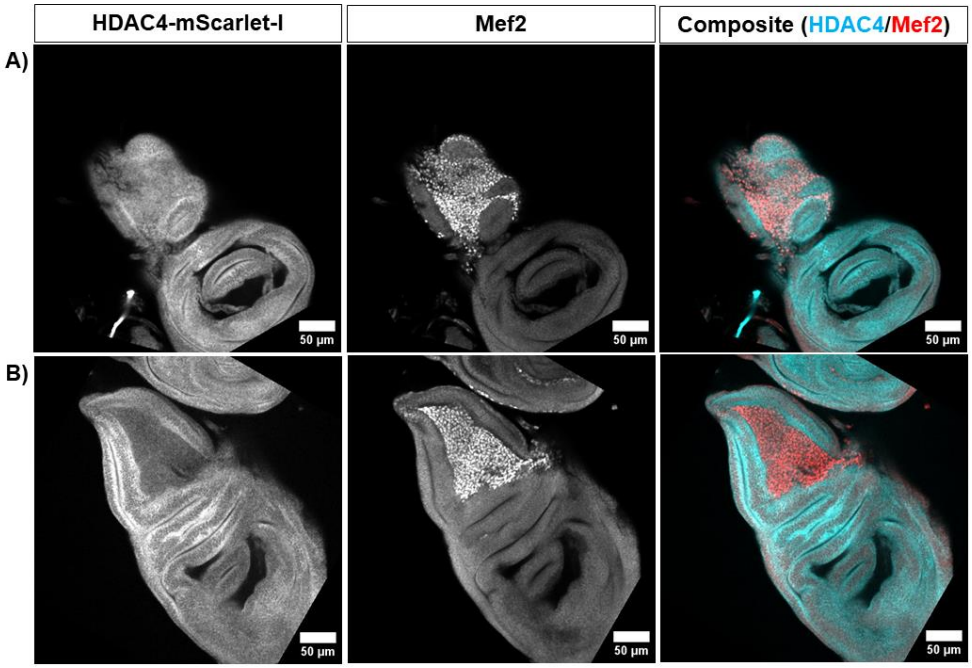


Figure 5. 22. Variable staining consistency against HDAC4-mScarlet-I in L3 wing imaginal discs.

Confocal micrograph of L3 wing imaginal discs expressing HDAC4-mScarlet-I, stained against HDAC4 (RFP), and Mef2. Single Z-planes (1µm) of two different wing discs (**A,B**) are shown. Samples display variable staining in the AMPs of the wing disc.

Results summary

In this chapter, I have successfully established two, novel *Mef2* insertion-ready deletion alleles which phenocopy that of existing *Mef2* null phenotypes. Notably, with $Mef2^{\Delta ATG}$, I show that deletion of only 54bp of CDS, encoding the 18 most N-terminal amino acids included in the MADS/Mef2 domain, is sufficient to completely abrogate Mef2 function, thus emphasising the importance of the MADS/Mef2 domain to Mef2 function. Interestingly however, and by contrast to the relative success of rescuing the *HDAC4* null allele in chapter 4, the *Mef2* alleles described here were less reliably rescued through the site specific recombination of either cDNA or gDNA into the engineered locus. This emphasises how the complex nature of gene structures makes them particularly vulnerable even carefully designed engineering strategies, indicating that the success of the insertion-ready deletion strategy may vary significantly depending on the target locus. I have also successfully generated fluorescently tagged *HDAC4 and Mef2* alleles. Mef2 was tagged with a C-terminal EGFP, which gives an incredibly bright signal in live, unfixed tissues at multiple stages of development. This enables the detection of Mef2 without the need for antibody staining and enables the use of live imaging techniques to study Mef2. I also generated *HDAC4* alleles tagged at either the N- or C-terminus with mScarlet-I, a bright red fluorescent protein, and used these alleles to attempt to study HDAC4 protein expression and localisation *in vivo*. However, based on initial analysis in the wing imaginal disc, *HDAC4* did not appear to be highly expressed in the AMPs, nor was there any obvious subcellular distribution observable. This may suggest, however pending further investigation, that endogenous HDAC4 in these muscle progenitor cells does not function analogous to that observed in *in vitro* models of muscle differentiation

Chapter 6: Identification of critical residues for Mef2 function *in vivo*, and initial investigation into the role of the p300 homologue *nejire* (*nej*) during IFM development

Introduction

As introduced in **1.4,** The highly conserved MADS/Mef2 domain mediates many functions necessary for Mef2 function, including DNA binding, dimerization, and transcriptional activation of target genes. While some residues have been identified as being required for DNA-binding and dimerization, the role of others is less well-characterised. For example, a region in the Mef2 domain is not necessary for Mef2C DNA-binding or dimerization, however is required for Mef2-dependent transcription *in vitro*. Therefore, It may be this region facilitates Mef2-protein interactions when bound to DNA, which is subsequently required for its ability to activate the expression of its target genes. Interestingly, the same Mef2 domain is responsible for the physical interaction of Mef2 with Class IIa HDAC corepressors, and has also been found to interact with p300/CBP histone acetyltransferase coactivators. However, given in **chapter 3** I have shown HDAC4 to be a Mef2 corepressor, it is unlikely that disruption of this interaction alone would abolish Mef2 function, meaning disrupting the binding of a transcriptional coactivator would more likely be the cause of this phenotype, while also possibly indicating a common binding site for proteins which differentially impact Mef2 function *in vivo*. Furthermore, none of these aforementioned critical residues have been formally tested in an *in vivo* model for Mef2 function. In this chapter, I use the ability for Mef2 to induce the premature differentiation of larval wing disc-associated AMPs, described in **chapter 3**, as an effective *in vivo* model for both Mef2 function and muscle differentiation. I then investigate the potential role of, Nejire, the *Drosophila* p300/CBP homologue, during *Drosophila* muscle differentiation *in vivo*. The research aims of this chapter can be described below:

1. Identify residues within the conserved MADS/Mef2 domain which are required for Mef2 transcriptional activity *in vivo*
2. Investigate the role of Nejire, the *Drosophila* p300/CBP orthologue, during muscle differentiation *in vivo* and whether it interacts with Mef2

6.1 Conserved residues within the MADS/Mef2 domain of Mef2 are implicated in cofactor recruitment

As previously described, the MADS/Mef2 domain is highly conserved between vertebrate and *Drosophila* Mef2 proteins (**Fig. 6.1A**). This is consistent with the many functions of this domain required for overall Mef2 activity, which include DNA-binding, dimerization, and transcriptional activation of target genes (Black and Olson 1998;

Potthoff and Olson 2007). A previous mutagenesis screen of vertebrate Mef2C identified many residues important for different aspects of Mef2C function (Molkentin et al. 1996a). Interestingly, while some mutants prevent Mef2C from dimerizing and/or binding DNA, others retain these functions but are unable to activate the transcription of a Mef2-dependent reporter. Specifically, a triple mutation valine-65, leucine-66, and leucine 67 caused this phenotype (Molkentin et al. 1996a) (**Fig. 6.1A**). One interpretation is that this region may interact with other proteins to required for Mef2 transcriptional activity. Indeed, this region of Mef2 has been shown to interact with both Class IIa HDACs (Han et al. 2005; Jayathilaka et al. 2012), and the histone acetyltransferase (HAT) p300 (He et al. 2011) via a conserved mechanism. The crystal structures show that residues 65-70 form a pocket in which the side chain orientations of these residues facilitate protein-protein interactions between Mef2 and either Class IIa HDACs, or p300. (**Fig. 6.1B-G**). This suggests that this domain may be important for interacting with other proteins, which may either positively, or negatively regulate Mef2 function. However, there has been no investigation into the functional significance of this region of the Mef2 protein during muscle differentiation *in vivo*. Although no structure of *Drosophila* Mef2 is available, the high degree of both sequence conservation in the MADS/Mef2 domain between Mef2 homologous, and the overall structural conservation observed in the crystal structures of Mef2A (**Fig. 6.1B-D)** and Mef2B (**Fig. 6.1E-G**) make it likely that *Drosophila* Mef2 adopts a similar structural organisation. Thus, *Drosophila* is an appropriate model for Mef2 structure-function analysis.

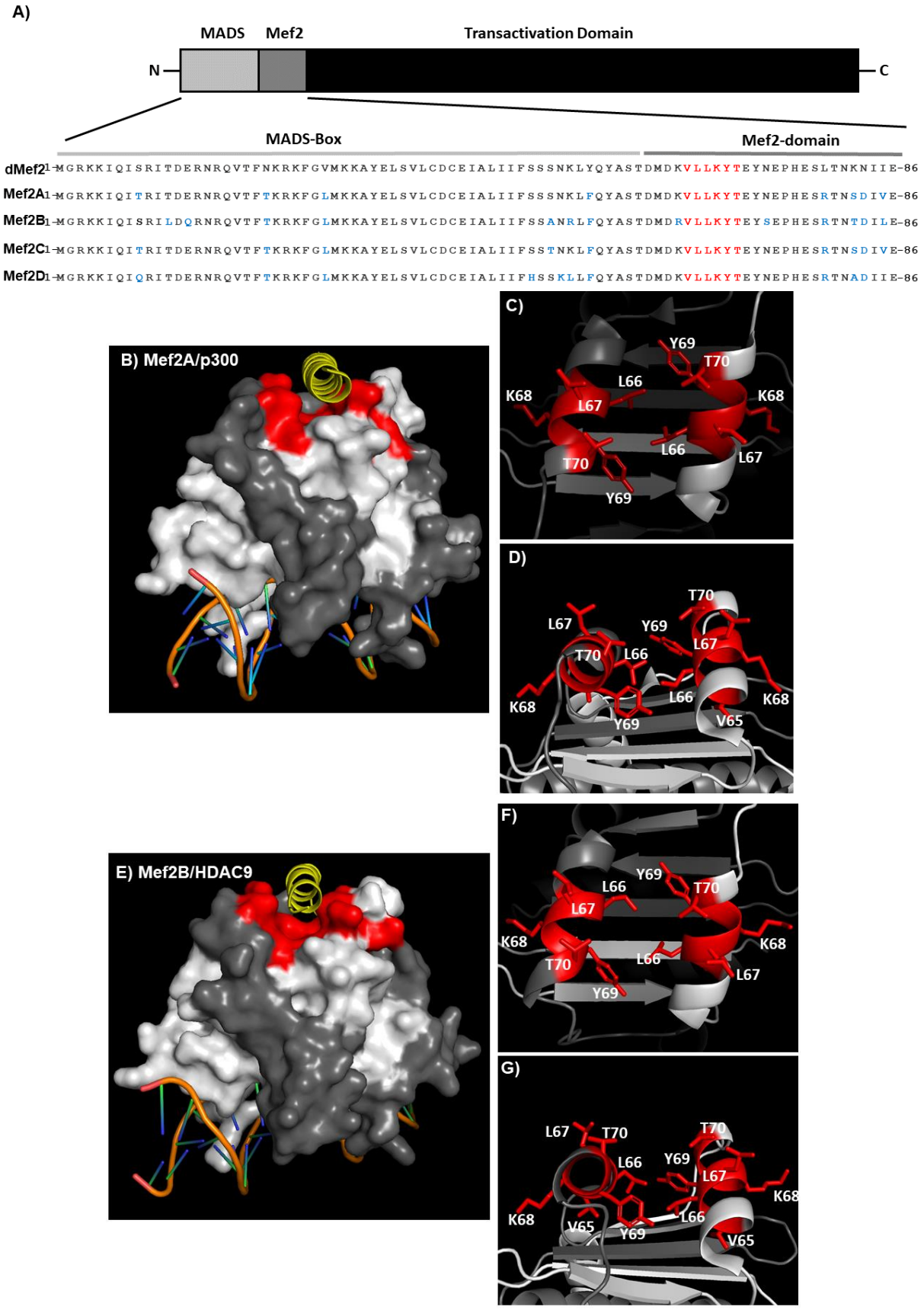


Figure 6. 1. Conserved residues implicated in regulating Mef2 function and cofactor interactions.

(A) MADS/Mef2 domain protein alignment of *Drosophila* Mef2 and the human Mef2A-D sequences. Conserved residues are in black, variable residues in blue (relative to dMef2). A region implicated to be involved in Mef2 function *in vitro* is shown in red. (**B-G**) Pymol renders of a Mef2A/p300 (PDB file: 3p57) (**B-D**) and Mef2B/HDAC9 (PDB file: 1TQE) **(E-G)** protein complexes bound to DNA, solved by X-ray crystallography. In both structures, only the MADS/Mef2 domain of Mef2 has been crystallised, along with only partial p300/HDAC9 fragments. Mef2 dimer is shown in greys, with the region described in (**A**) shown in red. P300/HDAC9 are shown in yellow. **(B,E)** The overall structure of the MADS/Mef2 domain of both Mef2 proteins is highly conserved, with the residues highlighted in (**A)** shown to be form a pocket which interacts with both p300 and HDAC9. **(C,G,F,E**) detailed structure of the region implicated in Mef2 function and cofactor interactions in Mef2A (**C,D**) and Mef2B **(F,G).** Side chains are shown as sticks.

6.2 Generation of mutants to target the MADS/Mef2 domain of Mef2

To study the effect of disrupting this region on Mef2 function, I generated a series of UAS-Mef2 constructs via site-directed mutagenesis. I opted to use the UAS-Mef2 CDS construct for my overexpression lines. This was because I wanted to generate constructs with a C-terminal, 3xHA epitope tag that may be useful for future protein immunoprecipitation experiments. Although Mef2 antibodies do exist, these are finite resources and thus an epitope tag with a variety of tested, commercially available antibodies is more desirable. Thus, if one were to use the UAS-Mef2 10t4a construct, the cloning workflow would have been more challenging due to the requirement to subclone multiple fragments: to clone in the 5' UTR, coding sequence, tag, and the 3' UTR to follow. Instead, and as described in materials and methods (**Chapter 2**), UAS-CDS+3xHA lines were more easily generated by fusing together two fragments: The first ~1kb of Mef2 CDS (from the original UAS-Mef2 10t4a construct), amplified by PCR; and the remaining C-terminal fragment, including the 3xHA epitope tag, which we had synthesised. They could be subcloned in separately and fused in frame by making use of a naturally occurring BglII restriction site within the Mef2 coding sequence. The sequence encoding the MADS/Mef2 domain lie within the N-terminal region that was amplified by PCR, reinforcing why synthesising the C-terminal region was also appropriate for our approach as it would be constant for all lines. Finally, including the HA tag within the coding sequence also meant we could use our standard pUAST-attB plasmid (as used for all other constructs), rather than also needing to obtain a vector already containing a HA tag within the vector backbone.

I generated five different constructs to study the effect of disrupting this region on Mef2 function, including that of Mef2 WT (**Fig. 6.2**). To investigate whether the Mef2C mutant identified *in vitro* (VLL65-67ASR) (Molkentin et al. 1996a) was also non-functional in *Drosophila*, I generated the same UAS-Mef2 VLL65-67ASR mutant (**Fig. 6.2**). I then made alanine point mutants of three other residues (**Fig. 6.2**). As per the initial focus of the project, these were designed based upon their implication in facilitating the HDAC/Mef2 interaction, however were also designed to potentially identify critical residue(s) for Mef2 function *in vivo*. The side chains of both L66 and L67 form the central part of the pocket that interacts with Class IIa HDACs *in* vitro (**Fig. 6.1F-G**) (Han et al. 2005), while mutation of L67 reduces the HDAC4/Mef2D interaction in a mammalian-2-hybrid assay (Jayathilaka et al. 2012) . Because both Mef2A and Mef2B crystal structures indicate V65 is oriented away from the binding interface, with its side chain protruding inwards towards the interior of the Mef2 structure, I decided not to mutate this residue as it was unlikely to be directly involved in Mef2-protein interactions (**Fig.6.1D,G**). In addition to residues within the 65-67 region, I also made an additional alanine mutation of Y69 (Y69A). Its large aromatic, polar side chain protrudes into the base of the pocket and I hypothesised it may also be important for Mef2 function (**Fig. 6.1D-F**). Furthermore, mutation of Y69 of Mef2B has been shown previously to reduce the physical interaction with HDAC4 when analysed by EMSA (Han et al. 2005). All constructs following site-directed mutagenesis were fully sequence verified prior to injection, and DNA sequence chromatograms verifying the desired changes can be visualised in **Figure 6.2**.

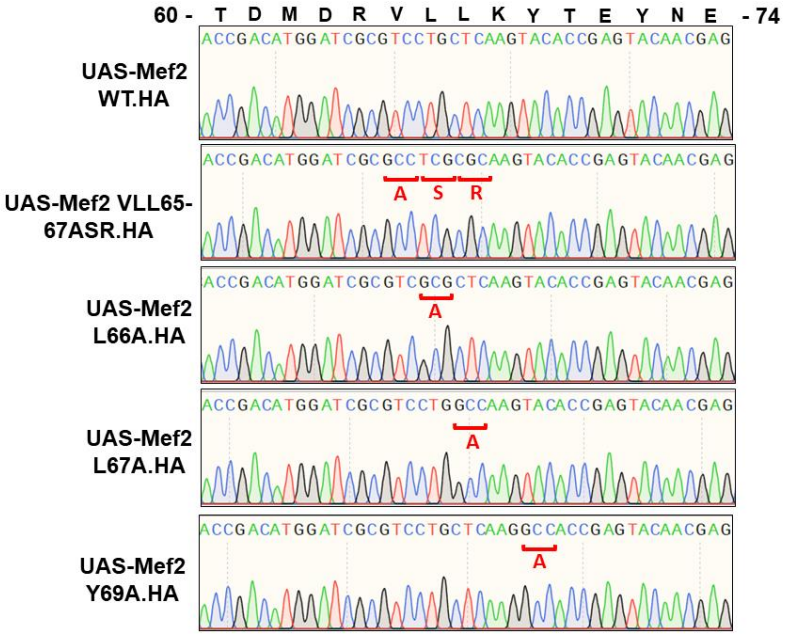


Figure 6. 2. Mutagenesis of the Mef2 domain in the generation of UAS-Mef2 constructs. DNA sequence chromatograms of the region encoding residues 60-74 in UAS-Mef2 construct generated in this project. Sequence of miniprepped constructs, prior to injection, are shown. Mutated codons to generate the desired amino acid change are shown in red.

6.3 Differential functionality of UAS-Mef2 mutants *in vivo*. Leucine-66 is a critical residue for Mef2 function.

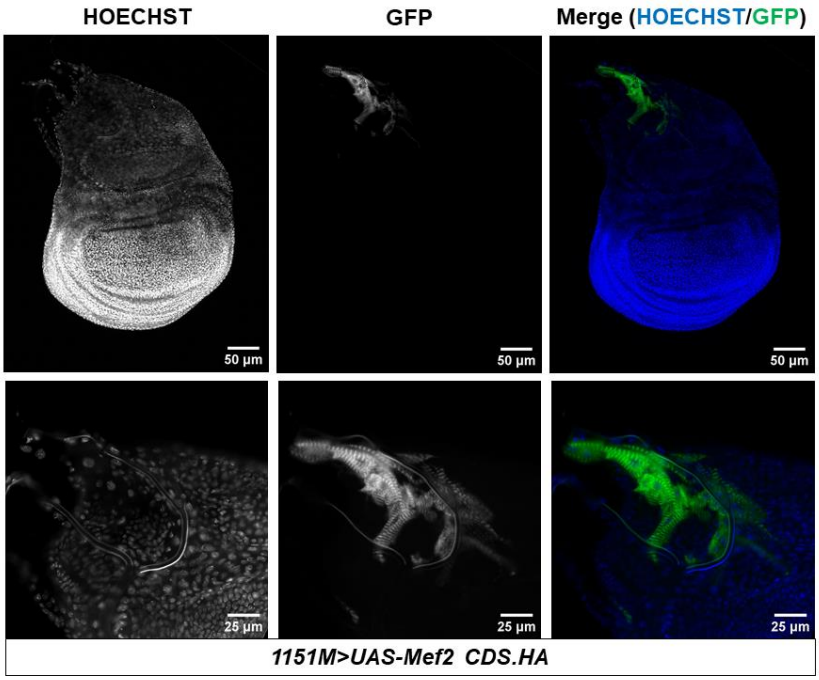


Figure 6. 3. Mef2 overexpression induces premature differentiation of L3 adult muscle progenitors.
Wing imaginal disc isolated from wandering L3 larva overexpressing UAS-Mef2 CDS.HA using the 1151;;Mhc-GFP-Gal4 driver. Overexpression induces the premature expression of Mhc in the AMPs**.** The Mhc also forms a striated appearance.

To assess the functional consequences of mutating these residues *in vivo*, UAS-Mef2 transgenes were expressed in the AMPs using the 1151;;MHC-GFP driver, and analysed for their ability to induce the premature differentiation phenotype previously described. To quantify any difference in the activity of mutants, the percentage of wing discs analysed that display the premature differentiation phenotype was recorded. As previously described, UAS-Mef2 CDS.HA could induce premature AMP differentiation in 100% wing discs (**Fig. 6.3**). In agreement with *in vitro* data, UAS-Mef2 VLL65-67ASR was completely unable to induce the premature differentiation phenotype, suggesting this mutant is also non-functional *in vivo* (Molkentin et al. 1996a)(**Fig. 6.4C**). This also suggests functional conservation between human Mef2C and *Drosophila* Mef2 as previously inferred.

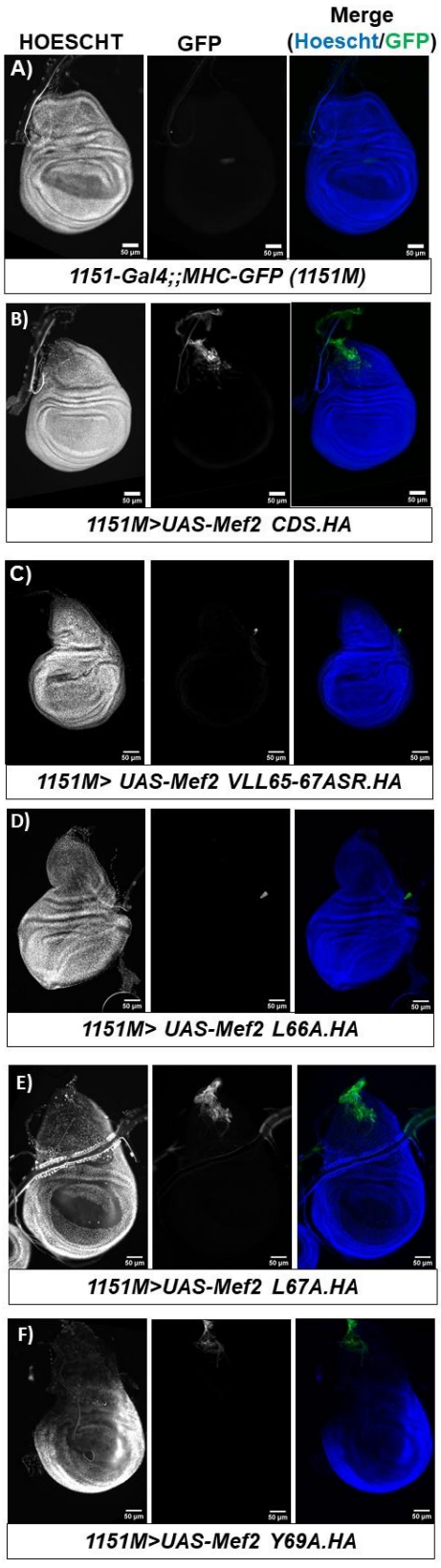


Figure 6. 4. Differential ability of UAS-Mef2 mutants to induce premature differentiation of L3 adult muscle progenitors

Wing imaginal discs isolated from either control (**A**) wandering L3 larvae, or larvae over-expressing UAS-Mef2 constructs using the 1151;;Mhc-GFP-Gal4 driver. Overexpression of UAS-Mef2.CDS induces the premature expression of Mhc in the AMPs **(B).** UAS-Mef2 VLL65-67ASR (**C**) and UAS-Mef2 L66A (**D**) fail to induce this phenotype. UAS-Mef2 L67A induces premature differentiation **(E)** akin to UAS-Mef2 WT.HA (**A**). UAS-Mef2 Y69A **(F)** induces premature differentiation, but less efficiently than UAS-Mef2 WT.

There are two possible reasons for this loss of function: A combinatorial effect caused by disruption to all three adjacent amino acids; or the mutagenesis of either one, or two critical amino acids, with other mutations present in the VLL65-67ASR mutant being passenger mutations that lack the ability to severely disrupt Mef2 function on their own. Interestingly, individual point mutants L66A and L67A had differential consequences on Mef2 function. UAS-Mef2 L66A was also unable to induce premature differentiation, analogous to the UAS-Mef2 VLL65-67ASR mutant (**Fig. 6.4D**); By contrast, UAS-Mef2 L67A retained the ability to induce premature differentiation in 100% of wing discs, with the phenotype indistinguishable of UAS-Mef2 WT (**Fig. 6.4E**). This could mean that L66 is the critical residue for Mef2 function, however the individual significance of V65 to Mef2 function *in vivo* is not yet known. Interestingly, the Y69 mutation has an intermediate effect on Mef2 function: While it still possesses the ability to induce premature differentiation akin to UAS-Mef2 WT (**Fig. 6.4F**), it does so in only 75% of discs. This suggests that Y69 is not necessary for Mef2 transcriptional activity *in vivo*, however the Y69A mutant is less active than both WT, and the L67A mutant.

6.4 Nejire (Nej): A candidate for regulating Mef2 during *Drosophila* muscle differentiation

Given the identification that residues 65-70 in the Mef2 domain are important for Mef2 transcriptional activation in a *Drosophila* model for muscle differentiation *in vivo* (**Fig. 6.4**), I next sought to investigate possible cofactors that may be required for Mef2 function. Because, as previously described, loss of HDAC4 did not yield phenotypes analogous to the of Mef2 loss-of-function, it is unlikely that HDAC4 is required for Mef2

transcriptional activity *in vivo*. Thus, I next wanted to investigate the possible reasons for why these mutations affect Mef2 function *in vivo*, outside of the context of the Mef2/HDAC axis directly. Having said this, the common bonding site for both proteins' may yet have functional significance.

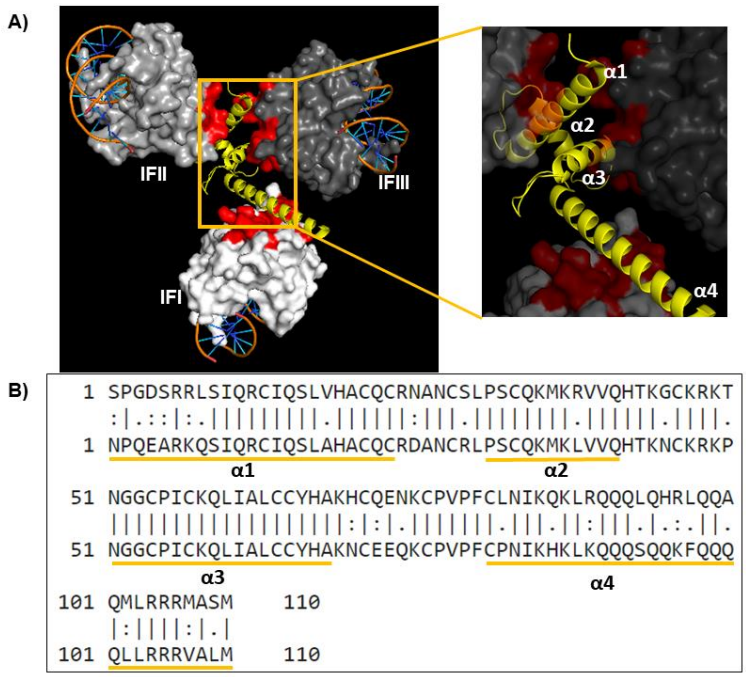


Figure 6. 5. The *Drosophila* p300 homologue, Nejire, is a candidate protein to regulate Mef2 in vivo.
(A) Pymol render of the human Mef2A/p300 crystal structure One domain of p300, the TAZ2 domain (yellow) can interact simultaneously with 3 Mef2 dimers bound to DNA. The TAZ2 domain forms four α-helices (α1-4), which together contribute to the formation of three interaction interfaces (IFI-III) with Mef2A. The TAZ2 domain interacts with residues 65-70 of Mef2A, which has been shown to be important for regulating Mef2 function. **(B)** Alignment of the TAZ2 domain of human p300 (top) and *Drosophila* Nejire (bottom). Residues corresponding to each of the four α-helices shown in **A** are indicated. The TAZ2 domain is highly conserved between both species.

p300 augments Mef2 transcriptional activity *in vitro* (Sartorelli et al. 1997; Ma et al. 2005; Angelelli et al. 2008). Furthermore, previous data have shown that vertebrate p300 physically interacts with Mef2 via the domain for which mutations affect Mef2 transcriptional activity. In the published crystal structure, the TAZ2 domain of p300 interacts with Mef2: Four α-helices of the TAZ2 domain can together interact with three Mef2 dimers (**Fig. 6.5A)** (He et al. 2011). Thus, it is plausible that the aforementioned mutations within the Mef2 domain disrupt the Mef2-p300 interaction, which in turn disrupt Mef2 transcriptional activity *in vivo*.

To investigate whether the function of p300 may be conserved with the *Drosophila* homologue, Neiire (Nej), I aligned the sequences for the TAZ2 domain (**Fig. 6.5B**). This domain is highly conserved, suggesting functional conservation between the two proteins. Because the MADS/Mef2 domain of Mef2 is also highly conserved, this supports an idea whereby Mef2 and Nej may interact in *Drosophila* muscle differentiation *in vivo.* This is supported by the report that both proteins physically interact in a yeast-2-hybrid (Lin and Baines 2019). However, the function of *nej* in *Drosophila* muscle development is not well understood, and this led me to investigate the possible role of nej as a transcriptional coactivator required for muscle development, as well as for Mef2 function.

6.5 *nej* knockdown inhibits DLM development

I chose to investigate the role of *nej* during IFM development, which has not been characterised previously. First, I knocked down *nej* in AMPs using 1151-Gal4 to investigate whether, akin to Mef2, it is also required for DLM development. I used two different nej RNAi lines with target sequences in different regions of the *nej* mRNA: one from the VDRC KK library (102885), and another from the second-generation TRiP library (BL37489). The latter generally yield stronger knockdown than the first generation TRiP lines and have no predicted off-target effects (OTEs) (Perkins et al. 2015). Both lines have also been used in the published literature. As seen in **Figure 6.6,** both RNAi lines inhibited DLM development, but to different extents. At 25°C, KK102885-mediated knockdown significantly reduced the mean fibre number to 3.4, with only 25% of hemithoraces scored having a wild-type number of DLM fibres (**Fig. 6.6C**). The phenotype was made significantly more severe when the cross was raised at 29°C, with a mean of 1.6 fibres and only 8% of hemithoraces presented with wild-type DLM numbers (**Fig. 6.6D**). Furthermore, I also saw a disruption to the normal patterning of the fibres that do form. In addition to fibres being absent, fibres also appeared misshapen and irregularly sized when compared to the DLM organisation within the thorax of a control fly (**Fig. 6.6A-B**). By contrast, knockdown with BL37489 inhibited DLM formation more strongly at both 25°C and 29°C. At both temperatures, 92% of hemithoraces displayed no differentiated DLM fibres (**Fig. 6.6E-G**). Although the two RNAi lines gave different phenotypic penetrance, the formation of an allelic series suggests this is most likely due to different knockdown efficiencies by different RNAi lines.

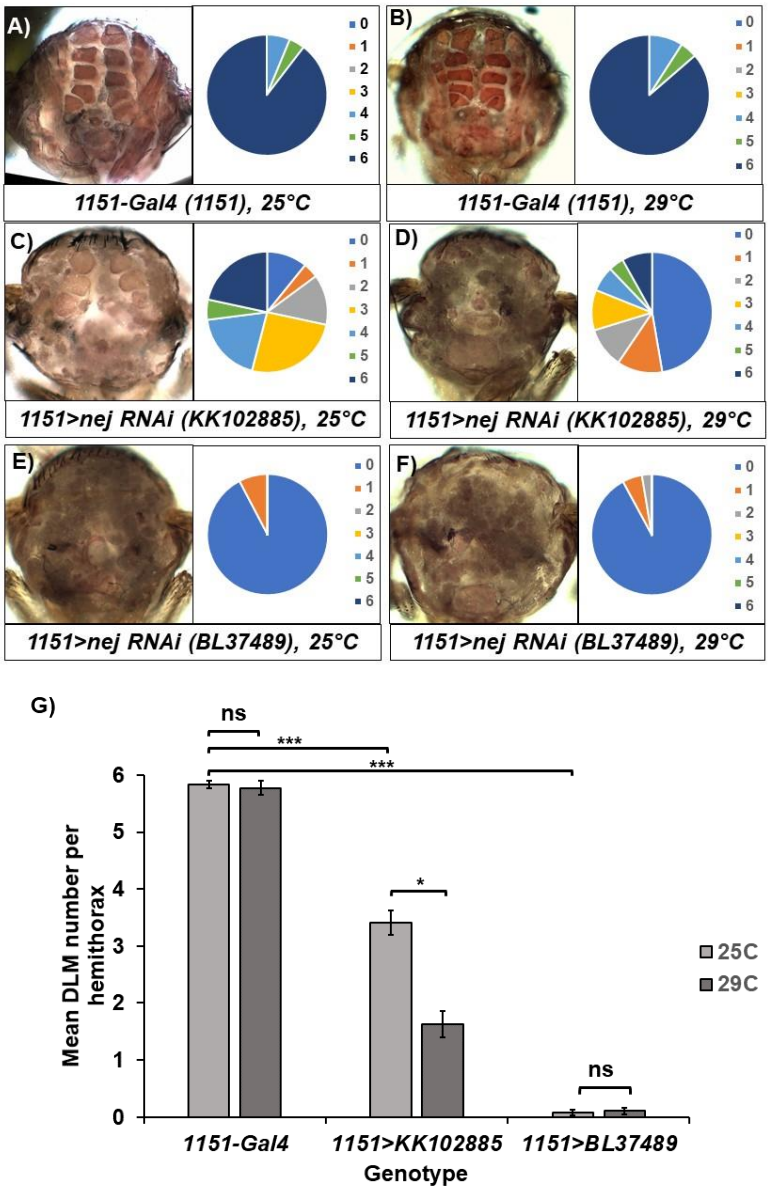


Figure 6. 6. nejire is required for DLM development.
(A-F) Hematoxylin-stained transverse cross-sections of control pupae 96hr APF **(A,B),** and pupae expressing two different RNAi constructs against *nej*: either KK102885 (VDRC KK library, **C,D**), or BL36489 (TRiP library, **E,F**), when raised at either 25°C or 29°C . Next to cross sections are pie-charts displaying distribution of the number of DLM fibres scored per hemithorax in each genotype sample. (**G**) Mean DLM fibre number per hemithorax for each genotype when raised when raised at either 25°C or 29°C. Error bars represent SEM. *** p<0.001, ** p<0.005, * p<0.01, ns = not significant. calculated from Kruskal-wallis followed by Dunn's multiple comparisons test.

6.6 *nej* overexpression inhibits DLM development

Given the effect observed by *nej* knockdown, I next used overexpression to further the understanding of the effect of *nej* dysregulation on DLM development. Interestingly, *nej* overexpression also inhibited DLM formation (**Fig. 6.7**). The mean fibre number was

significantly reduced to 3, however, in a similar phenotype observed to *nej* knockdown with KK102885, I observed a broad range of DLM fibre numbers between flies and only 17% of hemithoraces had a wild-type DLM number (**Fig. 6.7F-G**). As with *nej* knockdown, I again saw a disruption to the normal patterning of the fibres that do form (**Fig. 6.7A-D**). To see whether this phenotype may arise primarily through the hyperactivation of Mef2, these data were compared to the phenotype observed when Mef2 is overexpressed using the 1151-Gal4 driver. Interestingly, Mef2 overexpression only had a mild effect on DLM formation. The mean fibre number was 5.42 and was not significantly different from controls, while 70% of hemithoraces presented with a wild-type number of DLMs (**Fig. 6.7D-F**). These data show that both knockdown, and overexpression of *nej* severely disrupts muscle differentiation *in vivo*, indicating that a fine balance of nej is required for the normal coordination of DLM development. However, given the overexpression phenotype does not phenocopy that of Mef2 overexpression, this phenotype is not likely caused solely by a specific effect on Mef2 function.

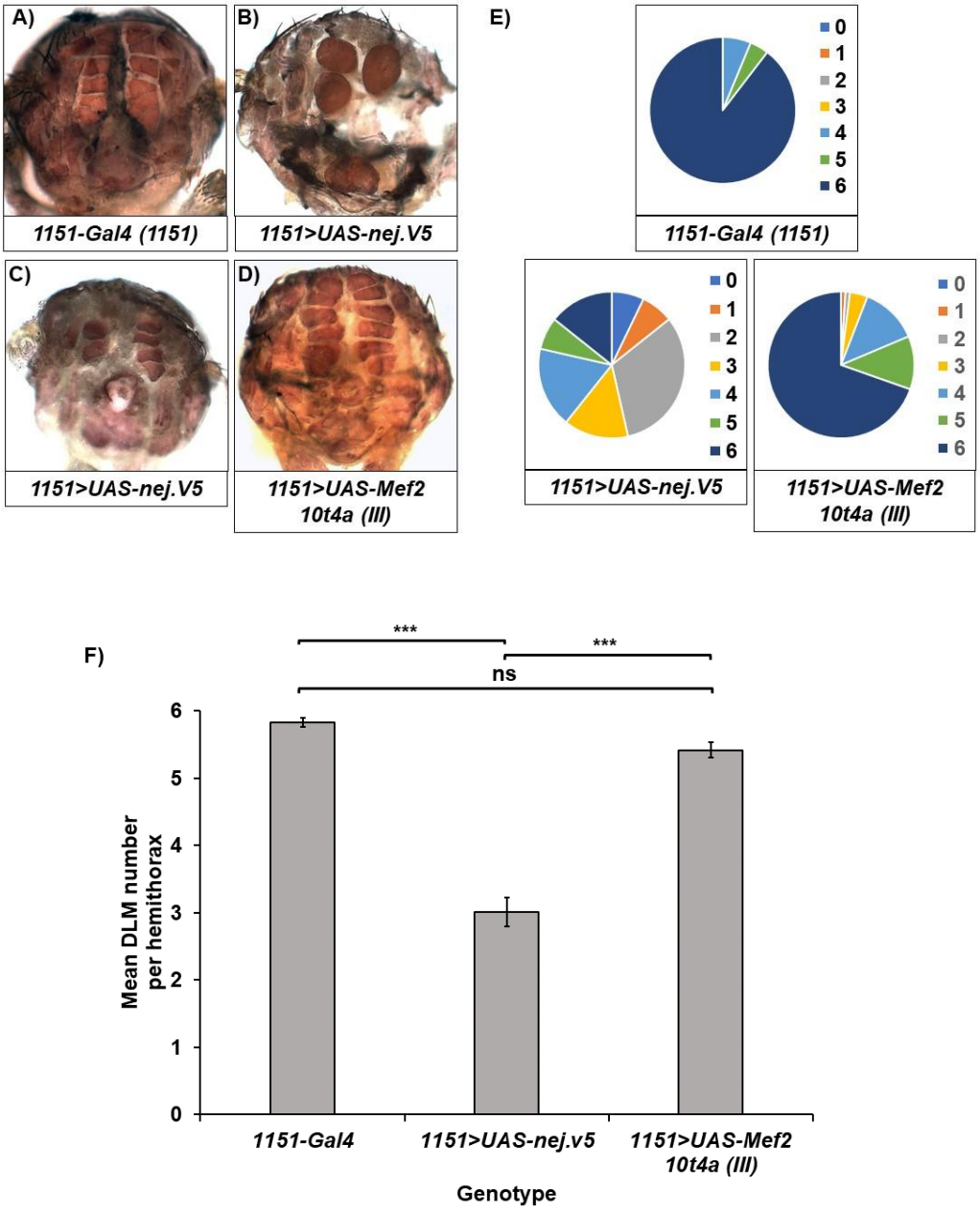


Figure 6. 7. *nejire* overexpression disrupts DLM formation.
Hematoxylin-stained transverse cross-sections of control flies **(A),** flies overexpressing *nejire* **(B-C),** and flies overexpressing *Mef2* (**D**), using the 1151-Gal4 driver. **(E)** Pie charts displaying the distribution of the number of DLM fibres per hemithorax scored for each genotype. **(F)** Mean DLM fibre number per hemithorax for each genotype when raised when raised at either 25°C or 29°C. Error bars represent SEM. Mean DLM fibre number per hemithorax for each genotype when raised when raised at either 25°C or 29°C. Error bars represent SEM. *** p<0.001, ** p<0.005, * p<0.01, ns = not significant. calculated from Kruskal-wallis followed by Dunn's multiple comparisons test.

6.7 Nej does not inhibit Mef2-induced muscle differentiation

Given that nej dysregulation disrupts DLM formation, I next sought to investigate the potential relationship nej has in the context of Mef2 function and regulation *in vivo*. The most appropriate model to study this *in vivo* is in the L3 AMPs, where I have shown throughout that Mef2 overexpression induces their premature differentiation.

Firstly, given that *nej* overexpression disrupted DLM development, it is possible that Nej also acts as a negative regulator of Mef2, despite the generally accepted understanding of p300 being a transcriptional coactivator (Sartorelli et al. 1997; Ma et al. 2005; Angelelli et al. 2008). Interestingly, one paper in *Drosophila* cells found that nej repressed transcription of a Mef2-dependent reporter (Lin and Baines 2019), leaving the role of nej in regulating Mef2 *in vivo* an open question. To address this. I expressed UAS-Mef2 and UAS-nej in the AMPs using 1151;;MHC-GFP-Gal4 (**Fig. 6.8**). If nej was acting as a transcriptional corepressor, akin to that of HDAC4, then one may expect the premature differentiation phenotype to be suppressed. Because the UAS-nej transgene is present on ChrIII, UAS-Mef2 10t4a on ChrII was used, which still efficiently induced the premature differentiation phenotype in 100% of wing discs (**Fig. 6.8B**). Similarly, over-expressing UAS-Mef2 alongside UAS-mCherry also induced premature differentiation in 100% of discs (**Fig. 6.8E**) showing that the phenotype is not affected by the presence of multiple transgenes potentially diluting the effect of the Gal4. Moreover, *nej* overexpression on its own did not induce premature differentiation as with Mef2 overexpression (**Fig. 6.8C**). Interestingly, Nej largely did not affect the ability for Mef2 to induce the premature differentiation of L3 AMPs, with the phenotype present in 96% of discs analysed (**Fig. 6.8D)**. This suggests that, generally, Nej does not repress Mef2 transcriptional activity *in vivo*. Because I did not observe 100% penetrance, this could mean that Nej also is not augmenting Mef2

activity, however this could also be due to protein dynamics and whether Nej protein levels are rate limiting or not.

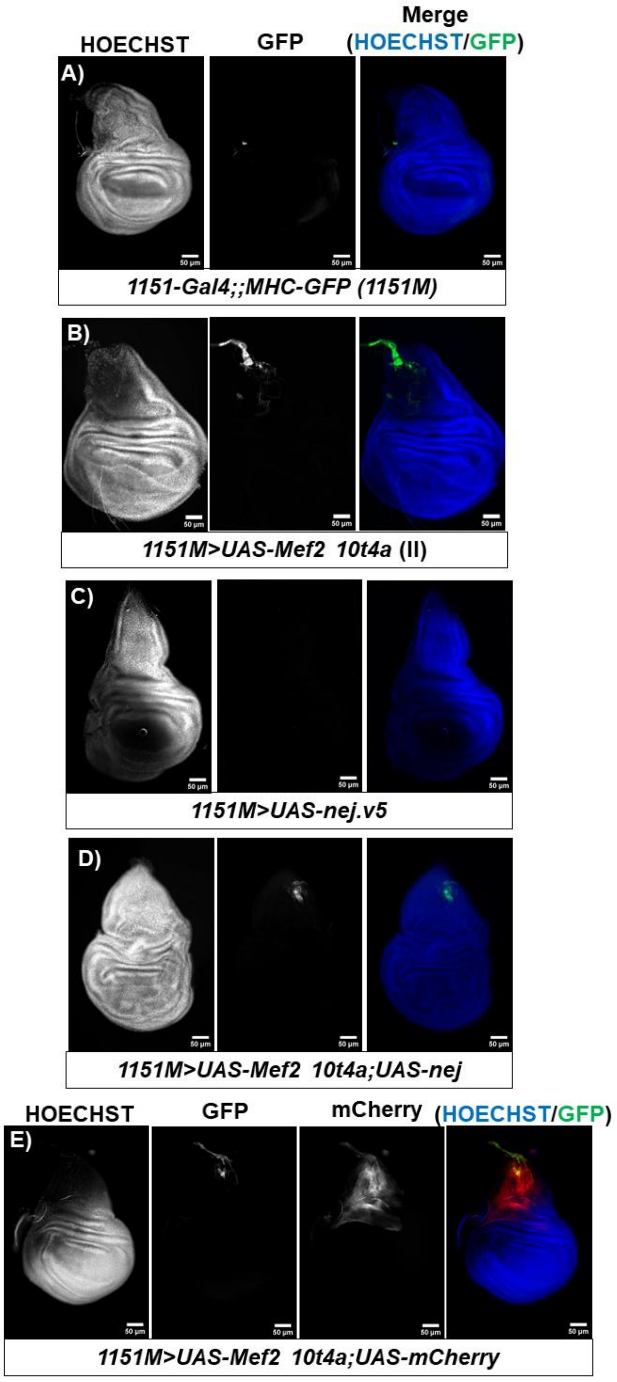


Figure 6. 8. Nej does not repress Mef2 induced differentiation of L3 adult muscle progenitors

L3 wing imaginal discs isolated from wandering L3 control larvae (**A**), or larvae expressing Mef2 **(B),** nejire (**C**) Mef2 and nejire **(D),** or Mef2 and mCherry **(E),** using the 1151;;Mhc-GFP-Gal4 driver. Overexpression of Mef2, but not nej alone, induces premature expression of Mhc in the AMP population (green). Co-expression of Mef2 and nejire, or Mef2 and mCherry does not affect premature Mhc expression.

6.8 *nej* may be required for Mef2-induced muscle differentiation

Next, I wanted to investigate whether nej is required for Mef2 function *in vivo*. Using the same model of Mef2-induced premature differentiation, I aimed to knock-down *nej*, already shown to inhibit DLM development (**Fig. 6.6**), in the presence of *Mef2* overexpression using 1151;;MHC-GFP. One would hypothesise that, if Nej was required for Mef2 function, knocking down nej in AMPs would prevent their premature differentiation, even in the context of overexpressing Mef2. To answer this question, I generated stocks with UAS-Mef2, and either UAS-nej RNAi (KK102885), or UAS-Nej RNAi (BL37489). I found that knockdown with either construct gave contrasting results on the ability for Mef2 to induce the premature differentiation phenotype. Knockdown with KK102885 did not affect the phenotype, with Mhc still detectable in 100% of wing discs (**Fig. 6.9B).** By contrast, knockdown with BL37489, which gave a stronger inhibition of DLM formation, abolished the ability for Mef2 to induce premature differentiation in the L3 AMPs, with 0% discs presenting with this phenotype (**Fig. 6.9C**). Given that: 1151-Gal4 expresses in AMPs from the second instar stage onwards, the differences between the two lines could be due to: the KK102885 not

driving a strong enough knockdown to deplete nej protein sufficiently in the AMPs. This may result in sufficient Nej protein to enable Mef2-dependent gene activation.

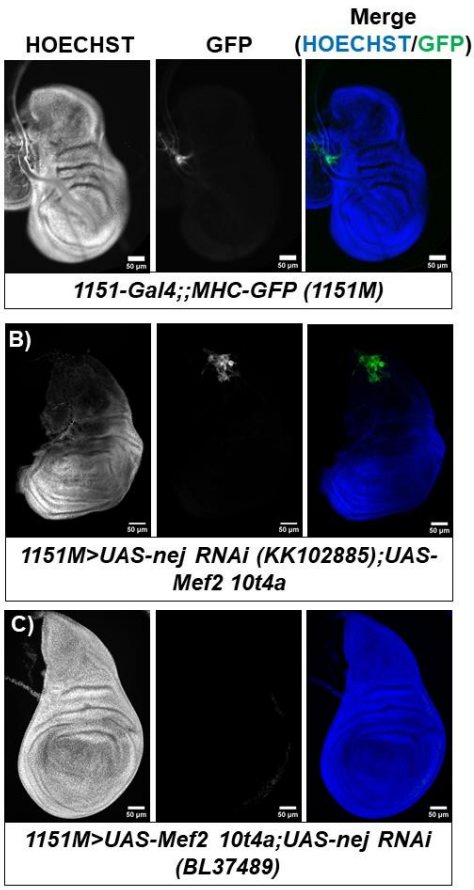


Figure 6. 9. Nej may be required for Mef2-induced premature differentiation of L3 adult muscle progenitors.
L3 wing imaginal discs isolated from wandering L3 control larvae (**A**), or larvae expressing Mef2, as well as one two different constructs to knock down *nej: either KK102885* (**B**), or BL37489 **(C)**, using the 1151;;mhc-GFP-Gal4 driver. Knockdown with KK102885 does not suppress the premature differentiation phenotype, whereas premature differentiation is completely inhibited by the co-expression of BL37489 alongside Mef2.

6.9 *nej* is required for proliferation and establishment of the AMP population

nej, as the homologue of p300, is not a Mef2-specific binding partner and is likely to interact with many proteins *in vivo*. As previously mentioned, p300 is predicted to have

over 400 directly interacting proteins (Dancy and Cole 2015) . Therefore, it is possible *nej* knockdown in AMPs may cause phenotypes unrelated to Mef2. Given that knockdown with BL37489 completely inhibited DLM formation, one possibility is that this phenotype may be caused through the failure of the AMP population to amplify during larval stages. If this was to occur, then the assay used to study a Mef2-induced differentiation phenotype within the L3 AMPs may not be informative, since the lack of AMPs would obviously prevent Mef2-induced differentiation. To support this possibility, it has also been observed that knockdown of another transcription factor, *zfh1*, also inhibits the establishment of the L3 AMP population (Rob Mitchell, unpublished).

To address this question, I knocked down nej using the 1151-Gal4 driver in combination with the Bl37489 line. Alongside, we also knocked down Mef2 which, although strongly inhibits DLM formation (Soler et al. 2012), is understood not to affect the AMP population and instead differentiation of these AMPs later in development (MVT lab unpublished). To analyse the AMP population, we antibody stained against Cut, a transcription factor that is an established marker for all AMPs (Sudarsan et al. 2001; Zappia et al. 2020), as well as an antibody against Mef2, which is also expressed in the AMPs (Cripps et al. 2004; Lovato et al. 2005; Soler and Taylor 2009). In the 1151-Gal4 control, both Mef2 and Cut are expressed throughout the AMP population with the notum of the wing disc (**Fig. 6.10A**). When Mef2 is knocked down, the AMPs are present, as seen by positive cut staining, however these cells lack Mef2 expression, highlighting the efficiency of this RNAi line in depleting Mef2 levels (**Fig. 6.10B**). By contrast, when *nej* is knocked down, we see no Cut or Mef2 staining (**Fig. 6.10C**), suggesting that *nej* knockdown with the stronger, BL37489 RNAi line prevents the proper establishment of the AMP population during the second and third larval instars. This ultimately suggests that we cannot confirm the requirement of nej for Mef2 transcriptional activation during muscle differentiation *in vivo* using these experiments.

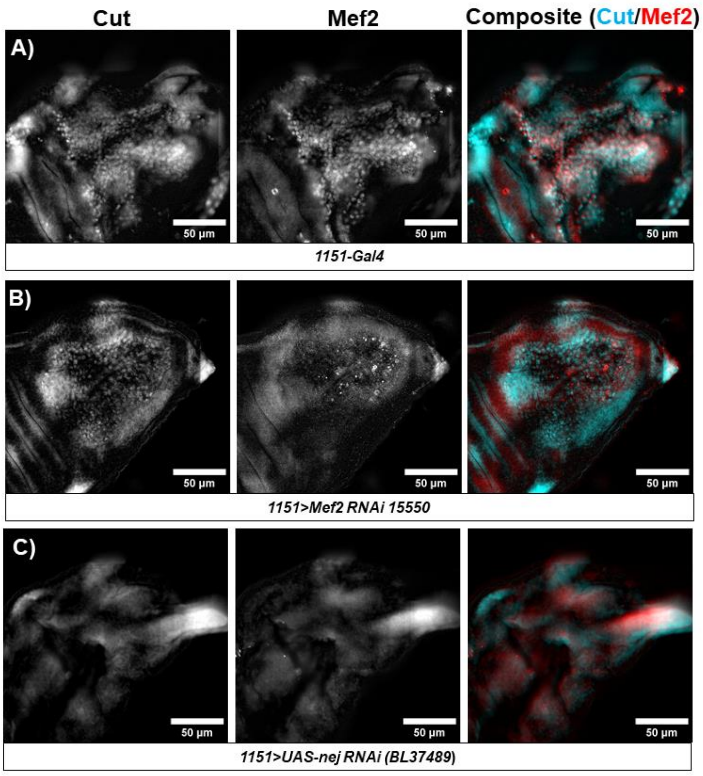


Figure 6. 10. *nej* is required for establishment of the AMP population.
L3 wing imaginal discs (notum shown) of *1151-Gal4* control (**A**), *Mef2* knockdown (**B**), and *nej* knockdown (**C**), stained against Cut and Mef2. *nej* knockdown causes a complete absence of Mef2 and Cut positive staining.

Results summary

In this chapter, using a Mef2-overexpression inducible model for muscle differentiation, I have identified critical residues required for Mef2 transcription activity *in vivo*. Different residues within the Mef2 domain display varying sensitivity to mutation and, notably, I have identified leucine-66 as a critical residue for Mef2 function *in vivo*. These data further *in vitro* studies and demonstrate that disruption to the Mef2 domain, known to interact with both class IIa HDACs and p300/CBP, can significantly disrupt Mef2 function during muscle differentiation *in vivo*. Given disruption to this domain could affect Mef2-protein interactions necessary for its function, and following the investigations of HDAC4 function in chapters 3 and 4, in this chapter I have demonstrated a critical role for the p300/CBP homologue, Nejire, during *Drosophila* adult muscle differentiation. I have shown that a fine balance of *nej* expression is required for normal DLM development, given either overexpression or RNAi-induced knockdown is sufficient to disrupt normal DLM development. However, with these current experiments, I have thus far been able to demonstrate a conclusive

functional link between Mef2 transcription and Nej *in vivo.* This was likely contributed to by the knowledge that Nej may interact with many different proteins aside from Mef2, and thus phenotypic effects arising from its deregulation would likely not exclusively affect Mef2 function during muscle differentiation.

Chapter 7: Discussion

7.1 CRISPR-Cas9 technology as an effective tool to study Mef2 function and regulation during muscle differentiation *in vivo*.

The ultimate goal of geneticists is to identify the function of endogenous genes in any given biological process. Of course, much progress has been, and continues to be made using the Gal4-UAS system to modulate gene expression in a tissue specific manner (Del Valle Rodríguez et al. 2012). However, these techniques have their inherent challenges. For example, in the context of the HDAC4-mediated inhibition of muscle differentiation, the generation of UAS-HDAC4 mutants has clearly been incredibly informative in providing insights in to the function of Class IIa HDACs in *Drosophila*. However, the caveat is that HDAC4 *can* inhibit muscle differentiation, rather than whether the endogenous gene actually *does*. Ideally, one would mutate specific residues within the endogenous gene. In this context, the resulting protein is expressed under the endogenous regulatory elements, and subsequently the most informative conclusions may be drawn with regards to protein function.

CRISPR-Cas9, revolutionised researchers' ability to investigate gene function. In this project, the goal was to apply previously developed CRISPR-Cas9 techniques to further our understanding of HDAC4 and Mef2 during *Drosophila* muscle differentiation *in vivo*. In this project, I have successfully developed novel, insertion-ready deletion alleles which may be used for investigating the function of the endogenous gene, in future experiments. I also successfully generated fluorescently-tagged HDAC4 and Mef2 alleles using an alternative, direct-tagging approach with CRISPR-cas9. These will be discussed in detail in the section below. In addition. I also designed and generated a similar allele by targeting the *Him* gene, which has also been established as an important gene for muscle differentiation, and has been implicated in the regulation of Mef2 (Liotta et al. 2007; Soler and Taylor 2009; Zappia et al. 2020). For this gene, which was not the focus of this project, the primers and schematic for the gene targeting can be found in **Appendix 1** and **Appendix 14**, respectively. Another gene of interest during his project, *nej*, was not targeted using CRISPR-Cas9, although remains a possibility. Nevertheless, as I will discuss, the tools I have generated will have significant benefit to both *Drosophila* muscle biologists, but also potentially the wider *Drosophila* research community. In this discussion, I have integrated multiple different aspects of this project into common themes, in a hope to provide clarity on

the both the most significant results arising from this project, but also the pressing questions which have arisen subsequently.

7.2 The role of HDAC4 during *Drosophila* muscle differentiation *in vivo*

7.2.1 *Drosophila* HDAC4 can inhibit Mef2 *in vivo*

One of the main goals of this project was to investigate Class IIa HDAC function in the context of muscle differentiation *in* vivo. Indeed, the presence of only a single Mef2 and Class IIa HDAC gene, both of which are well conserved with their vertebrate counterparts, made *Drosophila* an ideal model for this analysis. Here, I have demonstrated that, when overexpressed in undifferentiated myoblast populations, *Drosophila* HDAC4 can potently inhibit the differentiation of both the larval somatic musculature during embryogenesis, and the formation of the DLM indirect-flight muscles. These data extend previous *in vitro* data, where human HDAC4 or HDAC5 can inhibit both *in vitro* muscle differentiation models (Lu et al. 2000b; McKinsey et al. 2000a; McKinsey et al. 2000b; Miska et al. 2001).

Interestingly, when expressed in undifferentiated myoblast populations using either the TwipTwip- or 1151-Gal4 drivers, it appears this phenotype is largely a result of its direct inhibition of Mef2, since mutation of the Mef2 binding-domain is sufficient to abolish the inhibitory capacity of HDAC4. The Mef2 binding mutant, K165,L168,I172A, for instance, has already previously been published as a Mef2 binding mutant in *Drosophila* (Main et al. 2021). Indeed, mutation of either L175 or V179 of human HDAC4, the equivalent residues to L168 and I172 in *Drosophila* HDAC4, have previously been shown to abolish the interaction with Mef2 (Wang and Yang 2001; Han et al. 2005; Jayathilaka et al. 2012). To my knowledge, there remains no supporting biochemical data for the importance of the K165A mutation for Mef2 binding. In *Drosophila*, I have shown that a single point mutation, L168A, also completely abolished the ability for HDAC4 to inhibit muscle differentiation, analogous to that of the K165,L168,I172A mutant. This likely suggests that mutation of L168 is sufficient to completely abolish the HDAC4/Mef2 interaction in *Drosophila*, which supports biochemical data using the mammalian homologues described above. By contrast, the F171A mutation did still inhibit muscle differentiation in *Drosophila*, but to a lesser extent than the L168A mutation. This mutation *in vitro* has been shown to

reduce the HDAC4/Mef2B interaction by approximately ten-fold (Han et al. 2005). Thus, the formation of this allelic series of phenotype severity provides strong evidence that HDAC4 inhibits muscle differentiation by physically interacting with Mef2 *in vivo*, and also provides evidence for critical residues required for this interaction.

To bolster this conclusion, one may desire supporting Co-immunoprecipitation experiments from *Drosophila* tissue to quantify the physical interaction biochemically. To do so, a UAS-HDAC4 F171.myc epitope tagged line would need to be generated. This is because the F171A mutant I generated does not possess an epitope tag, nor is there a commercially available antibody against *Drosophila* HDAC4.

7.2.2 The potential mechanisms of HDAC4-mediated inhibition of Mef2 *in vivo*.

There question that remains is how HDAC4 inhibits Mef2 transcriptional activity *in vivo*. Firstly, Class IIa HDACs are perceived to lack direct catalytic activity, in large part to a tyrosine-histidine mutation within the HDAC domain of vertebrate proteins (Lahm et al. 2007; Bottomley et al. 2008). However, this catalytic tyrosine residue is conserved in *Drosophila* HDAC4, suggesting it may retain catalytic activity unlike it vertebrate counterparts. Indeed, HDAC4 has been implicated in the deacetylation of the FOXO transcription factor during lipid metabolism (Wang et al. 2011). It is possible HDAC4 could influence Mef2 histone deacetylation (Shvedunova and Akhtar 2022), or through the direct deacetylation of Mef2, since Mef2 acetylation increases Mef2 transcriptional activity *in vitro* (Ma et al. 2005; Angelelli et al. 2008). However, given that the UAS-HDAC4 Y1142H catalytic mutant was still capable of inhibiting DLM development as severely as wild-type, even if *Drosophila* HDAC4 was catalytically active, then this activity may not be required for Mef2 inhibition during muscle differentiation.

There are other possibilities for HDAC4-mediated inhibition of Mef2, which could be dependent upon its physical interaction. *In* vitro, Class IIa HDACs recruit class I HDACs the HDAC domain (Fischle et al. 2002). Therefore, HDAC4 could act as a scaffold for the formation of a larger repressor complex to shut down Mef2-dependent transcription. Indeed, *in vitro* the HDAC domain was required for HDAC4-mediated inhibition of C2 myoblast differentiation (Miska et al. 2001). It could be important to investigate whether a UAS-HDAC4 mutant lacking the HDAC domain could still inhibit *Drosophila* muscle formation. Moreover *in* vitro, HDAC4 has been found to stimulate the sumoylation of Mef2D (Grégoire and Yang 2005). Importantly, Mef2 sumoylation

represses Mef2 transcriptional activity (Grégoire and Yang 2005; Kang et al. 2006), while in *Drosophila,* sumoylation-deficient Mef2 may display higher transcriptional activity than wild-type (Rob Mitchell 2023, unpublished). Finally, Class IIa HDAC binding could compete with other coactivators required for Mef2 activation. As investigated in this project, both Class IIa HDACs, and the p300 HAT interact with the same region of Mef2, and based on structural analysis, it appears unlikely there is sufficient binding capacity for both proteins' to a single Mef2 dimer. Indeed, unlike HDACs, p300 is associated with augmenting Mef2 transcriptional activity (Sartorelli et al. 1997; Ma et al. 2005; Angelelli et al. 2008).

7.2.3 A Mef2-independent role for HDAC4 in developing muscle fibres, post-fusion?

Interestingly, I have possibly identified a different role in the negative regulation of muscle differentiation, not previously identified *in vitro*, which was limited to the nascent differentiation of myoblasts into syncytial myotubes that do not fully recapitulate muscle differentiation *in vivo*. Specifically, *in vivo*, following the fusion of myoblasts to myotubes, the *Drosophila* DLMs mature: growing dramatically to fill the entire thorax, attach to tendons, and undergo myofibrillogenesis, to ultimately generate contractile muscles required to power flight in the adult fly (Spletter et al. 2018). Interestingly, when overexpressed using the Act88F-Gal4 driver, HDAC4 could strongly inhibit DLM formation even with a defective Mef2-binding domain. This suggests Mef2-independent functions of HDAC4 in developing DLM fibres post fusion. Indeed, there roles emerging for HDAC4 beyond transcriptional regulation. In the *Drosophila* brain, over-expressed HDAC4 is pan-cellular, does not cause significant changes to gene expression, however inhibits long-term memory formation and neuronal development (Schwartz et al. 2016; Main et al. 2021). Moreover, HDAC4 genetically interacts with a number of cytoskeletal components in a rough-eye screen (Schwartz et al. 2016). Thus, HDAC4 could possibly function in the cytoplasm to disrupt the formation of the complex cytoskeletal architecture required for muscle formation. However, the exact cause of the apparent Mef2-binding independent phenotype during DLM muscle growth and maturation remains to be further explored. One possible experiment is to use RNA-seq to investigate the transcriptional changes induced by HDAC4, and HDAC4 mutant overexpression throughout muscle development *in* vivo.

7.2.4 A novel, rescuable *HDAC4* allele

My goal was to generate a system whereby the role of endogenous HDAC4 could be probed. I successfully generated a novel, rescuable *HDAC4*$^{\Delta att}$ allele that can be used to study HDAC4 function in *Drosophila in vivo*. *HDAC4*$^{\Delta att}$ was designed to remove over 90% of the CDS, which contained all annotated, functionally important domains, including the Mef2-binding domain, NLS, 14-3-3 binding sites, and the HDAC domain. Thus, it is incredibly likely that *HDAC4*$^{\Delta att}$ would be a null allele. Significantly, the *HDAC4* locus could be engineered to generate the *HDAC4*$^{\Delta att}$ allele. Indeed, this was a major achievement, given the initial deletion of 5.8kb was a compromise between minimising the deletion size in order to maximise the likelihood of generating the desired transgenic, but large enough to both generate a null, and provide maximum versatility in subsequent rescue experiments (Poernbacher et al. 2019)

HDAC4$^{\Delta att}$ LOF is lethal, indicating that *HDAC4* is an essential gene for *Drosophila* viability. Indeed, this novel finding reiterates the initial need for a null *HDAC4* allele in *Drosophila* because until this point, no published *HDAC4* null allele currently exists. Indeed, the current hypomorphs used in the published literature (with ascribed citations on Flybase) are *HDAC4*^{KG09091} (Bellen et al. 2004) and HDAC4^{e04575} ((Thibault et al. 2004). The former only reduced mRNA levels to 60% of that of wild-type HDAC4 (Choi et al. 2015). Furthermore, HDAC4^{KG09091} flies are homozygous viable, as are HDAC4^{KG09091}/HDAC4^{e04575} transheterozygotes (Wang et al. 2011; Choi et al. 2015). Of course, one could question whether the *HDAC4*$^{\Delta att}$ lethal phenotype was induced by some unknown off-target effect. However, this is unlikely, given that insertion of the genomic fragment to generate *HDAC4*$^{\Delta att\text{-}gDNA}$ rescued the lethality. Interestingly, *HDAC4*$^{\Delta att\text{-}gDNA}$ was not as healthy as *Oregon-R* controls. This could possibly be explained by genetic background, which is known to impact upon general physiological and fitness traits (Evangelou et al. 2019). In future, the original injection stock, *yw;;nos-cas9(attP2)* may be a more suitable control to further investigate the overall health of *HDAC4*$^{\Delta att\text{-}gDNA}$. Another explanation may that the scarless nature of the technique may have subtle effects on endogenous *HDAC4* function. Nevertheless, the nature of the insertion-ready deletion system, with *HDAC4*$^{\Delta att\text{-}gDNA}$ being derived from the *HDAC4*$^{\Delta att}$, allows subsequent experiments to be internally controlled. Therefore, this novel, *HDAC4*$^{\Delta att}$ insertion-ready deletion allele can be a powerful platform for investigating the role of Class IIa HDACs *in vivo*.

7.2.5 Is HDAC4 required for normal muscle differentiation *in vivo*?

My rescuable *HDAC4* allele system will allow the systematic functional analysis of *HDAC4 in vivo*. Of course, this system can be used for investigating HDAC4 function in any tissue of interest in the fly, however given the interest of our lab in the role of HDAC4 during muscle development, identification of any muscle phenotypes associated with HDAC4 LOF was the main priority. In this project, some preliminary data has identified two muscle phenotypes associated with HDAC4 LOF, which in turn warrants further investigation. While these phenotypes were measurable, there does not appear to be gross aberrations to muscle development as a direct result of the loss of *HDAC4*, at least indicating initially that it may not be critical for muscle differentiation *in vivo*.

Nevertheless, I observed defects in the patterning of the LT muscles of L3 larvae, as well as the morphology of DLM fibres in surviving late pupae. For the larval muscle phenotype, there are two possibilities for how this phenotype may have arisen. Firstly, the initial specification of the larval muscle pattern during embryogenesis may have been disrupted, possibly through the disruption to founder cell identity. For example, coordinated patterning of LT muscles is specified by the expression of identity transcription factors, whose expression could, in theory be disrupted in *HDAC4*$^{\Delta att}$ mutants (Dobi et al. 2015; Schulman et al. 2015; Junion and Jagla 2022). However, I did not observe any LT muscle phenotypes in the embryonic somatic musculature, although thus far my sample size was relatively small (n=8). Hence, given that the larval muscle phenotype observed was also reasonably subtle, it is possible not enough embryos were initially screened to identify this phenotype. The second possibility may be that larval muscle growth and/or maintenance during larval life may be affected, leading to either muscle fibre splitting, or degeneration. Indeed, mutations acting downstream of larval muscle specification and impacting LT muscle morphology have recently been described (Bertin et al. 2021), although these were observed in the embryo. This investigation into larval muscle phenotypes is also complicated by the fact that there appears to be a maternal contribution of *HDAC4* (Zeremski et al. 2003), which could mask embryonic and early larval muscle phenotypes. Indeed, this could be one limitation of the previous RNAi screen used to analyse muscle morphogenesis in *Drosophila* embryos, which identified no phenotype associated with

HDAC4 knockdown (Schnorrer et al. 2010). Nevertheless, by utilising the power of my rescuable *HDAC4*$^{\Delta att}$ allele, the role of *HDAC4* during embryonic development could be investigated further through generating maternal and zygotic *HDAC4*$^{\Delta att}$ mutants by utilising germline clone analysis (Chou and Perrimon 1992).

a key outstanding question is relating to general cause and effect. i.e. whether the *HDAC4* LOF phenotypes specifically caused by defects in the muscle differentiation pathway, or are they more general phenotypic consequences of global *HDAC4* loss. It is unlikely that lethality is caused solely by muscle specific effects. Indeed, *Drosophila* HDAC4 has been shown to function in other tissues regulation of lipid metabolism (Wang et al. 2011; Choi et al. 2015), long-term memory formation (Fitzsimons et al. 2013) and potassium buffering in the CNS (Li et al. 2021; Lones and DiAntonio 2023). Therefore, it is likely that the global deletion of HDAC4 may impact the development and function of other tissues, as well as muscle. Of course, the most obvious technique to investigate muscle-specific functions of would be using RNAi. I attempted to use RNAi to knockdown *HDAC4* expression in developing DLM fibres, however this did not phenocopy the DLM morphology defects observed in surviving *HDAC4*$^{\Delta att}$ mutants. However, the only RNAi line to have its efficiency validated, GD20522, only reduced HDAC4 protein to approximately 50% of wild-type levels in the brain (Fitzsimons et al. 2013). It is possible that knockdown was not sufficient to induce a phenotype. Indeed, the majority of LOF mutations in the genome are recessive, meaning 50% functional protein may be sufficient to ensure normal development. Furthermore, given the DLM phenotype observed was only in the small proportion of surviving *HDAC4*$^{\Delta att}$ mutants, it could also be a non-specific effect induced by the general poor health of these mutants, rather than a direct result of HDAC4 loss in the muscle. To address this, my *HDAC4*$^{\Delta att}$ is also capable of utilising a conditional knockout approach, by utilising a rescue (RIV) vector with rescue DNA fragments, excisable by FRT-FLP recombination (Poernbacher et al. 2019). In this approach, *HDAC4* would only be knocked out in a specific tissue, at a particular time. While also retaining the power and versatility of the insertion-ready deletion allele system, this would complement existing approaches to generate tissue-specific CRISPR mutants (Meltzer et al. 2019; Port et al. 2020; Koreman et al. 2021; Port and Boutros 2022).

Significantly, *HDAC4*$^{\Delta att}$ could also be used to further define the role of *HDAC4* during *Drosophila* muscle development, and in particular, to investigate its potential role in the context of Mef2 regulation. Indeed, the *HDAC4*$^{\Delta att}$ larval muscle phenotype could also be induced by *Mef2* overexpression. Although inconclusive, this could suggest a functional link between HDAC4 and Mef2 in the larval somatic musculature *in vivo*, warranting further investigation. To support this possibility, loss of Class IIa HDAC function in some vertebrate tissues have been associated with the Mef2 hyperactivation (Chang et al. 2004; Vega et al. 2004; Chang et al. 2006; Arnold et al. 2007). Specific changes, such as mutating the Mef2-binding domain, could be introduced into the gDNA rescue fragment to investigate this. Indeed, candidate residues which could be targeted in the endogenous locus, such as L168, have also been identified in this project using the Gal4/UAS overexpression strategy. However, given the lethality associated with *HDAC4*$^{\Delta att}$, such an approach may need to be incorporated into a conditional allele strategy (using this allele) to confine the desired mutation to developing muscle.

7.2.6 HDAC4 tagged alleles to probe protein expression, localisation, and interactions

It is also important to highlight the findings of the mScarlet-I tagged alleles, for which I have shown correct, in-frame insertion of the tagging cassette at both the N- and C-termini of the *HDAC4* locus. It is interesting that mScarlet-I-HDAC4 did not appear to act as a null allele and transcriptional reporter, given the allele was homozygous viable. Indeed, the N-terminus of *HDAC4* varies with different isoforms, and it is likely that the *HDAC4* locus mitigates the effects of the tagging cassette insertion through two possible ways to ensure sufficient *HDAC4* expression. This could occur through alternative splicing events may cause the skipping of exon two, as is annotated to occur in isoforms I and J. Or, secondly, transcription may be altered such that it is only initiated from the second possible TSS, as utilised by isoform B. Given the complexity of the *HDAC4* locus, this does emphasise the importance of generating my HDAC4-mScarlet-I C-terminally tagged allele.

Interestingly, initial analysis of HDAC4 expression in the L3 wing imaginal disc was inconclusive. Unlike with the analysis of Mef2 expression with Mef2-EGFP (see **7.6**), HDAC4 protein could not be detected in the absence of antibody staining. The AMPs

associate with the wing imaginal disc were chosen as a model to test whether any particular subcellular localisation could be observed in an undifferentiated myoblast pool *in vivo*. This is because the model based on *in vitro* analysis of Class IIa HDAC function would suggest that, if important for negatively regulating Mef2, HDAC4 would be predominantly nuclear in myoblasts, before being exported to the nucleus upon differentiation (McKinsey et al. 2000a). However, no obvious pattern of HDAC4 expression was observed and instead appeared to be pan-cellular. Furthermore, the difficulties associated with the detection of HDAC4 protein, which required both antibody staining and confocal microscopy, would indicate that HDAC4 may be lowly expressed in the AMPs. It is unclear as to why the signal obtained when staining for HDAC4 using HDAC4-mScarlet-I was variable, however could likely be a technical challenge since RNA-seq data confirms that *HDAC4* is expressed in the AMPs (Spletter et al. 2018; Zappia et al. 2020). Specifically, while GFP antibodies efficiently bind the YFP variant that tags HDAC4 in the CPTI line, it is not known how well the RFP antibody used binds the mScarlet-I fluorophore. Furthermore, no commercially available mScarlet-specific antibody currently exists. However, as HDAC4 appeared to be relatively lowly expressed, small variations in binding efficacy may have significant effects on protein detection.

The ultimate goal of these tagging approaches was not just to visualise HDAC4 expression and localisation, but also its physical relationship with Mef2. The fluorophore choice was chosen for its potential application in FLIM-FRET microscopy (**F**luorescence **l**ifetime **im**aging of **f**luorescence **r**esonance **e**nergy **t**ransfer), a high-resolution microscopy technique to quantify protein-protein interactions at distances fewer than 10nm (Wang et al. 2019). Indeed, due to their spectral overlap, GFP-RFP FRET pairs could be suitable for the quantification of protein-protein interactions *in vivo* (Lam et al. 2012; Bajar et al. 2016; McCullock et al. 2020). Subsequently, FLIM-FRET may be possible to investigate possible physical interactions *in vivo* between Mef2 (Mef2-EGFP) and HDAC4 (HDAC4-mScarlet-I).

7.3 *nejire*: complex roles in muscle differentiation *in vivo?*

By contrast to Class IIa HDACs, the p300/CBP family of HATs have previously shown positively regulate Mef2 transcriptional ability *in vitro* (Sartorelli et al. 1997; Ma et al. 2005; Angelelli et al. 2008). This could suggest a mechanism whereby p300 may stimulate Mef2 activity and contribute to activation of the muscle gene expression

program. However, these studies were limited to the study of p300 *in vitro*. Therefore, I wanted to investigate the role of the p300 homologue, nej, in an *in vivo* model for muscle differentiation. Indeed, *nej* is essential for normal muscle development *in vivo* s its knockdown inhibited DLM formation. I did not undertake qPCR analysis to confirm *nej* knockdown, however, the formation of an allelic series using two independent RNAi lines supports the requirement of nej for normal DLM development *in vivo*. Moreover, the stronger TRiP line, BL37489 has been previously validated (Jia et al. 2015), and has the most citations attributed to its use on flybase (Gramates et al. 2022). These data also support a more recent study in human cells, indicating p300/CBP play important roles in the activation of the myogenic gene expression program during the differentiation of primary human myoblasts (Fauquier et al. 2018).

Interestingly, a previous report in *Drosophila* cells *in vitro* suggests nej acts as a Mef2 transcriptional corepressor by inhibiting the expression of another Mef2 target gene, *pumilio* (Lin and Baines 2019). However, since *nej* overexpression did not inhibit the Mef2-induced premature differentiation of L3 AMPs, it does not appear to be functioning as a transcriptional corepressor during muscle differentiation *in vivo*. Having said that, nej is probably playing a more global role in the muscle differentiation pathway, rather than only influencing Mef2 target gene expression. *nej* overexpression also significantly inhibited normal DLM development *in vivo* and yielded a distinguishable phenotype from Mef2 overexpression alone. Therefore, this phenotype is not a result from Mef2 hyperactivation alone. Moreover, despite its canonical roles as a transcriptional coactivator, p300 may also have transcriptional repressive roles: sumoylation of vertebrate p300 converts it into a transcriptional co-repressor (Girdwood et al. 2003), while in *Drosophila*, nej has recently been shown to contribute to maintenance of transcriptional repressive states by polycomb-group proteins (Hunt et al. 2022). Given the establishment and growth of next generation sequencing technologies, RNA-seq following nej dysregulation may be the most appropriate method to investigate its transcriptional effects during the muscle differentiation pathway, especially given its role as a promiscuous cofactor with many potentially many binding partners (Bedford et al. 2010; Dancy and Cole 2015).

However, answering the question of whether nej is required for Mef2 function during muscle differentiation *in vivo* proved challenging, since: *nej* knockdown also inhibited establishment of the AMP pool associated with the L3 wing imaginal disc, which by

contrast is not caused by *Mef2* knockdown. This may be due to Nej possessing multiple roles throughout muscle differentiation, and not exclusive to Mef2 target gene activation. In particular, notch signalling is critical for DLM development in *Drosophila*, and loss of notch signalling inhibits the proliferation of the AMPs during larval stages (Anant et al. 1998; Gunage et al. 2014). Moreover, by interacting with the notch intracellular domain, vertebrate p300 has been shown to be an important coactivator in coordinating the transcriptional response to notch signalling (Oswald et al. 2001; Wallberg et al. 2002; Bray 2016; Sachan et al. 2023). This could explain why *nej* knockdown also lead to ablation of the AMP pool. Secondly, it could be that Nej is also required for the activation of Mef2. In vertebrates, Mef2 expression is directly activated by the MRFs, and the MRFs require p300 for their transcriptional activity (Sartorelli et al. 1997). It could be that Nej also functions to activate twist transcriptional activity *in vivo*. Ultimately, this could mean Nej may have roles upstream of Mef2 in the myogenic pathway.

7.4 The Mef2 domain: A common binding interface for coactivators and corepressors in *Drosophila*?

What led me to initially investigate Nej function during *Drosophila* muscle differentiation was the observation that mutating the Mef2 domain had differential consequences on the ability for Mef2 to induce premature differentiation of L3 AMPs. Importantly, the Mef2 domain contains the binding interface known to interact with both the p300 coactivator and Class IIa HDAC corepressors (Han et al. 2005; He et al. 2011). Furthermore, given the prior understanding of the p300 and Class IIa HDACs, this domain could be important for impacting Mef2 transcriptional activity depending on the which coregulator is bound (**Fig. 1.6**). My work furthers previous *in vitro* observations, whereby a VLL65-67ASR triple mutation inactivated Mef2C (Molkentin et al. 1996a). This phenotype was caused despite Mef2C retaining the ability to both dimerize, and bind to DNA. However, prior to this project, there had been little investigation into the functional importance of this domain in an *in vivo* model of differentiation.

In agreement with the functional conservation between *Drosophila* and vertebrate Mef2 proteins, the equivalent UAS-Mef2 VLL65-67ASR mutant also lacked the ability to induce premature differentiation of L3 AMPs in *Drosophila*. Here, I have furthered the understanding of the functional residues in this domain, showing that mutation of

L66, but not L67, completely abolishes Mef2 transcriptional activity *in vivo*. Moreover, I show that mutation of Y69 also reduces Mef2 transcriptional activity, suggesting this residue is also important for normal Mef2 function.

The important question in this context is, what is the functional significance of this domain, and why do some mutations affect Mef2 function more than others? Of course, I have not shown these mutations retain DNA binding and dimerization ability in *Drosophila*. Given the conservation between vertebrate and *Drosophila* Mef2, it is likely that these would retain the ability, as they do with Mef2C *in vitro* (Molkentin et al. 1996a). Having said that, recent molecular dynamics simulations of Mef2B indicate that a Y69H mutation may induce conformational changes which have subtle effects on both DNA-binding, and cofactor interactions (Zia and Rashid 2021). This suggests that although mutations in this domain do not completely abrogate Mef2 DNA-binding, there could be more subtle effects on Mef2 structure and function which could not previously be identified by Molkentin et al.

Given this region of Mef2 interacts with both Class IIa HDACs, and p300/CBP, it could be that mutations affect Mef2-coregulator interactions required for Mef2 transcriptional activity. However, as yet during this project, the data does not show any conclusive evidence whether there is differential binding of these coregulator proteins *in vivo*, and whether this directly affects Mef2 transcriptional activity. Thus, the immediate next goal would to investigate the effect of these mutations on the ability for Mef2 to interact with both HDAC4, and Nej. This would answer whether the mutations differentially affect Mef2's ability to physically interact with either coregulator, or whether the critical residues required for both interactions are shared. In vertebrate studies *in vitro*, mutations of L66, L67 and Y69 have all been shown to affect the Class IIa HDAC-Mef2 interaction (Han et al. 2005; Jayathilaka et al. 2012). By contrast, biochemical analysis of the relative importance of these residues for the Mef2-p300 interaction has not been undertaken (He et al. 2011). However, this question may in part be separate from that of why some Mef2 mutations abolish Mef2 activity: it is unlikely that the L66 mutation inactivates Mef2 directly as a result of an inability to bind HDAC4, given HDAC4 negatively regulates Mef2. It could therefore be more likely that these mutations prevent Mef2-Nej interactions, which may be required for Mef2 function.

7.5 Mef2-induced premature differentiation: a model system for studying Mef2 function *in vivo*.

7.5.1 Mef2 dysregulation affects the muscle differentiation gene expression program

A key result from this work, following on from previous observations (Lovato et al. 2005; Soler and Taylor 2009), is that the Mef2 overexpression disrupts the myogenic program by inducing premature gene expression and differentiation. Mef2 overexpression in undifferentiated AMPs associated with the L3 wing imaginal disc induced the premature expression of Mhc, an integral sarcomeric protein not normally expressed until approximately 30hr APF onwards (Spletter et al. 2018). Unlike previous work (Lovato et al. 2005; Soler and Taylor 2009), a striated appearance of Mhc was observed, suggesting that the AMPs do not only express terminal differentiation markers, but have also begun to mature and undergo physiological changes associated with muscle differentiation, notably sarcomerogenesis. Indeed, while not done by myself, further characterisation of this phenotype by Rob Mitchell also characterised sarcomere formation through staining for additional components, including the Z disc proteins Zasp66 and α-actinin.

One interesting observation associated with this phenotype is its spatial distribution. Mhc expression appeared generally restricted to the AMPs located in the most dorsal part of the wing disc notum. Why is it, therefore, that only a portion of AMPs express Mhc and differentiation prematurely? Moreover, given that DLM formation is only subtly affected when Mef2 is overexpressed, this suggests the majority of AMPs remain undifferentiated and retain the ability to fuse to LOM templates, since a significant reduction to the undifferentiated AMP pool would likely disrupt DLM development more severely (Vishal et al. 2017).

One reason or this phenotype could be related to the heterogeneous nature of the AMP population, recently identified (Zappia et al. 2020). They identified 5 distinct clusters of AMPs believed to at different stages of the differentiation pathway. It may be that some AMPs are more ‘primed’ to differentiate, and therefore are most vulnerable to Mef2-induced differentiation. This may make sense given that some clusters display higher levels of active notch signalling, which inhibits differentiation, as well as other repressors of muscle differentiation such as Zfh1 and Him (Soler and

Taylor 2009; Boukhatmi and Bray 2018; Zappia et al. 2020). Furthermore, both Zfh1 and Him have been shown to repress Mef2 activity (Postigo et al. 1999; Liotta et al. 2007; Soler and Taylor 2009). This could mean that, even in the presence of excess Mef2, the balance of positive and negative regulators of differentiation in different clusters may determine the ultimate fate of the AMPs.

Secondly, it could be influenced by Mef2 autoregulation, whereby Mef2 has been shown to activate the expression of negative regulators of Mef2 including the aforementioned Zfh1 (Postigo et al. 1999), microRNAs (Chen et al. 2012), Him (Sandmann et al. 2006), and Class IIa HDACs (Haberland et al. 2007). it is plausible these mechanisms may limit the penetrance of Mef2 overexpression in the AMPs.

Finally, there may also be some more general limits to Mef2 transcriptional activity. For example, the local chromatin environment not be permissive to Mef2 binding. Conversely, endogenous Mef2 in AMPs may already be occupying many of its target sites. If this were the case, then the question would be: why is it not activating the expression of target genes? It could be the absence of a necessary cofactor required for Mef2 target gene activation, target genes requiring multiple regulatory inputs (i.e bound TFs) to coordinate their expression, or other regulatory mechanisms, such as PTMs and negative regulator complexes bound to Mef2 which restrict its activity. For example, in the embryo, Mef2 expression can drive the ectopic expression of some target genes in the ectoderm on its own, while other target genes require the co-expression of another transcription factor, lame-duck (Sandmann et al. 2006). Moreover, Mef2 cooperates with another transcription factor, Chorion-factor 2, in the activation of muscle gene expression in the embryo (Tanaka et al. 2008). Conversely, one recent result also reported in that a Mef2 S98A phospho-null allele displayed also resulted in premature Mhc expression in the AMPs, suggesting that PTMs may also play a significant role in modulating endogenous Mef2 activity *in vivo* (Vishal et al. 2023).

These data emphasise that understanding of Mef2 target gene activation during adult muscle differentiation remains relatively understudied. Given its integral role in adult muscle differentiation, high-throughput sequencing should be adopted to obtain a more global understanding of Mef2 function *in vivo*. Furthermore, given the phenotype induced by Mef2 overexpression, the L3 AMPs in the wing imaginal disc could be a suitable model to probe Mef2 function and regulation during adult myogenesis.

7.5.2 The premature differentiation assay as a model to probe Mef2 function

The premature differentiation assay also proved useful as an assay to probe the relationship between Mef2 structure and function. However, while this was a clear-cut and easily interpretable result when investigating binary phenotypes, further work could further refine the assay to gain more quantifiable insights into Mef2 activity. For example, in situations where the ability to induce premature differentiation of the AMPs is retained, but to a greater, or lesser extent than wild-type. Firstly, while UAS-Mef2 L67A retained the ability to induce premature differentiation in 100% of samples, it is not fully understood whether the transcriptional activity of Mef2 is equivalent. Similarly, the relative transcriptional activity of UAS-Mef2 10t4a and UAS-Mef2 CDS.HA, given then former includes 3'UTR region known to possess an miRNA target sequence (Chen et al. 2012). Thirdly, in the quantification of the effect additional transgenes on Mef2 transcriptional activity, such as with nej co-expression.

There are two possibilities of quantifying Mef2 transcriptional activity *in vivo* using this approach. Firstly, by direct quantification of the amount of GFP in each disc. However, this would possibly require confocal analysis to enable signal quantification through the entire sample, rather than just across a single plane of focus when detecting standard fluorescence on the Olympus BX53 compound microscope; or secondly, a biochemical approach by quantifying the activity of a lacZ reporter gene under the regulation of a Mef2-dependent promoter(Tanaka et al. 2008; Chechenova et al. 2015). The former may be preferential over the latter due to the ability to retain the spatial information associated with the Mef2-dependent phenotype in the disc.

However, there are challenges associated with this assay that researchers must be aware of, especially when utilising it to probe Mef2-interacting proteins. For example, one could attempt to utilise this assay to investigate the hypothesis that UAS-HDAC4 mutants defective in Mef2 binding should not be capable of suppressing the premature differentiation phenotype. However, given that Mef2 overexpression induced AMP terminal differentiation, and overexpression of UAS-HDAC4 mutants could still inhibit terminal differentiation of the DLMs when overexpressed using the Actin88F-Gal4 driver, this could complicate the application and interpretation of the assay. Equally, as could the investigation of Mef2-interacting factors with multiple roles in muscle differentiation such as Nej. For example, without assessing for the presence of the

AMPs, one may have concluded that nejire is required for Mef2-induced premature differentiation of the AMPs. However, in this case, where nej appears to be required for establishment of the AMP population, this assay may not be the most appropriate and emphases the challenges associated with investigating gene function *in vivo*.

7.6 Engineering the *Mef2* locus: Critical tools for future investigation

7.6.1 Mef2-EGFP

A major goal of the project was to develop CRISPR-based tools to facilitate furthering our understanding of Mef2 function *in vivo* through targeted engineering of the endogenous *Mef2* locus. Indeed, when my project first started, there had ben no published CRISPR mutant of *Drosophila Mef2,* and the recently published phospo-mutant was published after my tools were generated (Vishal et al. 2023).

Mef2-EGFP has been a major development for the field. Generally speaking, the analysis of protein localisation *in vivo* is limited to those proteins for which there is an antibody. Only approximately 450, or 3% of all protein-coding genes in *Drosophila*, are afforded such a luxury (Nagarkar-Jaiswal et al. 2015). Moreover, even when antibodies are available, their large size (approx. 150-160kDa) affects the spatial resolution obtainable (Xu et al. 2022). Of course, Mef2 does have an available antibody. However, while this has been beneficial, there are limitations. Firstly, it has obviously been limited to studying Mef2 expression in fixed tissue. Secondly, the antibody itself was also generated by a single lab, and its availability is probably finite. The development of Mef2-EGFP was designed to address both of these issues. Mef2-EGFP yielded an incredibly bright signal which could be detected across multiple stages of *Drosophila* development, in the absence of fixation or antibody procedures. This achievement should not be understated, for two main reasons. Firstly, it was unknown whether the 26.9kda EGFP fluorophore would disrupt Mef2 function *in vivo.* Secondly, the presence of a fluorophore does not necessarily guarantee the ability to visualise endogenous protein without antibody staining. For the former, although one could further confirm endogenous Mef2 function by carefully analysing muscle structure and function, its homozygous viability is extremely convincing. For the latter point, this is emphasised not only in the study of HDAC4, which neither I, or others

(Fitzsimons et al. 2013) could detect in the absence of antibody staining, but also for other proteins, many of which also cannot be detected in the absence of staining (Hens and Kittelmann, Personal communication). Unlike other CRISPR lines generated in this project, I did not sequence verify the insertion due to the compelling evidence from expression analysis, however this of course could be readily obtained if desired.

It would be interesting to also tag the N-terminus of Mef2 as well. The aim of such an experiment may be more aligned to investigating Mef2 structure-function, given the N-terminus' close proximity to both the MADS/Mef2 domain, and to the bound DNA in crystal structures (Han et al. 2005; He et al. 2011). Having said that, while Mef2-EGFP tags the majority of isoforms, isoform H is annotated on Flybase to utilise a STOP codon downstream of the one predominantly used as a result of stop-codon suppression (Gramates et al. 2022). However, nothing is annotated or published as to the expression or function of this isoform. One could investigate the expression of this isoform specifically by generating an isoform H-specific C-terminal EGFP tag. Alternatively, because the N-terminus is common to all isoforms, a functional N-terminally tagged Mef2 would ensure all isoforms are tagged.

In any case, there are some immediate possibilities for future investigation into Mef2 function *in vivo* with Mef2-EGFP. Firstly, Mef2 protein dynamics can be followed in an individual, live sample over time. Secondly, a major interest will be top further our understanding of Mef2-interacting proteins. This could firstly be achieved through immunoprecipitation experiments to pull down endogenous Mef2 using the GFP tag. Of course, this can in theory be done using the existing Mef2 antibody, however may be technically easier using commercially available antibodies against GFP, especially when attempting to pull down endogenously expressed protein. Indeed, a relatively recent coIP experiment aiming to pull down endogenous protein required a sample of over 1'200 wing discs, emphasising the challenge associated with generating enough tissue for such experiments (Nakajima et al. 2019). Equally, FLIM-FRET microscopy could be used to probe protein-protein interactions *in vivo*, and has been applied to *Drosophila* muscle biology (Lemke et al. 2019). However, this of course relies upon endogenously tagged candidate interacting proteins with complimentary fluorophores. For example, the complimentary HDAC4 and Mef2 protein fusions could utilised to probe potential HDAC4-Mef2 interactions *in vivo* using FLIM-FRET.

7.6.2 Mef2 insertion-ready deletion alleles

A second major aim was to generate versatile platform which could be used to investigate the function and regulation of endogenous Mef2. Unlike *HDAC4,* there is a greater amount of information surrounding critical domains for Mef2 function, which made satisfying the first criteria, the generation a null allele, less complex. In addition to removing the entire CDS, I found that deletion of only a 250bp region (in $Mef2^{\Delta ATG}$) containing the coding sequence for 18 amino acids for the MADS-Mef2 domain, was sufficient to generate a *Mef2* null. Possibly this is not surprising, given that an in-tact MADS/Mef2 domain is critical for Mef2 function (Molkentin et al. 1996a; Nguyen et al. 2002). Interestingly however, I saw variable results in terms of the ability to rescue $Mef2^{\Delta CDS}$ and $Mef2^{\Delta ATG}$. One allele, $Mef2^{\Delta ATG\text{-}cDNA}$ was successfully rescued to homozygous viability, which also displayed phenotypically normal embryonic somatic musculature and DLM fibres. Taken together, these results indicate successful rescue of the $Mef2^{\Delta ATG}$ null allele. Moreover, given that this cDNA encodes only a single Mef2 isoform (isoform C), it also suggests that a single *Mef2* isoform may be sufficient to recapitulate the majority of Mef2 function *in vivo*, and possibly reiterates the question of the functional roles of different *Mef2* isoforms *in vivo*. Indeed, this is a significant result, and could allow the utilisation of the $Mef2^{\Delta ATG}$ for more functional analysis, such as cDNA rescue with different isoforms, or the mutagenesis of particular residues of Mef2 implicated in regulating its function. For example, mutations resulting in differing ability to induce premature differentiation of L3 AMPs could be incorporated into $Mef2^{\Delta ATG}$ cDNA rescue alleles to further understand their functional significance in terms of endogenous Mef2 function.

What remains interesting however is why none of the other alleles appeared to rescue *Mef2* function, given they were not homozygous viable and Mef2 function is essential for viability. Of course, no additional phenotyping was undertaken with these alleles to assess for any partial restoration of function. Neither, as for $Mef2^{\Delta ATG\text{-}cDNA}$, were the recued alleles molecularly verified by PCR, however the red eye phenotype in the stable balanced stock does reliably inform of integration of the RIV^{white} construct. There are multiple possibilities for the lack of efficient rescue. For the cDNA rescue of $Mef2^{\Delta CDS}$, the loss of intronic sequences could indicate important regulatory sequences are located within them. For $Mef2^{\Delta ATG\text{-}gDNA}$, the excess integrated marker DNA could initially have disrupted mRNA processing, although cre-mediated excision

of this DNA also did not appear to restore Mef2 function. In theory, it could be the residual loxP site within the second intron (of isoform C) may introduce a cryptic splice site which could equally disrupt mRNA processing. Finally, the initial 5' CRISPR site placed within the 5' UTR, common to both *Mef2*$^{\Delta CDS}$ and *Mef2*$^{\Delta ATG}$ could cause general disruption to Mef2 expression, such as through affecting ribosomal entry (Poernbacher et al. 2019). However, this would not explain the successful rescue with *Mef2*$^{\Delta ATG\text{-}cDNA}$. Thus, it remains relatively unknown the exact mechanisms by which Mef2 function is not recued in some lines, and emphasises the challenges associated with engineering endogenous loci.

7.7 Outlook and concluding remarks

The work undertaken in this thesis has its central focus towards understanding gene expression regulation during somatic muscle differentiation *in vivo*. I have focused on the master transcription factor Mef2, who's activity must be regulated in order to correctly coordinate gene expression throughout myogenesis. While many regulatory mechanisms have been identified, this work has been directed towards investigating how Mef2-protein interactions may contribute to this regulation. My work has demonstrated for the first time *in vivo* that the *Drosophila* Class IIa HDAC, HDAC4 can negatively regulate Mef2 in undifferentiated myoblast populations during muscle differentiation *in vivo*, causing severe muscle phenotypes when overexpressed. Of course, while no experiments, such as immunoprecipitations or Yeast-2-Hybrids were undertaken to confirm the effect on Mef2 binding in the biological context here, these data fit appropriately with the existing data on the Mef2-Class IIa HDAC physical interaction, and the existence of a recognised UAS-HDAC4 Mef2 binding mutant being available in the published literature (Main et al. 2021). Of course, one could perform the aforementioned physical interaction experiments were it to be desirable to bolster these conclusions. Although, in the context of this project, it was, and remains most appropriate to address *HDAC4* function in the context of the endogenously expressed gene, as opposed to the phenotypic effects of overexpressing different UAS-HDAC4 constructs in developing muscle.

Nevertheless, it remains important to reiterate the fundamental goal of developmental geneticists: to establish whether a gene product *does* function in a particular context during normal development, as opposed to whether it *can*. Indeed, in this project I have developed valuable genetic tools, primarily for the study of HDAC4 and/or Mef2

function, using CRISPR-Cas9, while also emphasising some of the challenges associated with engineering endogenous loci. The latter is of particular significance when developing insertion-ready deletion alleles. I have developed the first rescuable *HDAC4* null allele, which benefits over more classical gene traps which are not full LOF, and *HDAC4* CRIMIC/MiMiC lines which do not allow the flexibility in terms of genomic rescue that my allele allows (Bellen et al. 2004; Thibault et al. 2004; Kanca et al. 2019). This is a major advantage of implementing the insertion-ready deletion allele system, and will continue to be used extensively in the lab (along with my tagged alleles) in order to further investigate the potential role of *HDAC4* during muscle differentiation. However, although effective in the context of engineering the *HDAC4* locus, its suitability as a reliable tool for high throughput analysis of gene function more generally remains unclear, highlighted through its inconsistent results observed when engineering the *Mef2* locus. Coupled with the extensive time and financial commitments required to establish the platform for each gene, this will mean researchers contemplating this technology should consider the probability of success, with the impact and importance of such a tool in the short-to medium term. CRISPR-Cas9 and its derivatives however remain the most widely used and effective tool for genome engineering, and continues to have a significant impact on researchers' ability to understand endogenous gene function *in vivo*.

Finally, my observations that mutations in the HDAC4-interacting domain of Mef2 affect Mef2 transcriptional activity *in vivo* could suggest that this domain is required for coordinating other protein-protein interactions necessary for Mef2 transcriptional activity. This is because disrupting the binding of exclusively a negative regulator (HDAC4) would unlikely abolish the ability for Mef2 to direct target gene activation. Indeed, my additional work identified that the p300 homologue *nejire* is necessary for muscle differentiation *in vivo*, a critical point when understood that Nej also physically interacts with Mef2 via the same interface as HDAC4. Indeed, these data could possibly point towards this region of Mef2 being important for regulating Mef2 through the recruitment of Nej, and possibly other coregulator proteins more generally. Of course, the promiscuous nature of Nej, meaning it likely interacts with many other proteins as well as Mef2 ultimately caused challenges in this work when investigating a possible direct functional link between Mef2 and Nejire during muscle differentiation *in vivo*. Although many interacting partners of Mef2 have thus far been identified in

different contexts, one could undertake a higher throughput approach to screening Mef2-interacting proteins and their functions, using approaches such as Yeast-2-Hybrid, and high throughput sequencing to investigate genome-wide occupancy of Mef2, and potential interacting partners *in vivo*. The latter remains particularly attractive in the case of Mef2, which activates the expression of many different genes at different developmental stages. Finally however, although protein interactions are of high importance for Mef2 function and regulation, it remains likely that multiple different mechanisms converge in order to correctly regulate Mef2 transcriptional activity during muscle development *in vivo*.

Appendix

Name	Sequence (5'-3')	Purpose
Insertion-ready deletion (ΔAtt) alleles		
Mef2		
Mef2_5_CDS/Wg_CRISPR-F	CCCATGGCAACAGATTTC	PCR and sequence verification of CRISPR-Cas9 target sites in the yw;;nos-cas9(III-attP2) injection line.
Mef2_5_CDS/Wg_CRISPR-R	CCTGGGGACAATTTATAC	
Mef2_3_CDS_CRISPR-F	GCTCAAATTTCTCGGCAG	
Mef2_3_CDS_CRISPR-R	CTCGTAAGGACATAGGTC	
Mef2_3_Wg_CRISPR-F	GCCATGAATCGATTCACC	
Mef2_3_Wg_CRISPR-R	CAAATATCACGCATCACC	
Mef2_5HA_CDS_F-NotI	ATAATGCGGCCGCGAATTTGTTGAGTGTGTC	Amplification of 5' homology arm upstream of the 5' Cas9 site used in both CRISPR designs for the insertion of the Att cassette
Mef2_5HA_Wg_F-NotI	AGCTAGCGGCCGCTGAAGTTAAATTAACCTC	
Mef2_5HA_CDS/Wg_R-NheI	GATTAGCTAGCGTTGCTGCCCTGCGAATT	
Mef2_3HA_CDS_F-SpeI	AATTAACTAGTGTCCTCCGTTTCCGGACC	Amplification of 3' homology arm downstream of 3' Cas9 site used in the removal of entire coding sequence (CDS) CRISPR approach
Mef2_3HA_CDS_R-AatII	ATTATGACGTCCGGTCTTGGAGTTACTGG	
Mef2_3HA_Wg_F-SpeI	GCCGGACTAGTCGTAGCTATAATAATACA	Amplification of 3' homology arm downstream of 3' Cas9 site used in the 'Wg' style CRISPR approach
Mef2_3HA_Wg_R-AatII	ATTATGACGTCTTGTACTCGGTGTACTTG	
Mef2_gRNA_5'_CDS/Wg	TATATATAGGAAAGATATCCGGGTGAACTTC GATTCGCAGGGCAGCAACTAAGTTTTAGAGC TAGAAATAGCAAG	
Mef2_gRNA_3'_CDS	ATTTTAACTTGCTATTTCTAGCTCTAAAA	

	CCGCGTCCTCCGTTTCCGGACGACGTTAAA TTGAAAATAGGTC	Amplification and cloning of gRNA sequences into pCFD4
Mef2_gRNA_3'_Wg	ATTTTAACTTGCTATTTCTAGCTCTAAAA CATCCGTAGCTATAATAATACGACGTTAAA TTGAAAATAGGTC	
Mef2 CRISPR – rescue constructs		
Mef2_rscu_5'UTR-F+EcoRI	AGGGCGAATTCTAACGGAGTTTCCCCG	PCR amplification of fragments for rescue of Mef2 insertion-ready deletion alleles
Mef2_gDNArscu_CDS-R+XhoI	GGCCGCTCGAGATGTGAAATGTACAGAGAGC	
Mef2_gDNArscu_ATG-R+XhoI	TTATACTCGAGGATTGGCAAACACATTGTGAC	
Mef2_cDNArscu_CDS-R+XhoI	ATGAACTCGAGAATCTCAGTAGAGCG	
Mef2_cDNArscu_3'UTR-F+XhoI	CTGAGATTCTCGAGTTCATCATCGAACC	
Mef2_cDNArscu_3'UTR-R+KpnI	GTCAGGTACCCAGTTGCCAAAAAGG	
Mef2 CRISPR – Sequencing primers		
Mef2_rscu_Seq1	ATGAAGAAGGCCTACGAG	
Mef2_rscu_Seq2	TCGATGATGTTCTTGTTG	
Mef2_rscu_Seq3	GAGTGTGAATGCGGATAG	
Mef2_rscu_Seq4	AGACATCGCTTAACACGC	
Mef2_rscu_Seq5	TCAGCAACAGCAACTAGG	
Mef2_3UTR-F	CACGCTCTACTGAGATTC	
Mef2_5HA_seq1	TGGAGCTGATGAAATGGC	
Mef2_5HA_seq2	ATTGGAATGCTGTGTTGC	
Mef2_5HA_seq3	TTCAATGCACCGCTAATG	
Mef2_5HA_seq4	GGGATAGATCTAGGGTCC	
Mef2_3HA_CDS_seq1	AACACTGAGTACATCTGC	
Mef2_3HA_CDS_seq2	AACCATGCAAGTCATACG	

Mef2_3HA_CDS_seq3	TCGAAATGGGCAATTGAC	
Mef2_3HA_CDS_seq4	TTCACATTTCACTCGGCC	
Mef2_3HA_Wg_seq1	ACGATGACGCGTTTATGG	
Mef2_3HA_Wg_seq2	CGACTTTTCGTTCCTTTC	
HDAC4 CRISPR – Gene targeting		
HDAC4 5_Att_CRISPR-F	CTTATCGTGGTGGAGAAG	PCR and sequence verification of CRISPR-Cas9 target sites in the yw;;nos-cas9(III-attP2) injection line.
HDAC4 5_Att_CRISPR-R	GGTGACGATGATAGTGTG	
HDAC4_3_Att_CRISPR-F	CTCGATGACTTCTACAAC	
HDAC4_3_Att_CRISPR-R	TACAGAACTCTATCCCGC	
HDAC4_NT_CRISPR-F	GGCATAGGATTTTCTAGG	
HDAC4_NT_CRISPR-R	GCTACAAGACTCTTCTAC	
HDAC4_5HA_F-NotI	CCAAAGCGGCCGCAGAAACCAAAGATTA	Amplification of 5' and 3 homology arms for cloning into pTV3 targeting vector
HDAC4_5HA_R-NheI	CTCCGGCTAGCGATAGGTTTGTATTGATA	
HDAC4_3HA_F-AatII	ATGCAGACGTCTATTGCGGAAGAGGAG	
HDAC4_3HA_R-AgeI	CCAACCGGTCCATCTCTCGTTGTGCT	
HDAC4_gRNA_5'	TATATATAGGAAAGATATCCGGGTGAACTTC GTATATATGGGTAATGCAGATGTTTTAGAGCTAGAAATAGCAAG	Amplification and cloning of gRNA sequences into pCFD4
HDAC4_gRNA_3'	ATTTTAACTTGCTATTTCTAGCTCTAAAA CCAATATACGTCTGCATCGCCGACGTTAAATTGAAAATAGGTC	
HDAC4 CRISPR – Rescue constructs		
HDAC4_gDNArscu_F+NotI	ATACAGCGGCCGCTGCATTACCCATATATAC	PCR amplification of gDNA construct for HDAC$^{\Delta att}$ rescue
HDAC4_gDNArscu_R+NheI	ACCACGCTAGCCCATTCTGTATAATTAC	
HDAC4 CRISPR – Sequencing primers		
HDAC4_5HA_SEQ1	CGATTTGTACGTGAATGC	

HDAC4_5HA_SEQ2	GTATGTACGTTGACGGTC	
HDAC4_5HA_SEQ3	ACGTTAGTTATCGGCGATCG	
HDAC4_3HA_SEQ1	TGGTGAATACACATGTGC	
HDAC4_3HA_SEQ2	GCCGAGATCTTATATACG	
HDAC4_3HA_SEQ3	CAACCTCAAAGTTCCCAAGTC	
HDAC4rscuSeq1	ACCCTCACTGCTCAAATATG	
HDAC4rscuSeq2	ATACCCAACTAACCCAACAC	
HDAC4rscuSeq3	GGCTACTATAATCCACTGGG	
HDAC4rscuSeq4	GTCTTAGTCAGATTCCCGAC	
HDAC4rscuSeq5	TGGTCTCGGCTTTAACGTG	
HDAC4rscuSeq6	GATCGTCTATAGCCACATC	
Him CRISPR – gene targeting		
Him_Att_CRISPR-F	AGTTTGGCGCGCAATGTG	PCR and sequence verification of CRISPR-Cas9 target sites in the yw;;nos-cas9(III-attP2) injection line.
Him_Att_CRISPR-R	GAACGAAGGCAGATGGAG	
Him_Att_5HA_F+EcoRI	GATTGGAATTCATCTTGCGGCACTACGAT	Amplification of 5' and 3 homology arms for cloning into pTV3 targeting vector
Him_Att_5HA_R+KpnI	ATTAGGTACCTGCGAGCGCGTCGAGC	
Him_Att_3HA_F+AatII(2)	CCATTTTCCGACGTCTCTGCAAACCGAC	
Him_Att_3HA_R+AgeI	AGCAACCGGTACTGTTCAGTGATGTC	
Him_gRNA_5	TATATAGGAAAGATATCCGGGTGAACTTCGAGCTCGAC GCGCTCGCATATGTTTTAGAGCTAGAAATAGCAAG	Amplification and cloning of gRNA sequences into pCFD4
Him_gRNA_3	ATTTTAACTTGCTATTTCTAGCTCTAAACCTGCAAAC CGACTTTATTTTGACGTTAAATTGAAAATAGGTC	
Common sequencing primers for CRISPR constructs		

M13uni-43	AGGGTTTTCCCAGTCACGACGTT	
pTV3_5_Arm-R	GAGAACTCAAAGGTTACC	
M13rev-29	CAGGAAACAGCTATGACC	
pTV3_3_Arm-F	CAATGTATCTTAGCCCCG	
pCFD4_fwd	GACACAGCGCGTACGTCCTTCG	
pCFD_rev	GCCTACCTGGAGCCTGAGAGTT	
RIV_fwd	CAATTCCGCTCGAATTGC	
RIV_rev	AGAGAATAACGGGGCATG	
pUAST-attB constructs		
HDAC4		
UAS-HDAC4_CDS_F-BglII	ATCGAAGATCTATGTCTAGTCCCGACGAT	Amplify HDAC4 CDS from FI19806 cDNA for cloning into pUAST-attB
UAS-HDAC4_CDS_R-XbaI	GTCGATCTAGACTATTTGGTTTCATCCTG	
HDAC4_SDM_LEGGY/H-F	CGGCCAGATCGTGGCCGCCCTCGAG	Site-directed mutagenesis of HDAC4 cDNA and sequence verification
HDAC4_SDM_LEGGY/H-R	CTCGAGGGCGGCCACGATCTGGCCG	
HDAC4_SDM_F171A-F	GGACTTTCTGCTCAGGATGGCGCAGTTGAGAATCTGTTTG	
HDAC4_SDM_F171A-R	CAAACAGATTCTCAACTGCGCCATCCTGAGCAGAAAGTCC	
HDAC4_SDM_L168A-F	CCAGTCCCGAGGTCAAACAGATTGCCAACTGCTTCATC	
HDAC4_SDM_L168A-R	GATGAAGCAGTTGGCAATCTGTTTGACCTCGGGACTGG	
HDAC4_CDS_Seq1	CGAGAATAGTGAGGCGGAGC	
HDAC4_CDS_Seq2	GACGTACTGGTGGCTATCGT	
HDAC4_CDS_Seq3	AATTGGCCTCCGCATGATGG	
pUAST-Rev	CCATTCATCAGTTCCATAGG	
Mef2		
UAS-Mef2_CDS_F+EcoRI	CTACGGAATTCATGGGCCGCAAAAAAATTC	PCR amplification of 1090BP of Mef2 cDNA for cloning into pUAST-attB
UAS-Mef2_CDS_R+BglII	GGAAAGATCTCTCGGCGGCGACTG	
Mef2_SDM_L66A-F	CGACATGGATCGCGTCGCGCTCAAGTACACCGAG	
Mef2_SDM_L66A-R	CTCGGTGTACTTGAGCGCGACGCGATCCATGTCG	

Mef2_SDM_L67A-F	CATGGATCGCGTCCTGGCCAAGTACACCGAGTAC	Site-directed mutagenesis of Mef2 cDNA
Mef2_SDM_L67A-R	GTACTCGGTGTACTTGGCCAGGACGCGATCCATG	
Mef2_SDM_Y69A-F	TCGCGTCCTGCTCAAGGCCACCGAGTACAACGAG	
Mef2_SDM_Y69A-R	CTCGTTGTACTCGGTGGCCTTGAGCAGGACGCGA	
Mef2_SDM_VLL65-67ASR-F	GCACCGACATGGATCGCGCCTCGCGCAAGTACACCGAGTACAA	
Mef2_SDM_VLL65-67ASR-R	TTGTACTCGGTGTACTTGCGCGAGGCGCGATCCATGTCGGTGC	
Common sequencing primers for UAS constructs		
pUAST-Fwd	CAACTACTGAAATCTGCCAAG	
pUAST-ups UAS	TGCTTGGATTTCACTGGAAC	
pUASt_attB_fwd	CACTGCATTCTAGTTGTGGTTTGTCCAAAC	
pUAST-Rev	CCATTCATCAGTTCCATAGG	
Mini-White R	GCGATAAAGTCGACTTCG	
Direct tagging of HDAC4 and Mef2		
HDAC4		
HDAC4_NT_HR1F	GAAAGACTGGGCCTTTCGCCCGGGCTAACGT ACGAGTTTGATGGAATCCTTCTTACTTTG	PCR amplification of homology arms for direct tagging of HDAC4
HDAC4_NT_HR1R	CTCCTTGATCACTGCCTCGCCCTTGCTCACC ATTTTGCATATGGAATTGGGGGA	
HDAC4_NT_HR2F	AGTTCGGGGTCCAGCGGTTCTTCAGGCAGTTCTAGTC TCGACGATAGAATACCAATACAC	
HDAC4_NT_HR2R	GCCCTTGAACTCGATTGACGCTCTTCGACCCCACAA TTTCCCACCATCTATATCTATTC	
HDAC4_NT_pCFD3-F	GTCGTGTATTGGTATTCTATCGT	Oligos for cloning gRNAs into pCFD3
HDAC4_NT_pCFD3-R	AAACACGATAGAATACCAATACA	
HDAC4_CT_HR1F	GAAAGACTGGGCCTTTCGCCCGGGCTAACGTTGTTTCTTCAATTC GATAGCCATTGCGGC	PCR amplification of homology arms
HDAC4_CT_HR1R	CTCCTTGATCACTGCCTCGCCCTTGCTCACTTTGGTTTCATCCTGA TCCATCGGCTC	

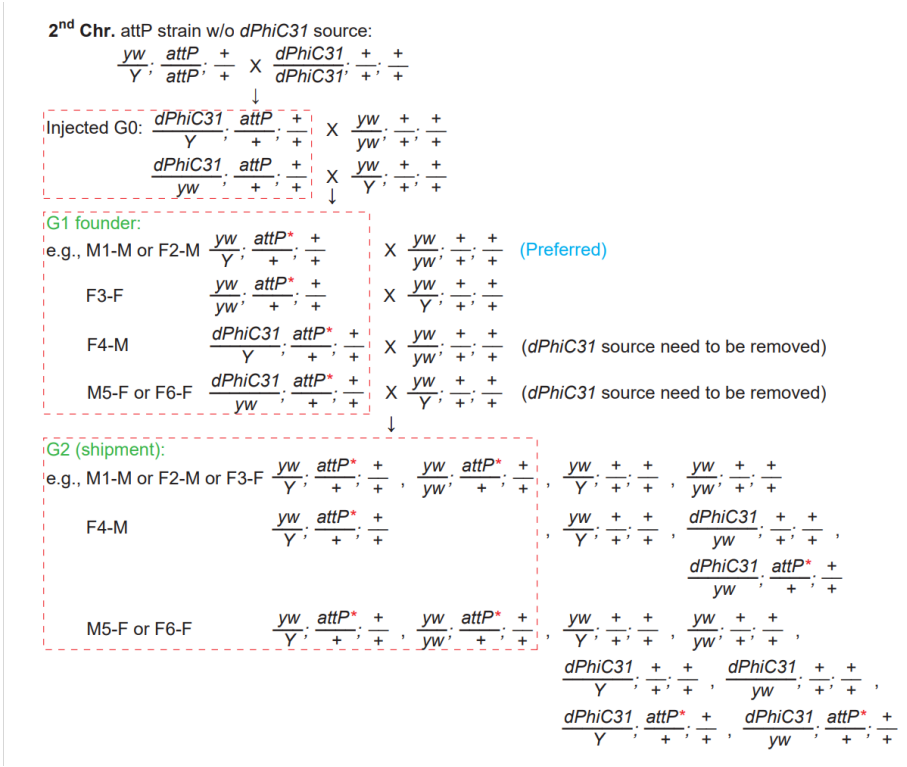
HDAC4_CT_HR2F	GTATAATGTATGCTATACGAAGTTATGGCAGAAGGCGATGCAGA CGTATATTGCG	for direct tagging of HDAC4
HDAC4_CT_HR1R	GCCCTTGAACTCGATTGACGCTCTTCGACCGTTACACACATGGTG AATACACATG	
HDAC4_CT_pCFD3-F	GTCGCGTCTGCATCGCCTTCTATT	Oligos for cloning gRNAs into pCFD3
HDAC4_CT_pCFD3-R	AAACAATAGAAGGCGATGCAGACG	
Mef2		
Mef2_CDT_HR1F	CCCGGGCTAATTATGGGGTGTCGCCCTTCGATAGCGGC ATGATTTTAGTGCTTAACGC	PCR amplification of homology arms for direct tagging of Mef2
Mef2_CDT_HR1R	CCCGGTGAACAGCTCCTCGCCCTTGCTCACTGTGCCC CATCCGCCCGATATTCTC	
Mef2_CDT_HR2F	GTATAATGTATGCTATACGAAGTTATGGCAGCCGGCACGAGCT GCTGATAGCAG	
Mef2_CDT_HR2R	GCCCTTGAACTCGATTGACGCTCTTCGACAAGTGTGTATGCCT TGCTTACACAAGGCTGT	
Mef2_pCFD3_CT_F	GTCGACATAGCCGGCACGAGCTGC	Oligos for cloning gRNAs into pCFD3
Mef2_pCFD3_CT_R	AAACGCAGCTCGTGCCGGCTATGT	
Fluorophore mutagenesis		
Scarlet-H H164M-F	GCTGAAGGGCGACATTAAGATGGCCCTGCGCCTGA	Mutagenesis of mScarlet-H to mScarlet-I within pN-mScarlet
Scarlet-H H164M-R	TCAGGCGCAGGGCCATCTTAATGTCGCCCTTCAGC	
Scarlet-I T74I-F	CCAGGGCCTTCATCAAGCACCCCGC	
Scarlet-I T74I-R	GCGGGGTGCTTGATGAAGGCCCTGG	
Scarlet_Fwd+SmaI	TTTCGCCCGGGCTAACGCGAGACCTATGGTCTCTGTGAGCAAG	PCR amplificiation of mScarlet-I for cloning into pC vector
Scarlet_Rev+BsiWI	CCCACGTACGATAACTTCGTATAGCATACATTATACGAAGTTAT TTACTTGTACAGC	
CRISPR allele genotyping		
Allele (figure)	**Primer pair**	

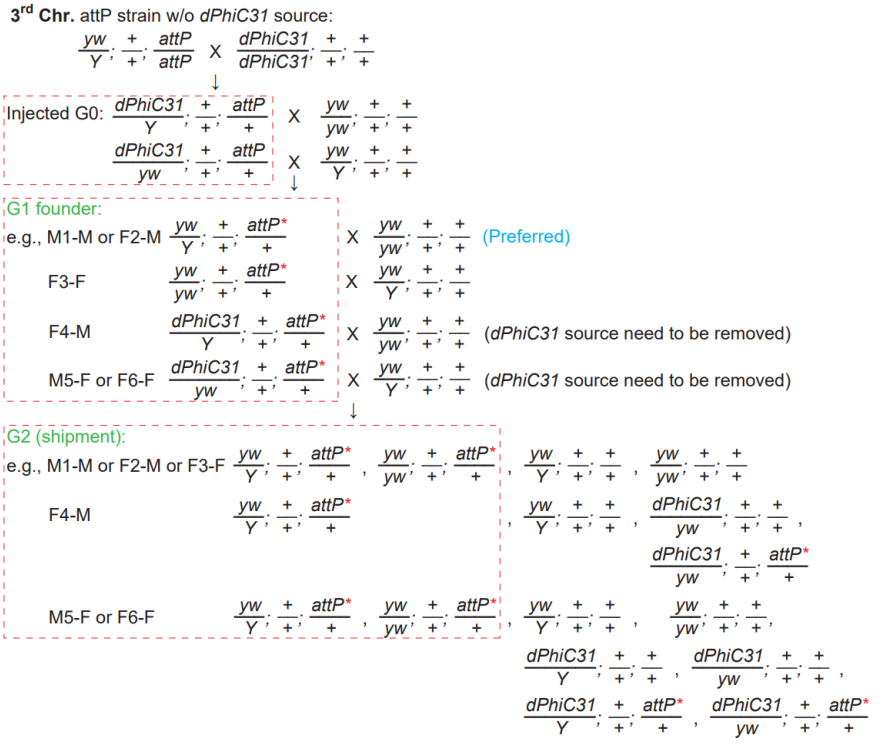
HDAC4^{Att}	HDAC4 5_Att_CRISPR-R & HDAC4_3_Att_CRISPR-F	
HDAC4$^{Att\text{-}gDNA}$	HDAC4 5HA_SEQ3 & Mini-White R, HDAC4rscuSeq4& HDAC4_3_Att_CRISPR-F, HDAC4rscuSeq3& HDAC4_3_Att_CRISPR-F.	
Mef2^{CDS}	Mef2_5_CDS/Wg_CRISPR-R & Mef2_3_CDS_CRISPR-F	
Mef2^{ATG}	Mef2_5_CDS/Wg_CRISPR-R & Mef2_3_Wg_CRISPR-F	
Mef2$^{ATG\text{-}cDNA}$	Mef2_5_CDS/Wg_CRISPR-R & Mef2III_CDS2_Rev	
mScarlet-I-HDAC4	HDAC4_NT_CRISPR-F & HDAC4_NT_CRISPR-F	
HDAC4-mScarlet-I	HDAC4rscuSeq5 & HDAC4_3_Att_CRISPR-F	

Appendix. 1. Primers used during this project

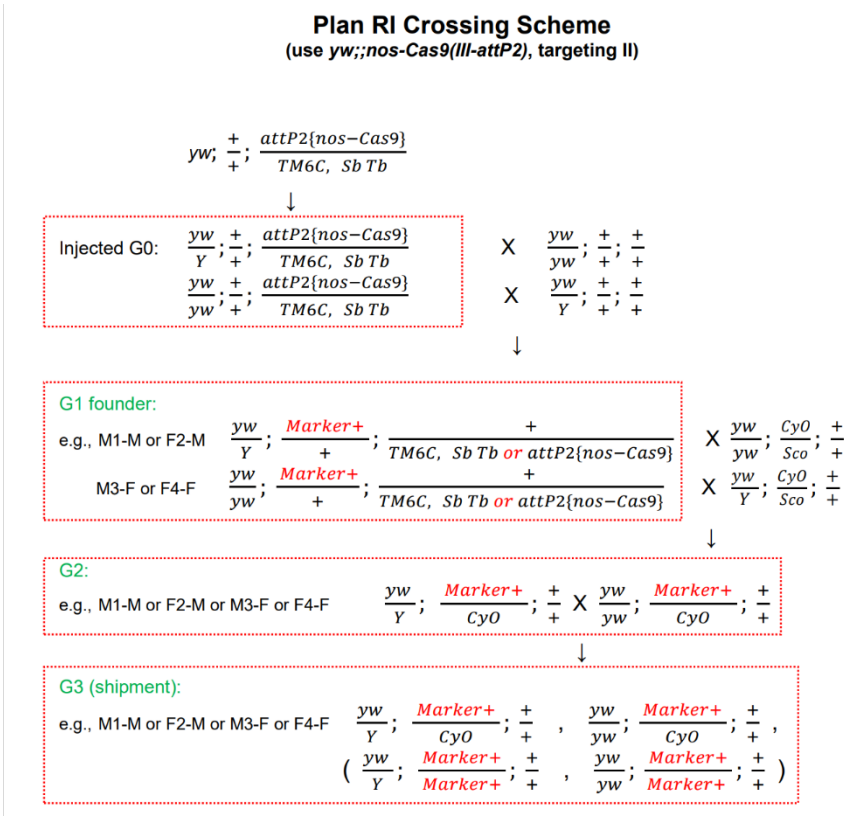
Antibodies			
Primary			
Name	**Host**	**Working dilution**	**Source**
α-Mhc (Myosin heavy chain)	Mouse	1:100	DSHB #3E8-3D3
α-Mef2	Rabbit	1:500	Gift from Eileen Furlong
α-GFP	Chicken	1:500	Abcam #13970
α-GFP	Rabbit	1:500	Proteintech #PABG1
α-RFP	Rat	1:500	Proteintech #5F8 (Chromotek)
α-Cut	Mouse	1:50	DSHB #2B10
α-myc	mouse	1:1000	Gift from Prof. Helen White-Cooper
Secondary/direct stains			
α-Rabbit Alexa Fluor™ 488	Goat	1:500	Invitrogen #A-11008, gift from Prof. Helen White-Cooper
α-Rabbit Alexa Fluor™ 555	Goat	1:500	Invitrogen #A-21428, gift from Prof. Helen White-Cooper
α-Mouse Alexa Fluor™ 555	Goat	1:500	Invitrogen #A-21422, gift from Prof. Helen White-Cooper
α-Rat Alexa Fluor™ 488	Goat	1:500	Invitrogen #A-11006, gift from Prof. Helen White-Cooper
α-Rat Alexa Fluor™ 555	Goat	1:500	Invitrogen #A-21434, gift from Prof. Helen White-Cooper
α-Chicken Alexa Fluor™ 488	Goat	1:500	Jackson laboratories # 703-545-155
HOECHST		1:3000	ThermoFisher #H3569
Phalloidin		1:140	Cytoskeleton Inc. # PHDH1-A

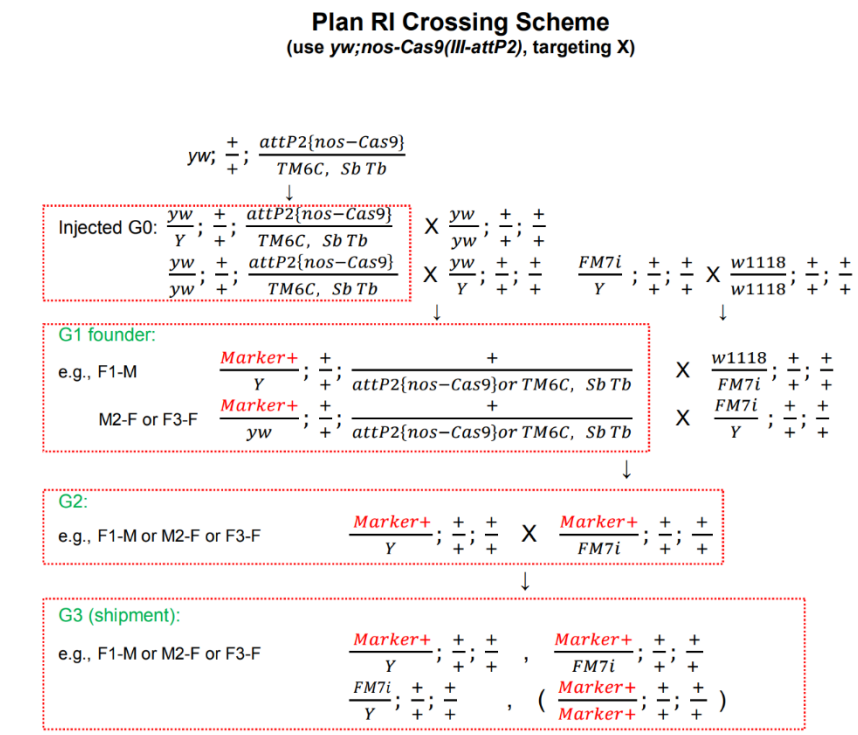
Appendix. 2. Antibody list.



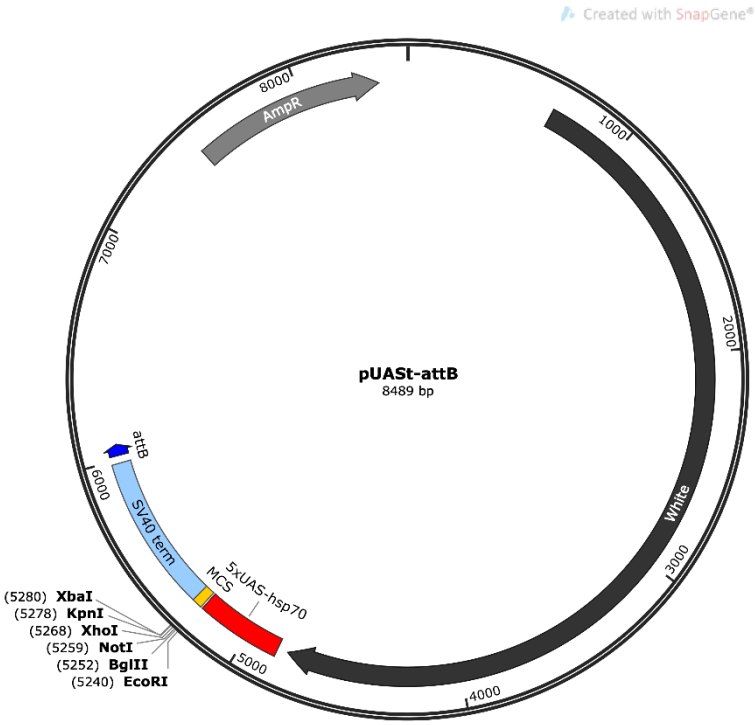


Appendix. 3. General crossing scheme used by BestGene to generate transgenic UAS overexpression lines (plan H)

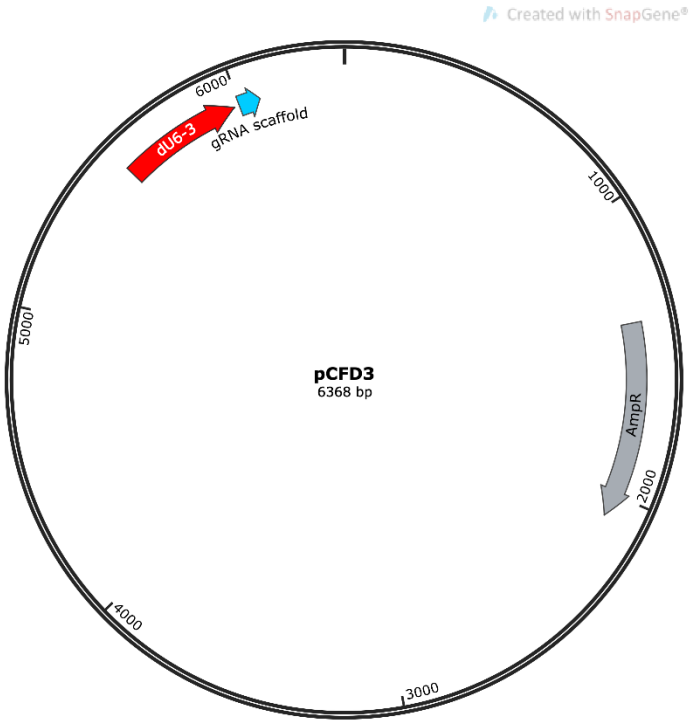




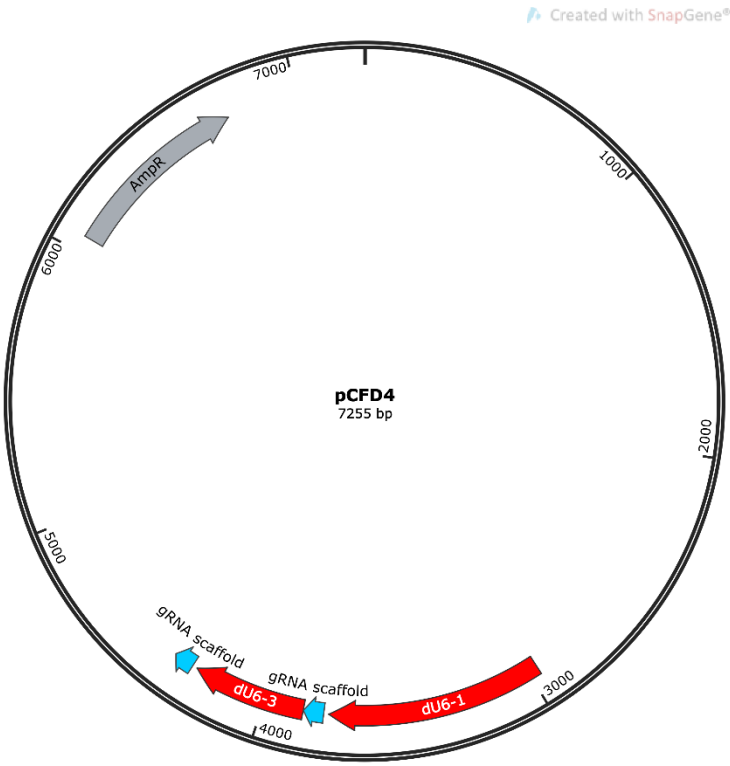
Appendix. 4. General crossing scheme used by BestGene to generate genome engineered alleles by CRISPR-Cas9, either targeting X (top), or II (bottom) (Plan RI).



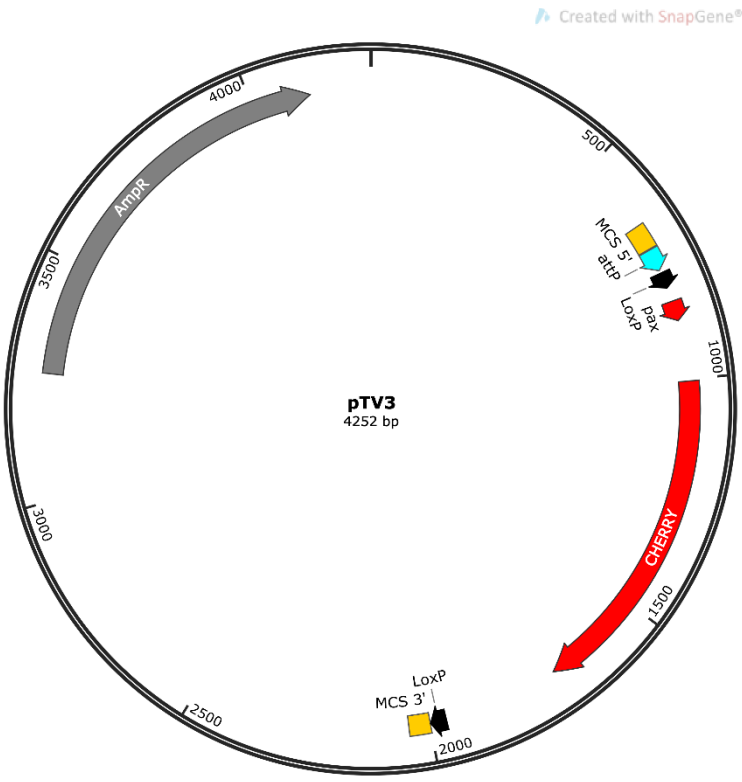
Appendix. 5. pUAST-attB vector.



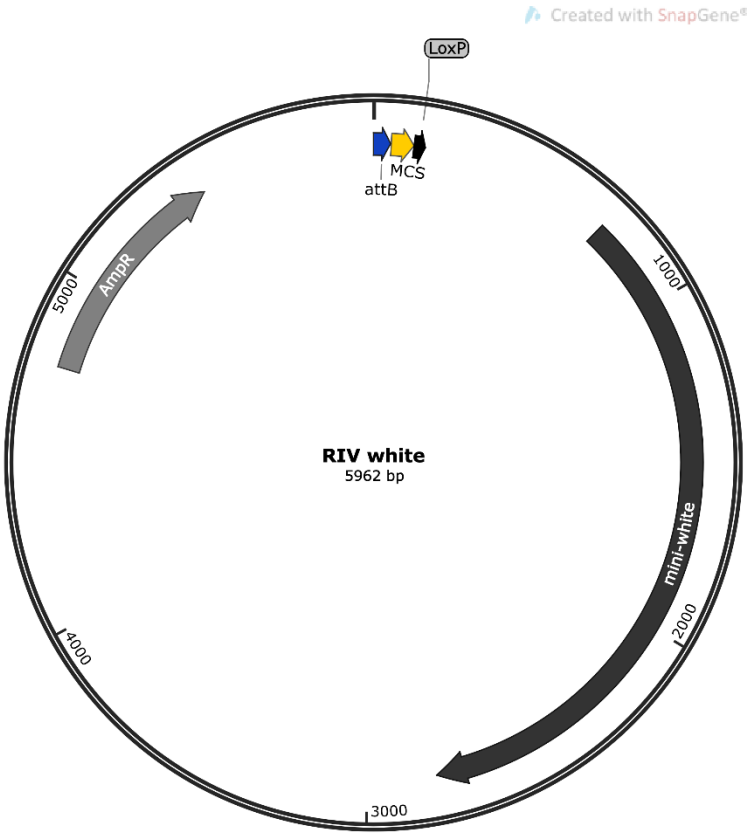
Appendix. 6. pCFD3 vector



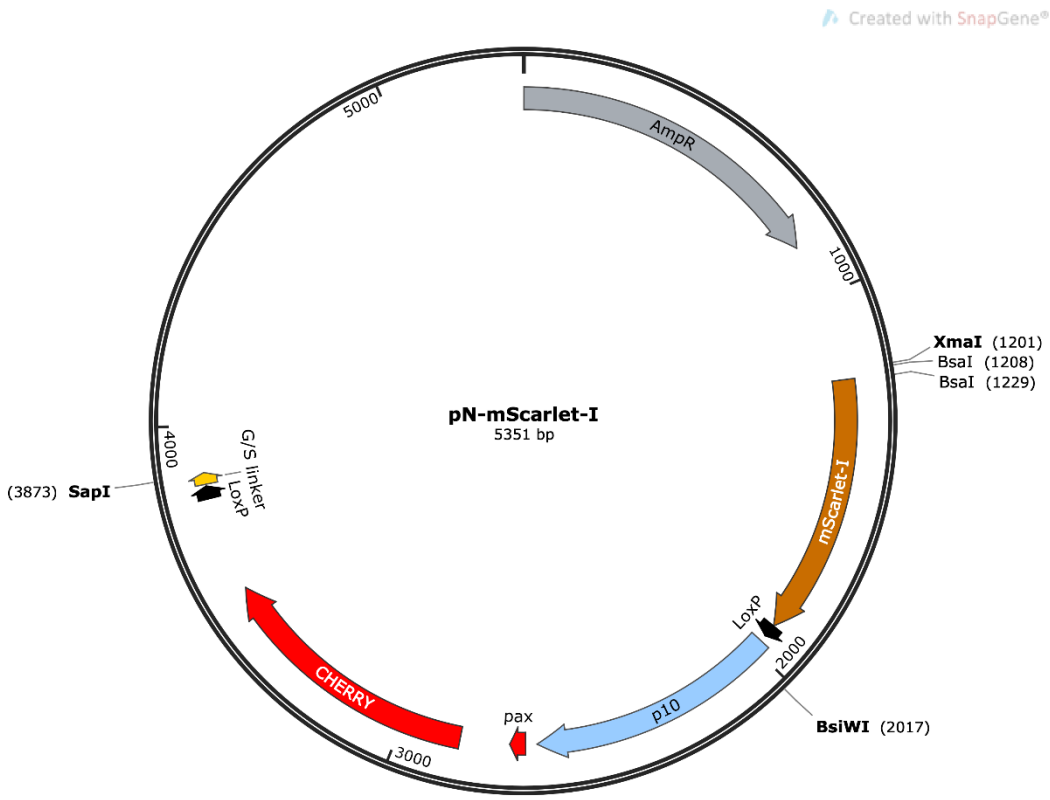
Appendix. 7. pCFD4 vector



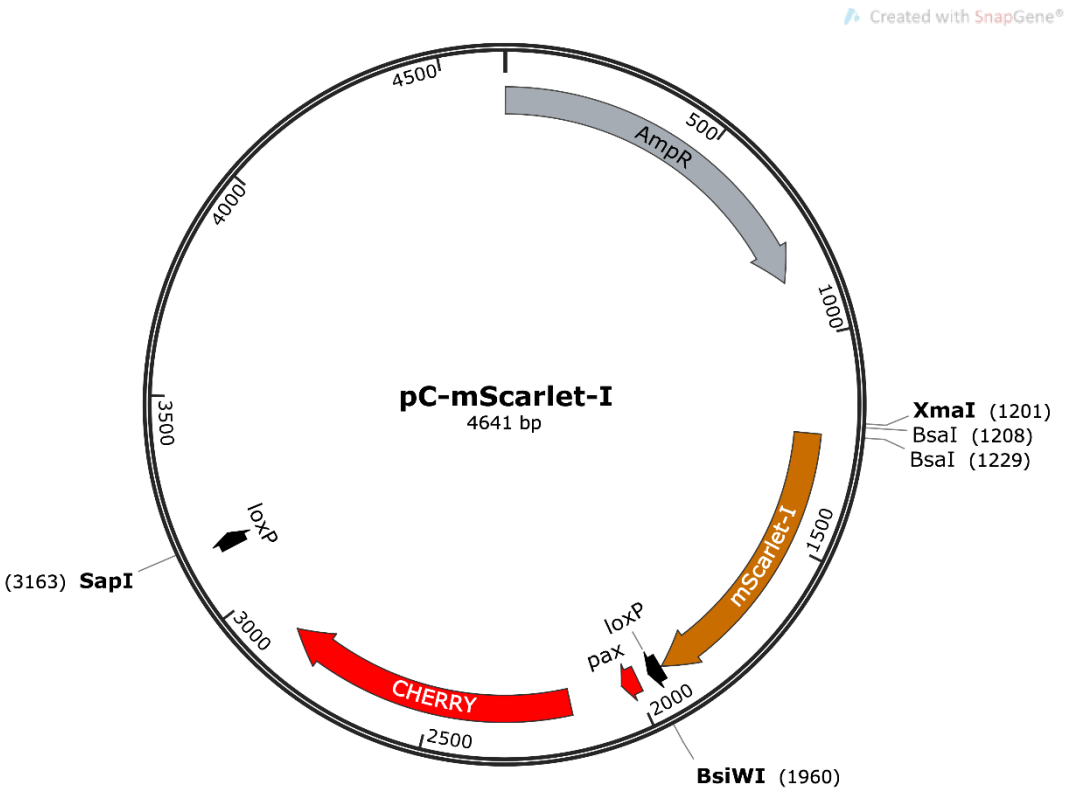
Appendix. 8. pTV3 vector



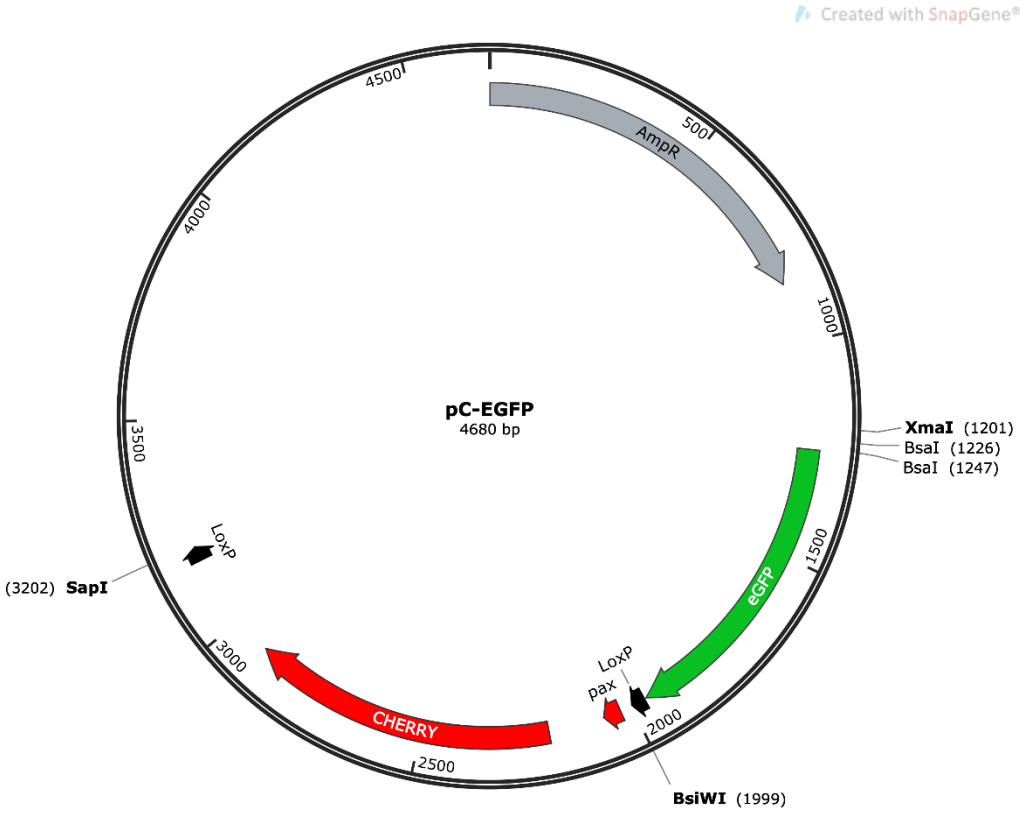
Appendix. 9. RIV^{White} vector



Appendix. 10. pN-mScarlet-I vector



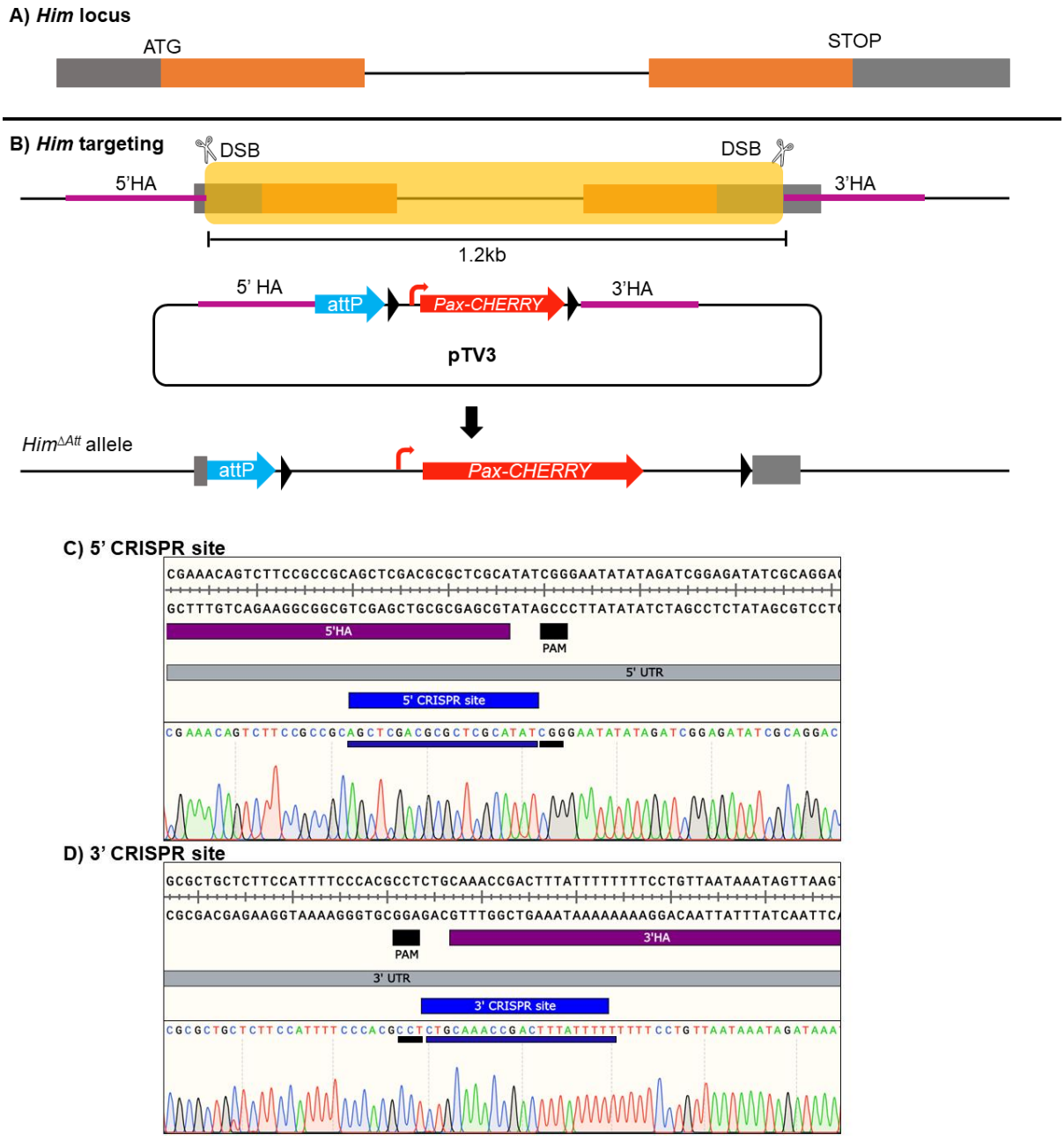
Appendix. 11. pC-mScarlet-I vector.



Appendix. 12. pC-EGFP vector.

CAATCCGCCCTCACTACAACCGGAGATCTTTCCGCCAGCGGTCATCAGCAGAATAGCAATGGTTCCACGG
GCAGCGGCGGATCCAGCAGCAGCACCAGTAGCAACGCCAGCGGAGGAGCAGGAGGCGGTGGAGCCGTC
AGCGCAGCCAATGTCATCACGCACTTGAACAACGTCAGTGTCCTGGCGGGAGGTCCTTCGGGGCAGGGA
GGAGGAGGCGGAGGCGGCGGCAGCAACGGGAATGTCGAACAGGCCACCAATCTTAGCGTACTGAGCCA
CGCGCAGCAACATCACCTGGGCATGCCCAACTCGCGTCCCTCGTCCACGGGCCACATCACACCCACTCCAG
GTGCGCCGAGCAGCGACCAGGATGTGCGTCTGGCAGCCGTCGCCGTGCAGCAGCAACAGCAGCAGCCAC
ATCAGCAACAGCAACTAGGCGACTACGATGCCCCCAACCACAAACGGCCGAGAATATCGGGCGGATGGG
GCACAGGCGGCTCCTACCCATACGATGTTCCTGACTATGCGGGCTATCCCTATGACGTCCCGGACTATGCA
GGATCCTATCCATATGACGTTCCAGATTACGCTTAGGGTACCTCTAGACTACTCTGGCGTCGATGAGGGA

Appendix. 13. Mef2 fragment sequence, generated by Twist Bioscience, for the generation of UAS-Mef2 transgenic lines.



Appendix. 14. targeting the Him locus to generate an insertion-ready deletion allele

References

Adams, M. D. et al. 2000. The genome sequence of Drosophila melanogaster. *Science* 287(5461), pp. 2185-2195. doi: 10.1126/science.287.5461.2185

Alexandre, C., Baena-Lopez, A. and Vincent, J. P. 2014. Patterning and growth control by membrane-tethered wingless. *Nature* 505(7482), pp. 180-185. doi: 10.1038/nature12879

Anant, S., Roy, S. and Vijayraghavan, K. 1998. Twist and Notch negatively regulate adult muscle differentiation in Drosophila. *Development* 125(8), pp. 1361-1369.

Anderson, C. M., Hu, J., Barnes, R. M., Heidt, A. B., Cornelissen, I. and Black, B. L. 2015. Myocyte enhancer factor 2C function in skeletal muscle is required for normal growth and glucose metabolism in mice. *Skeletal Muscle* 5(1), doi: 10.1186/s13395-015-0031-0

Angelelli, C. et al. 2008. Differentiation-dependent lysine 4 acetylation enhances MEF2C binding to DNA in skeletal muscle cells. *Nucleic Acids Research* 36(3), pp. 915-928. doi: 10.1093/nar/gkm1114

Arnold, M. A. et al. 2007. MEF2C Transcription Factor Controls Chondrocyte Hypertrophy and Bone Development. *Developmental Cell* 12(3), pp. 377-389. doi: 10.1016/j.devcel.2007.02.004

Baena-Lopez, L. A., Alexandre, C., Mitchell, A., Pasakarnis, L. and Vincent, J. P. 2013. Accelerated homologous recombination and subsequent genome modification in Drosophila. *Development (Cambridge)* 140(23), pp. 4818-4825. doi: 10.1242/dev.100933

Bainbridge, S. P. and Bownes, M. 1981. Staging the metamorphosis of Drosophila melanogaster. *Journal of Embryology and Experimental Morphology* Vol.66, pp. 57-80.

Bajar, B. T., Wang, E. S., Zhang, S., Lin, M. Z. and Chu, J. 2016. A guide to fluorescent protein FRET pairs. *Sensors (Switzerland)* 16(9), doi: 10.3390/s16091488

Balakrishnan, M., Sisso, W. J. and Baylies, M. K. 2021. Analyzing muscle structure and function throughout the larval instars in live Drosophila. *STAR Protocols* 2(1), doi: 10.1016/j.xpro.2020.100291

Bannister, A. J. and Kouzarides, T. 2011. Regulation of chromatin by histone modifications. *Cell Research* 21(3), pp. 381-395. doi: 10.1038/cr.2011.22

Bassett, A. R., Tibbit, C., Ponting, C. P. and Liu, J. L. 2013. Highly Efficient Targeted Mutagenesis of Drosophila with the CRISPR/Cas9 System. *Cell Reports* 4(1), pp. 220-228. doi: 10.1016/j.celrep.2013.06.020

Baylies, M. K. and Bate, M. 1996. twist: A myogenic switch in Drosophila. *Science* 272(5267), pp. 1481-1484. doi: 10.1126/science.272.5267.1481

Baylies, M. K., Bate, M. and Gomez, M. R. 1998. Myogenesis: A view from Drosophila. *Cell* 93(6), pp. 921-927. doi: 10.1016/S0092-8674(00)81198-8

Beckett, K. and Baylies, M. K. 2006. The Development of The Drosophila Larval Body Wall Muscles. *International Review of Neurobiology*.

Bedford, D. C., Kasper, L. H., Fukuyama, T. and Brindle, P. K. 2010. Target gene context influences the transcriptional requirement for the KAT3 family of CBP and p300 histone acetyltransferases. *Epigenetics* 5(1), pp. 9-15. doi: 10.4161/epi.5.1.10449

Bellen, H. J. et al. 2004. The BDGP gene disruption project: Single transposon insertions associated with 40% of Drosophila genes. *Genetics* 167(2), pp. 761-781. doi: 10.1534/genetics.104.026427

Bentzinger, C. F., Wang, Y. X. and Rudnicki, M. A. 2012. Building muscle: molecular regulation of myogenesis. *Cold Spring Harbor perspectives in biology* 4(2), doi: 10.1101/cshperspect.a008342

Bertin, B. et al. 2021. Gelsolin and dCryAB act downstream of muscle identity genes and contribute to preventing muscle splitting and branching in Drosophila. *Scientific Reports* 11(1), doi: 10.1038/s41598-021-92506-3

Bier, E., Harrison, M. M., O'connor-Giles, K. M. and Wildonger, J. 2018. Advances in engineering the fly genome with the CRISPR-Cas system. *Genetics* 208(1), pp. 1-18. doi: 10.1534/genetics.117.1113

Bindels, D. S. et al. 2016. MScarlet: A bright monomeric red fluorescent protein for cellular imaging. *Nature Methods* 14(1), pp. 53-56. doi: 10.1038/nmeth.4074

Bischof, J., Björklund, M., Furger, E., Schertel, C., Taipale, J. and Basler, K. 2012. A versatile platform for creating a comprehensive UAS-ORFeome library in Drosophila. *Development (Cambridge)* 140(11), pp. 2434-2442. doi: 10.1242/dev.088757

Bischof, J., Maeda, R. K., Hediger, M., Karch, F. and Basler, K. 2007. An optimized transgenesis system for Drosophila using germ-line-specific φC31 integrases. *Proceedings of the National Academy of Sciences of the United States of America* 104(9), pp. 3312-3317. doi: 10.1073/pnas.0611511104

Black, B. L. and Olson, E. N. 1998. Transcriptional control of muscle development by myocyte enhancer factor-2 (MEF2) proteins. *Annual Review of Cell and Developmental Biology* 14, pp. 167-196. doi: 10.1146/annurev.cellbio.14.1.167

Blake, J. A. and Ziman, M. R. 2014. Pax genes: Regulators of lineage specification and progenitor cell maintenance. *Development (Cambridge)* 141(4), pp. 737-751. doi: 10.1242/dev.091785

Bothe, I. and Baylies, M. K. 2016. Drosophila myogenesis. *Current Biology* 26(17), pp. R786-R791. doi: 10.1016/j.cub.2016.07.062

Bottomley, M. J. et al. 2008. Structural and functional analysis of the human HDAC4 catalytic domain reveals a regulatory structural zinc-binding domain. *Journal of Biological Chemistry* 283(39), pp. 26694-26704. doi: 10.1074/jbc.M803514200

Boukhatmi, H. and Bray, S. 2018. A population of adult satellite-like cells in Drosophila is maintained through a switch in RNA-isoforms. *eLife* 7, doi: 10.7554/eLife.35954

Bour, B. A. et al. 1995. Drosophila MEF2, a transcription factor that is essential for myogenesis. *Genes and Development* 9(6), pp. 730-741. doi: 10.1101/gad.9.6.730

Brand, A. H. and Perrimon, N. 1993. Targeted gene expression as a means of altering cell fates and generating dominant phenotypes. *Development* 118(2), pp. 401-415.

Bray, S. J. 2016. Notch signalling in context. *Nature Reviews Molecular Cell Biology* 17(11), pp. 722-735. doi: 10.1038/nrm.2016.94

Bryantsev, A. L., Baker, P. W., Lovato, T. L., Jaramillo, M. S. and Cripps, R. M. 2012. Differential requirements for Myocyte Enhancer Factor-2 during adult myogenesis in Drosophila. *Developmental Biology* 361(2), pp. 191-207. doi: 10.1016/j.ydbio.2011.09.031

Bu, L. and Cripps, R. M. 2022. Promoter architecture of Drosophila genes regulated by Myocyte enhancer factor-2. *PLoS ONE* 17(7 July), doi: 10.1371/journal.pone.0271554

Carmena, A., Bate, M. and Jimenez, F. 1995. lethal of scute, a proneural gene, participates m the specification of muscle progenitors during Drosophila embryogenesis. *Genes and Development* 9(19), pp. 2373-2383. doi: 10.1101/gad.9.19.2373

Carmena, A., Murugasu-Oei, B., Menon, D., Jiménez, F. and Chia, W. 1998. inscuteable and numb mediate asymmetric muscle progenitor cell divisions during Drosophila myogenesis. *Genes and Development* 12(3), pp. 304-315. doi: 10.1101/gad.12.3.304

Chan, J. K. L., Sun, L., Yang, X. J., Zhu, G. and Wu, Z. 2003. Functional characterization of an amino-terminal region of HDAC4 that possesses MEF2 binding and transcriptional repressive activity. *Journal of Biological Chemistry* 278(26), pp. 23515-23521. doi: 10.1074/jbc.M301922200

Chang, S., McKinsey, T. A., Zhang, C. L., Richardson, J. A., Hill, J. A. and Olson, E. N. 2004. Histone deacetylases 5 and 9 govern responsiveness of the heart to a subset of stress signals and play redundant roles in heart development. *Molecular and Cellular Biology* 24(19), pp. 8467-8476. doi: 10.1128/MCB.24.19.8467-8476.2004

Chang, S., Young, B. D., Li, S., Qi, X., Richardson, J. A. and Olson, E. N. 2006. Histone Deacetylase 7 Maintains Vascular Integrity by Repressing Matrix Metalloproteinase 10. *Cell* 126(2), pp. 321-334. doi: 10.1016/j.cell.2006.05.040

Chaturvedi, D., Reichert, H., Gunage, R. D. and VijayRaghavan, K. 2017. Identification and functional characterization of muscle satellite cells in Drosophila. *eLife* 6, doi: 10.7554/eLife.30107

Chechenova, M. B., Maes, S. and Cripps, R. M. 2015. Expression of the Troponin C at 41C Gene in Adult Drosophila Tubular Muscles Depends upon Both Positive and Negative Regulatory Inputs. *PLoS ONE* 10(12), doi: 10.1371/journal.pone.0144615

Chen, H. and Pugh, B. F. 2021. What do Transcription Factors Interact With? *Journal of Molecular Biology* 433(14), doi: 10.1016/j.jmb.2021.166883

Chen, Z., Liang, S., Zhao, Y. and Han, Z. 2012. miR-92b regulates Mef2 levels through a negative-feedback circuit during Drosophila muscle development. *Development (Cambridge)* 139(19), pp. 3543-3552. doi: 10.1242/dev.082719

Choi, S., Lim, D. S. and Chung, J. 2015. Feeding and Fasting Signals Converge on the LKB1-SIK3 Pathway to Regulate Lipid Metabolism in Drosophila. *PLoS Genetics* 11(5), doi: 10.1371/journal.pgen.1005263

Chou, T. B. and Perrimon, N. 1992. Use of a yeast site-specific recombinase to produce female germline chimeras in drosophila. *Genetics* 131(3), pp. 643-653.

Chyb, S. and Gompel, N. 2013. *Atlas of Drosophila Morphology: Wild-type and Classical Mutants*.

Clark, R. I. et al. 2013. XMEF2 is an in vivo immune-metabolic switch. *Cell* 155(2), p. 435. doi: 10.1016/j.cell.2013.09.007

Clocchiatti, A., Florean, C. and Brancolini, C. 2011. Class IIa HDACs: From important roles in differentiation to possible implications in tumourigenesis. *Journal of Cellular and Molecular Medicine* 15(9), pp. 1833-1846. doi: 10.1111/j.1582-4934.2011.01321.x

Cong, L. et al. 2013. Multiplex genome engineering using CRISPR/Cas systems. *Science* 339(6121), pp. 819-823. doi: 10.1126/science.1231143

Cripps, R. M., Black, B. L., Zhao, B., Lien, C. L., Schulz, R. A. and Olson, E. N. 1998. The myogenic regulatory gene Mef2 is a direct target for transcriptional activation by Twist during Drosophila myogenesis. *Genes and Development* 12(3), pp. 422-434. doi: 10.1101/gad.12.3.422

Cripps, R. M., Lovato, T. L. and Olson, E. N. 2004. Positive autoregulation of the Myocyte enhancer factor-2 myogenic control gene during somatic muscle development in Drosophila. *Developmental Biology* 267(2), pp. 536-547. doi: 10.1016/j.ydbio.2003.12.004

Dancy, B. M. and Cole, P. A. 2015. Protein lysine acetylation by p300/CBP. *Chemical Reviews* 115(6), pp. 2419-2452. doi: 10.1021/cr500452k

Davidson, E. H. 2010. Emerging properties of animal gene regulatory networks. *Nature* 468(7326), pp. 911-920. doi: 10.1038/nature09645

Davidson, E. H. and Erwin, D. H. 2006. Gene regulatory networks and the evolution of animal body plans. *Science* 311(5762), pp. 796-797. doi: 10.1126/science.1113832

Del Valle Rodríguez, A., Didiano, D. and Desplan, C. 2012. Power tools for gene expression and clonal analysis in Drosophila. *Nature Methods* 9(1), pp. 47-55. doi: 10.1038/nmeth.1800

Dietzl, G. et al. 2007. A genome-wide transgenic RNAi library for conditional gene inactivation in Drosophila. *Nature* 448(7150), pp. 151-156. doi: 10.1038/nature05954

Dobi, K. C., Schulman, V. K. and Baylies, M. K. 2015. Specification of the somatic musculature in Drosophila. *Wiley Interdisciplinary Reviews: Developmental Biology* 4(4), pp. 357-375. doi: 10.1002/wdev.182

Dodou, E., Xu, S. M. and Black, B. L. 2003. mef2c is activated directly by myogenic basic helix-loop-helix proteins during skeletal muscle development in vivo. *Mechanisms of Development* 120(9), pp. 1021-1032. doi: 10.1016/S0925-4773(03)00178-3

Dong, C. et al. 2017. Myocyte enhancer factor 2C and its directly-interacting proteins: A review. *Progress in Biophysics and Molecular Biology* 126, pp. 22-30. doi: 10.1016/j.pbiomolbio.2017.02.002

Doudna, J. A. and Charpentier, E. 2014. The new frontier of genome engineering with CRISPR-Cas9. *Science* 346(6213), doi: 10.1126/science.1258096

Edmondson, D. G., Lyons, G. E., Martin, J. F. and Olson, E. N. 1994. Mef2 gene expression marks the cardiac and skeletal muscle lineages during mouse embryogenesis. *Development* 120(5), pp. 1251-1263.

Elgar, S. J., Han, J. and Taylor, M. V. 2008. mef2 activity levels differentially affect gene expression during Drosophila muscle development. *Proceedings of the National Academy of Sciences of the United States of America* 105(3), pp. 918-923. doi: 10.1073/pnas.0711255105

Estrella, N. L., Desjardins, C. A., Nocco, S. E., Clark, A. L., Maksimenko, Y. and Naya, F. J. 2015. MEF2 transcription factors regulate distinct gene programs in mammalian skeletal muscle differentiation. *Journal of Biological Chemistry* 290(2), pp. 1256-1268. doi: 10.1074/jbc.M114.589838

Evangelou, A. et al. 2019. Unpredictable effects of the genetic background of transgenic lines in physiological quantitative traits. *G3: Genes, Genomes, Genetics* 9(11), pp. 3877-3890. doi: 10.1534/g3.119.400715

Fauquier, L. et al. 2018. CBP and P300 regulate distinct gene networks required for human primary myoblast differentiation and muscle integrity. *Scientific Reports* 8(1), doi: 10.1038/s41598-018-31102-4

Fernandes, J., Bate, M. and Vijayraghavan, K. 1991. Development of the indirect flight muscles of Drosophila. *Development* 113(1), pp. 67-77.

Fischle, W., Dequiedt, F., Hendzel, M. J., Guenther, M. G., Lazar, M. A., Voelter, W. and Verdin, E. 2002. Enzymatic activity associated with class II HDACs is dependent on a multiprotein complex containing HDAC3 and SMRT/N-CoR. *Molecular Cell* 9(1), pp. 45-57. doi: 10.1016/S1097-2765(01)00429-4

Fitzsimons, H. L., Schwartz, S., Given, F. M. and Scott, M. J. 2013. The histone deacetylase HDAC4 regulates long-term memory in Drosophila. *PLoS ONE* 8(12), doi: 10.1371/journal.pone.0083903

Gallo, S. M., Gerrard, D. T., Miner, D., Simich, M., Des Soye, B., Bergman, C. M. and Halfon, M. S. 2011. REDfly v3.0: Toward a comprehensive database of transcriptional regulatory elements in Drosophila. *Nucleic Acids Research* 39(SUPPL. 1), pp. D118-D123. doi: 10.1093/nar/gkq999

Girdwood, D. et al. 2003. p300 transcriptional repression is mediated by SUMO modification. *Molecular Cell* 11(4), pp. 1043-1054. doi: 10.1016/S1097-2765(03)00141-2

Gramates, L. S. et al. 2022. FlyBase: a guided tour of highlighted features. *Genetics* 220(4), doi: 10.1093/genetics/iyac035

Gratz, S. J. et al. 2013. Genome engineering of Drosophila with the CRISPR RNA-guided Cas9 nuclease. *Genetics* 194(4), pp. 1029-1035. doi: 10.1534/genetics.113.152710

Gratz, S. J., Ukken, F. P., Rubinstein, C. D., Thiede, G., Donohue, L. K., Cummings, A. M. and Oconnor-Giles, K. M. 2014. Highly specific and efficient CRISPR/Cas9-catalyzed homology-directed repair in Drosophila. *Genetics* 196(4), pp. 961-971. doi: 10.1534/genetics.113.160713

Green, E. W., Fedele, G., Giorgini, F. and Kyriacou, C. P. 2014. A Drosophila RNAi collection is subject to dominant phenotypic effects. *Nature Methods* 11(3), pp. 222-223. doi: 10.1038/nmeth.2856

Grégoire, S. et al. 2006. Control of MEF2 transcriptional activity by coordinated phosphorylation and sumoylation. *Journal of Biological Chemistry* 281(7), pp. 4423-4433. doi: 10.1074/jbc.M509471200

Grégoire, S. and Yang, X. J. 2005. Association with class IIa histone deacetylases upregulates the sumoylation of MEF2 transcription factors. *Molecular and Cellular Biology* 25(6), pp. 2273-2287. doi: 10.1128/MCB.25.6.2273-2287.2005

Groth, A. C., Fish, M., Nusse, R. and Calos, M. P. 2004. Construction of Transgenic Drosophila by Using the Site-Specific Integrase from Phage φC31. *Genetics* 166(4), pp. 1775-1782. doi: 10.1534/genetics.166.4.1775

Gunage, R. D., Dhanyasi, N., Reichert, H. and VijayRaghavan, K. 2017. Drosophila adult muscle development and regeneration. *Seminars in Cell and Developmental Biology* 72, pp. 56-66. doi: 10.1016/j.semcdb.2017.11.017

Gunage, R. D., Reichert, H. and VijayRaghavan, K. 2014. Identification of a new stem cell population that generates Drosophila flight muscles. *eLife* 3(August2014), pp. 1-25. doi: 10.7554/eLife.03126

Gunthorpe, D., Beatty, K. E. and Taylor, M. V. 1999. Different levels, but not different isoforms, of the Drosophila transcription factor DMEF2 affect distinct aspects of muscle differentiation. *Developmental Biology* 215(1), pp. 130-145. doi: 10.1006/dbio.1999.9449

Haberland, M., Arnold, M. A., McAnally, J., Phan, D., Kim, Y. and Olson, E. N. 2007. Regulation of HDAC9 gene expression by MEF2 establishes a negative-feedback loop in the transcriptional circuitry of muscle differentiation. *Molecular and Cellular Biology* 27(2), pp. 518-525. doi: 10.1128/MCB.01415-06

Haberland, M., Montgomery, R. L. and Olson, E. N. 2009. The many roles of histone deacetylases in development and physiology: Implications for disease and therapy. *Nature Reviews Genetics* 10(1), pp. 32-42. doi: 10.1038/nrg2485

Hales, K. G., Korey, C. A., Larracuente, A. M. and Roberts, D. M. 2015. Genetics on the fly: A primer on the drosophila model system. *Genetics* 201(3), pp. 815-842. doi: 10.1534/genetics.115.183392

Han, A., He, J., Wu, Y., Liu, J. O. and Chen, L. 2005. Mechanism of recruitment of class II histone deacetylases by myocyte enhancer factor-2. *Journal of Molecular Biology* 345(1), pp. 91-102. doi: 10.1016/j.jmb.2004.10.033

Hanahan, D. and Weinberg, R. A. 2011. Hallmarks of cancer: The next generation. *Cell* 144(5), pp. 646-674. doi: 10.1016/j.cell.2011.02.013

He, J. et al. 2011. Structure of p300 bound to MEF2 on DNA reveals a mechanism of enhanceosome assembly. *Nucleic Acids Research* 39(10), pp. 4464-4474. doi: 10.1093/nar/gkr030

Heigwer, F., Port, F. and Boutros, M. 2018. Rna interference (RNAi) screening in Drosophila. *Genetics* 208(3), pp. 853-874. doi: 10.1534/genetics.117.300077

Holmqvist, P. H. and Mannervik, M. 2013. Genomic occupancy of the transcriptional co-activators p300 and CBP. *Transcription* 4(1), pp. 18-23. doi: 10.4161/trns.22601

Horn, C., Schmid, B. G. M., Pogoda, F. S. and Wimmer, E. A. 2002. Fluorescent transformation markers for insect transgenesis. *Insect Biochemistry and Molecular Biology* 32(10), pp. 1221-1235. doi: 10.1016/S0965-1748(02)00085-1

Hunt, G., Boija, A. and Mannervik, M. 2022. p300/CBP sustains Polycomb silencing by non-enzymatic functions. *Molecular Cell* 82(19), pp. 3580-3597.e3589. doi: 10.1016/j.molcel.2022.09.005

Jawkar, S. and Nongthomba, U. 2020. Indirect flight muscles in Drosophila melanogaster as a tractable model to study muscle development and disease. *International Journal of Developmental Biology* 64(1-3), pp. 177-183. doi: 10.1387/ijdb.190333un

Jayathilaka, N. et al. 2012. Inhibition of the function of class IIa HDACs by blocking their interaction with MEF2. *Nucleic Acids Research* 40(12), pp. 5378-5388. doi: 10.1093/nar/gks189

Jia, D. et al. 2015. A large-scale in vivo RNAi screen to identify genes involved in Notch-mediated follicle cell differentiation and cell cycle switches. *Scientific Reports* 5, doi: 10.1038/srep12328

Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J. A. and Charpentier, E. 2012. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* 337(6096), pp. 816-821. doi: 10.1126/science.1225829

Jones, P. et al. 2008. Probing the elusive catalytic activity of vertebrate class IIa histone deacetylases. *Bioorganic and Medicinal Chemistry Letters* 18(6), pp. 1814-1819. doi: 10.1016/j.bmcl.2008.02.025

Junion, G. and Jagla, K. 2022. Diversification of muscle types in Drosophila embryos. *Experimental Cell Research* 410(1), doi: 10.1016/j.yexcr.2021.112950

Junion, G., Jagla, T., Duplant, S., Tapin, R., Da Ponte, J. P. and Jagla, K. 2005. Mapping Dmef2-binding regulatory modules by using a ChIP-enriched in silico targets approach. *Proceedings of the National Academy of Sciences of the United States of America* 102(51), pp. 18479-18484. doi: 10.1073/pnas.0507030102

Kanca, O. et al. 2019. An efficient CRISPR-based strategy to insert small and large fragments of DNA using short homology arms. *eLife* 8, doi: 10.7554/eLife.51539

Kang, J., Gocke, C. B. and Yu, H. 2006. Phosphorylation-facilitated sumoylation of MEF2C negatively regulates its transcriptional activity. *BMC Biochemistry* 7, p. 14P. doi: 10.1186/1471-2091-7-5

Kato, Y. et al. 2000. Big mitogen-activated kinase regulates multiple members of the MEF2 protein family. *Journal of Biological Chemistry* 275(24), pp. 18534-18540. doi: 10.1074/jbc.M001573200

Khalil, A. M. 2020. The genome editing revolution: review. *Journal of Genetic Engineering and Biotechnology* 18(1), doi: 10.1186/s43141-020-00078-y

Kondo, S. and Ueda, R. 2013. Highly Improved gene targeting by germline-specific Cas9 expression in Drosophila. *Genetics* 195(3), pp. 715-721. doi: 10.1534/genetics.113.156737

Koonin, E. V. and Makarova, K. S. 2019. Origins and evolution of CRISPR-Cas systems. *Philosophical Transactions of the Royal Society B: Biological Sciences* 374(1772), doi: 10.1098/rstb.2018.0087

Koreman, G. T. et al. 2021. Upgraded CRISPR/Cas9 tools for tissue-specific mutagenesis in Drosophila. *Proceedings of the National Academy of Sciences of the United States of America* 118(14), doi: 10.1073/pnas.2014255118

Koyama, T. and Mirth, C. K. 2021. Ecdysone quantification from whole body samples of drosophila melanogaster larvae. *Bio-protocol* 11(3), doi: 10.21769/BioProtoc.3915

Lahm, A. et al. 2007. Unraveling the hidden catalytic activity of vertebrate class IIa histone deacetylases. *Proceedings of the National Academy of Sciences of the United States of America* 104(44), pp. 17335-17340. doi: 10.1073/pnas.0706487104

Lam, A. J. et al. 2012. Improving FRET dynamic range with bright green and red fluorescent proteins. *Nature Methods* 9(10), pp. 1005-1012. doi: 10.1038/nmeth.2171

Lander, E. S. et al. 2001. Initial sequencing and analysis of the human genome. *Nature* 409(6822), pp. 860-921. doi: 10.1038/35057062

Laurichesse, Q. and Soler, C. 2020. Muscle development : a view from adult myogenesis in Drosophila. *Seminars in Cell and Developmental Biology* 104, pp. 39-50. doi: 10.1016/j.semcdb.2020.02.009

Lee, H. H. and Frasch, M. 2000. Wingless effects mesoderm patterning and ectoderm segmentation events via induction of its downstream target sloppy paired. *Development* 127(24), pp. 5497-5508.

Lemercier, C., Verdel, A., Galloo, B., Curtet, S., Brocard, M. P. and Khochbin, S. 2000. mHDA1/HDAC5 histone deacetylase interacts with and represses MEF2A transcriptional activity. *Journal of Biological Chemistry* 275(20), pp. 15594-15599. doi: 10.1074/jbc.M908437199

Lemke, S. B. and Schnorrer, F. 2017. Mechanical forces during muscle development. *Mechanisms of Development* 144, pp. 92-101. doi: 10.1016/j.mod.2016.11.003

Lemke, S. B., Weidemann, T., Cost, A. L., Grashoff, C. and Schnorrer, F. 2019. A small proportion of talin molecules transmit forces at developing muscle attachments in vivo. *PLoS Biology* 17(3), doi: 10.1371/journal.pbio.3000057

Leptin, M. 1991. twist and snail as positive and negative regulators during Drosophila mesoderm development. *Genes and Development* 5(9), pp. 1568-1576.

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

Levine, M. and Davidson, E. H. 2005. Gene regulatory networks for development. *Proceedings of the National Academy of Sciences of the United States of America* 102(14), pp. 4936-4942. doi: 10.1073/pnas.0408031102

Li, H., Lones, L. and Diantonio, A. 2021. Bidirectional regulation of glial potassium buffering – glioprotection versus neuroprotection. *eLife* 10, doi: 10.7554/eLife.62606

Lilly, B., Zhao, B., Ranganayakulu, G., Paterson, B. M., Schulz, R. A. and Olson, E. N. 1995. Requirement of MADS domain transcription factor D-MEF2 for muscle formation in Drosophila. *Science* 267(5198), pp. 688-693. doi: 10.1126/science.7839146

Lin, M. H., Bour, B. A., Abmayr, S. M. and Storti, R. V. 1997a. Ectopic expression of MEF2 in the epidermis induces epidermal expression of muscle genes and abnormal muscle development in Drosophila. *Developmental Biology* 182(2), pp. 240-255. doi: 10.1006/dbio.1996.8484

Lin, Q., Schwarz, J., Bucana, C. and Olson, E. N. 1997b. Control of mouse cardiac morphogenesis and myogenesis by transcription factor MEF2C. *Science* 276(5317), pp. 1404-1407. doi: 10.1126/science.276.5317.1404

Lin, W. H. and Baines, R. A. 2019. Myocyte enhancer factor-2 and p300 interact to regulate the expression of homeostatic regulator Pumilio in Drosophila. *European Journal of Neuroscience* 50(1), pp. 1727-1740. doi: 10.1111/ejn.14357

Liotta, D., Han, J., Elgar, S., Garvey, C., Han, Z. and Taylor, M. V. 2007. The Him Gene Reveals a Balance of Inputs Controlling Muscle Differentiation in Drosophila. *Current Biology* 17(16), pp. 1409-1413. doi: 10.1016/j.cub.2007.07.039

Logan, C. Y. and Nusse, R. 2004. The Wnt signaling pathway in development and disease. *Annual Review of Cell and Developmental Biology* 20, pp. 781-810. doi: 10.1146/annurev.cellbio.20.010403.113126

Lones, L. and DiAntonio, A. 2023. SIK3 and Wnk converge on Fray to regulate glial K+ buffering and seizure susceptibility. *PLoS Genetics* 19(1), doi: 10.1371/journal.pgen.1010581

Lovato, T. L., Benjamin, A. R. and Cripps, R. M. 2005. Transcription of Myocyte enhancer factor-2 in adult Drosophila myoblasts is induced by the steroid hormone ecdysone. *Developmental Biology* 288(2), pp. 612-621. doi: 10.1016/j.ydbio.2005.09.007

Lowe, N. et al. 2014. Analysis of the expression patterns, Subcellular localisations and interaction partners of drosophila proteins using a pigp protein trap library. *Development (Cambridge)* 141(20), pp. 3994-4005. doi: 10.1242/dev.111054

Lu, J., McKinsey, T. A., Nicol, R. L. and Olson, E. N. 2000a. Signal-dependent activation of the MEF2 transcription factor by dissociation from histone deacetylases. *Proceedings of the National Academy of Sciences of the United States of America* 97(8), pp. 4070-4075. doi: 10.1073/pnas.080064097

Lu, J., McKinsey, T. A., Zhang, C. L. and Olson, E. N. 2000b. Regulation of skeletal myogenesis by association of the MEF2 transcription factor with class II histone deacetylases. *Molecular Cell* 6(2), pp. 233-244. doi: 10.1016/S1097-2765(00)00025-3

Luck, K. et al. 2020. A reference map of the human binary protein interactome. *Nature* 580(7803), pp. 402-408. doi: 10.1038/s41586-020-2188-x

Ma, K., Chan, J. K. L., Zhu, G. and Wu, Z. 2005. Myocyte enhancer factor 2 acetylation by p300 enhances its DNA binding activity, transcriptional activity, and myogenic differentiation. *Molecular and Cellular Biology* 25(9), pp. 3575-3582. doi: 10.1128/MCB.25.9.3575-3582.2005

MacKay, T. F. C. et al. 2012. The Drosophila melanogaster Genetic Reference Panel. *Nature* 482(7384), pp. 173-178. doi: 10.1038/nature10811

Main, P., Tan, W. J., Wheeler, D. and Fitzsimons, H. L. 2021. Increased Abundance of Nuclear HDAC4 Impairs Neuronal Development and Long-Term Memory. *Frontiers in Molecular Neuroscience* 14, doi: 10.3389/fnmol.2021.616642

Makarova, K. S. et al. 2011. Evolution and classification of the CRISPR-Cas systems. *Nature Reviews Microbiology* 9(6), pp. 467-477. doi: 10.1038/nrmicro2577

Makarova, K. S. et al. 2015. An updated evolutionary classification of CRISPR-Cas systems. *Nature Reviews Microbiology* 13(11), pp. 722-736. doi: 10.1038/nrmicro3569

Maqbool, T. and Jagla, K. 2007. Genetic control of muscle development: Learning from Drosophila. *Journal of Muscle Research and Cell Motility* 28(7-8), pp. 397-407. doi: 10.1007/s10974-008-9133-1

Markstein, M., Pitsouli, C., Villalta, C., Celniker, S. E. and Perrimon, N. 2008. Exploiting position effects and the gypsy retrovirus insulator to engineer precisely expressed transgenes. *Nature Genetics* 40(4), pp. 476-483. doi: 10.1038/ng.101

Marmorstein, R. and Zhou, M. M. 2014. Writers and readers of histone acetylation: Structure, mechanism, and inhibition. *Cold Spring Harbor perspectives in biology* 6(7), doi: 10.1101/cshperspect.a018762

McCullock, T. W., MacLean, D. M. and Kammermeier, P. J. 2020. Comparing the performance of mScarlet-I, mRuby3, and mCherry as FRET acceptors for mNeonGreen. *PLoS ONE* 15(2), doi: 10.1371/journal.pone.0219886

McGuire, S. E., Mao, Z. and Davis, R. L. 2004. Spatiotemporal gene expression targeting with the TARGET and gene-switch systems in Drosophila. *Science's STKE : signal transduction knowledge environment* 2004(220), p. pl6. doi: 10.1126/stke.2202004pl6

McKinsey, T. A., Chun, L. Z. and Olson, E. N. 2001. Identification of a signal-responsive nuclear export sequence in class II histone deacetylases. *Molecular and Cellular Biology* 21(18), pp. 6312-6321. doi: 10.1128/MCB.21.18.6312-6321.2001

McKinsey, T. A., Zhang, C. L., Lu, J. and Olson, E. N. 2000a. Signal-dependent nuclear export of a histone deacetylase regulates muscle differentiation. *Nature* 408(6808), pp. 106-111. doi: 10.1038/35040593

McKinsey, T. A., Zhang, C. L. and Olson, E. N. 2000b. Activation of the myocyte enhancer factor-2 transcription factor by calcium/calmodulin-dependent protein kinase-stimulated binding of 14-3-3 to histone deacetylase 5. *Proceedings of the National Academy of Sciences of the United States of America* 97(26), pp. 14400-14405. doi: 10.1073/pnas.260501497

Meltzer, H. et al. 2019. Tissue-specific (ts)CRISPR as an efficient strategy for in vivo screening in Drosophila. *Nature Communications* 10(1), doi: 10.1038/s41467-019-10140-0

Mercuri, E., Bönnemann, C. G. and Muntoni, F. 2019. Muscular dystrophies. *The Lancet* 394(10213), pp. 2025-2038. doi: 10.1016/S0140-6736(19)32910-1

Miller, D. E., Cook, K. R. and Scott Hawley, R. 2019. The joy of balancers. *PLoS Genetics* 15(11), doi: 10.1371/journal.pgen.1008421

Miska, E. A., Karlsson, C., Langley, E., Nielsen, S. J., Pines, J. and Kouzarides, T. 1999. HDAC4 deacetylase associates with and represses the MEF2 transcription factor. *EMBO Journal* 18(18), pp. 5099-5107. doi: 10.1093/emboj/18.18.5099

Miska, E. A., Langley, E., Wolf, D., Karlsson, C., Pines, J. and Kouzarides, T. 2001. Differential localization of HDAC4 orchestrates muscle differentiation. *Nucleic Acids Research* 29(16), pp. 3439-3447. doi: 10.1093/nar/29.16.3439

Molkentin, J. D., Black, B. L., Martin, J. F. and Olson, E. N. 1995. Cooperative activation of muscle gene expression by MEF2 and myogenic bHLH proteins. *Cell* 83(7), pp. 1125-1136. doi: 10.1016/0092-8674(95)90139-6

Molkentin, J. D., Black, B. L., Martin, J. F. and Olson, E. N. 1996a. Mutational analysis of the DNA binding, dimerization, and transcriptional activation domains of MEF2C. *Molecular and Cellular Biology* 16(6), pp. 2627-2636. doi: 10.1128/MCB.16.6.2627

Molkentin, J. D. et al. 1996b. MEF2B is a potent transactivator expressed in early myogenic lineages. *Molecular and Cellular Biology* 16(7), pp. 3814-3824. doi: 10.1128/MCB.16.7.3814

Molkentin, J. D. and Olson, E. N. 1996. Combinatorial control of muscle development by basic helix-loop-helix and MADS-box transcription factors. *Proceedings of the National Academy of Sciences of the United States of America* 93(18), pp. 9366-9373. doi: 10.1073/pnas.93.18.9366

Morgan, T. H. 1910. Sex limited inheritance in drosophila. *Science* 32(812), pp. 120-122. doi: 10.1126/science.32.812.120

Nagarkar-Jaiswal, S. et al. 2015. A library of MiMICs allows tagging of genes and reversible, spatial and temporal knockdown of proteins in Drosophila. *eLife* 2015(4), doi: 10.7554/eLife.05338

Nakajima, Y. I., Lee, Z. T., McKinney, S. A., Swanson, S. K., Florens, L. and Gibson, M. C. 2019. Junctional tumor suppressors interact with 14-3-3 proteins to control planar spindle alignment. *Journal of Cell Biology* 218(6), pp. 1824-1838. doi: 10.1083/JCB.201803116

Narita, T., Weinert, B. T. and Choudhary, C. 2019. Functions and mechanisms of non-histone protein acetylation. *Nature Reviews Molecular Cell Biology* 20(3), pp. 156-174. doi: 10.1038/s41580-018-0081-3

Nelson, C. M. 2022. Mechanical Control of Cell Differentiation: Insights from the Early Embryo. *Annual Review of Biomedical Engineering* 24, pp. 307-322. doi: 10.1146/annurev-bioeng-060418-052527

Nguyen, H. T., Bodmer, R., Abmayr, S. M., McDermott, J. C. and Spoerel, N. A. 1994. D-mef2: A Drosophila mesoderm-specific MADS box-containing gene with a biphasic expression profile during embryogenesis. *Proceedings of the National Academy of Sciences of the United States of America* 91(16), pp. 7520-7524. doi: 10.1073/pnas.91.16.7520

Nguyen, T., Wang, J. and Schulz, R. A. 2002. Mutations within the conserved MADS box of the D-MEF2 muscle differentiation factor result in a loss of DNA binding ability and lethality in Drosophila. *Differentiation* 70(8), pp. 438-446. doi: 10.1046/j.1432-0436.2002.700806.x

Ni, J. Q. et al. 2011. A genome-scale shRNA resource for transgenic RNAi in Drosophila. *Nature Methods* 8(5), pp. 405-407. doi: 10.1038/nmeth.1592

Nikonova, E., Kao, S. Y. and Spletter, M. L. 2020. Contributions of alternative splicing to muscle type development and function. *Seminars in Cell and Developmental Biology* 104, pp. 65-80. doi: 10.1016/j.semcdb.2020.02.003

Nurk, S. et al. 2022. The complete sequence of a human genome. *Science* 376(6588), pp. 44-53. doi: 10.1126/science.abj6987

Nüsslein-volhard, C. and Wieschaus, E. 1980. Mutations affecting segment number and polarity in drosophila. *Nature* 287(5785), pp. 795-801. doi: 10.1038/287795a0

Ornatsky, O. I., Andreucci, J. J. and McDermott, J. C. 1997. A dominant-negative form of transcription factor MEF2 inhibits myogenesis. *Journal of Biological Chemistry* 272(52), pp. 33271-33278. doi: 10.1074/jbc.272.52.33271

Ornatsky, O. I. et al. 1999. Post-translational control of the MEF2A transcriptional regulatory protein. *Nucleic Acids Research* 27(13), pp. 2646-2654. doi: 10.1093/nar/27.13.2646

Oswald, F. et al. 2001. p300 acts as a transcriptional coactivator for mammalian Notch-1. *Molecular and Cellular Biology* 21(22), pp. 7761-7774. doi: 10.1128/MCB.21.22.7761-7774.2001

Panigrahi, A. and O'Malley, B. W. 2021. Mechanisms of enhancer action: the known and the unknown. *Genome biology* 22(1), doi: 10.1186/s13059-021-02322-1

Pennacchio, L. A., Bickmore, W., Dean, A., Nobrega, M. A. and Bejerano, G. 2013. Enhancers: Five essential questions. *Nature Reviews Genetics* 14(4), pp. 288-295. doi: 10.1038/nrg3458

Perkins, L. A. et al. 2015. The transgenic RNAi project at Harvard medical school: Resources and validation. *Genetics* 201(3), pp. 843-852. doi: 10.1534/genetics.115.180208

Perrimon, N., Pitsouli, C. and Shilo, B. Z. 2012. Signaling mechanisms controlling cell fate and embryonic patterning. *Cold Spring Harbor perspectives in biology* 4(8), doi: 10.1101/cshperspect.a005975

Poernbacher, I., Crossman, S., Joachim Kurth, Hisashi Nojima, Alberto Baena-Lopez, Cyrille Alexandre and Vincent, J.-P. 2019. Lessons in genome engineering: opportunities, tools and pitfalls. doi: http://dx.doi.org/10.1101/710871

Poovathumkadavil, P. and Jagla, K. 2020. Genetic Control of Muscle Diversification and Homeostasis: Insights from Drosophila. *Cells* 9(6), doi: 10.3390/cells9061543

Port, F. and Boutros, M. 2022. Tissue-Specific CRISPR-Cas9 Screening in Drosophila.*Methods in Molecular Biology*. Vol. 2540. pp. 157-176.

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

Port, F. et al. 2020. A large-scale resource for tissue-specific CRISPR mutagenesis in Drosophila. *eLife* 9, doi: 10.7554/eLife.53865

Postigo, A. A., Ward, E., Skeath, J. B. and Dean, D. C. 1999. zfh-1, the Drosophila homologue of ZEB, is a transcriptional repressor that regulates somatic myogenesis. *Molecular and Cellular Biology* 19(10), pp. 7255-7263. doi: 10.1128/mcb.19.10.7255

Potthoff, M. J., Arnold, M. A., McAnally, J., Richardson, J. A., Bassel-Duby, R. and Olson, E. N. 2007a. Regulation of skeletal muscle sarcomere integrity and postnatal muscle function by Mef2c. *Molecular and Cellular Biology* 27(23), pp. 8143-8151. doi: 10.1128/MCB.01187-07

Potthoff, M. J. and Olson, E. N. 2007. MEF2: A central regulator of diverse developmental programs. *Development* 134(23), pp. 4131-4140. doi: 10.1242/dev.008367

Potthoff, M. J. et al. 2007b. Histone deacetylase degradation and MEF2 activation promote the formation of slow-twitch myofibers. *Journal of Clinical Investigation* 117(9), pp. 2459-2467. doi: 10.1172/JCI31960

Preibisch, S., Saalfeld, S. and Tomancak, P. 2009. Globally optimal stitching of tiled 3D microscopic image acquisitions. *Bioinformatics* 25(11), pp. 1463-1465. doi: 10.1093/bioinformatics/btp184

Procter, J. B. et al. 2021. Alignment of Biological Sequences with Jalview.*Methods in Molecular Biology*. Vol. 2231. pp. 203-224.

Ran, F. A., Hsu, P. D., Wright, J., Agarwala, V., Scott, D. A. and Zhang, F. 2013. Genome engineering using the CRISPR-Cas9 system. *Nature Protocols* 8(11), pp. 2281-2308. doi: 10.1038/nprot.2013.143

Ranganayakulu, G., Zhao, B., Dokidis, A., Molkentin, J. D., Olson, E. N. and Schulz, R. A. 1995. A Series of Mutations in the D-MEF2 Transcription Factor Reveal Multiple Functions in Larval and Adult Myogenesis in Drosophila. *Developmental Biology* 171(1), pp. 169-181. doi: 10.1006/dbio.1995.1269

Reed, S. M. and Quelle, D. E. 2014. P53 acetylation: Regulation and consequences. *Cancers* 7(1), pp. 30-69. doi: 10.3390/cancers7010030

Reeves, R. G. and Tautz, D. 2017. Automated phenotyping indicates pupal size in Drosophila is a highly heritable trait with an apparent polygenic basis. *G3: Genes, Genomes, Genetics* 7(4), pp. 1277-1286. doi: 10.1534/g3.117.039883

Rembold, M. et al. 2014. A conserved role for Snail as a potentiator of active transcription. *Genes and Development* 28(2), pp. 167-181. doi: 10.1101/gad.230953.113

Ren, X. et al. 2013. Optimized gene editing technology for Drosophila melanogaster using germ line-specific Cas9. *Proceedings of the National Academy of Sciences of the United States of America* 110(47), pp. 19012-19017. doi: 10.1073/pnas.1318481110

Ren, X. et al. 2014. Enhanced specificity and efficiency of the CRISPR/Cas9 system with optimized sgRNA parameters in Drosophila. *Cell Reports* 9(3), pp. 1151-1162. doi: 10.1016/j.celrep.2014.09.044

Reynolds, N., O'Shaughnessy, A. and Hendrich, B. 2013. Transcriptional repressors: Multifaceted regulators of gene expression. *Development (Cambridge)* 140(3), pp. 505-512. doi: 10.1242/dev.083105

Rivera, J., Keränen, S. V. E., Gallo, S. M. and Halfon, M. S. 2019. REDfly: The transcriptional regulatory element database for Drosophila. *Nucleic Acids Research* 47(D1), pp. D828-D834. doi: 10.1093/nar/gky957

Ruiz-Gómez, M. 1998. Muscle patterning and specification in Drosophila. *International Journal of Developmental Biology* 42(3), pp. 283-290.

Ruiz-Gómez, M. and Bate, M. 1997. Segregation of myogenic lineages in Drosophila requires Numb. *Development* 124(23), pp. 4857-4866.

Ruiz-Gómez, M., Romani, S., Hartmann, C., Jäckle, H. and Bate, M. 1997. Specific muscle identities are regulated by Kruppel during Drosophila embryogenesis. *Development* 124(17), pp. 3407-3414.

Sachan, N., Sharma, V., Mutsuddi, M. and Mukherjee, A. 2023. Notch signalling: multifaceted role in development and disease. *FEBS Journal*, doi: 10.1111/febs.16815

Sandmann, T., Girardot, C., Brehme, M., Tongprasit, W., Stolc, V. and Furlong, E. E. M. 2007. A core transcriptional network for early mesoderm development in Drosophila melanogaster. *Genes and Development* 21(4), pp. 436-449. doi: 10.1101/gad.1509007

Sandmann, T., Jensen, L. J., Jakobsen, J. S., Karzynski, M. M., Eichenlaub, M. P., Bork, P. and Furlong, E. E. M. 2006. A Temporal Map of Transcription Factor Activity: Mef2 Directly Regulates Target Genes at All Stages of Muscle Development. *Developmental Cell* 10(6), pp. 797-807. doi: 10.1016/j.devcel.2006.04.009

Sarov, M. et al. 2016. A genome-wide resource for the analysis of protein localisation in Drosophila. *eLife* 5, doi: 10.7554/eLife.12068

Sartorelli, V., Huang, J., Hamamori, Y. and Kedes, L. 1997. Molecular mechanisms of myogenic coactivation by p300: Direct interaction with the activation domain of MyoD and with the MADS box of MEF2C. *Molecular and Cellular Biology* 17(2), pp. 1010-1026. doi: 10.1128/MCB.17.2.1010

Sasai, N., Toriyama, M. and Kondo, T. 2019. Hedgehog Signal and Genetic Disorders. *Frontiers in Genetics* 10, doi: 10.3389/fgene.2019.01103

Schindelin, J. et al. 2012. Fiji: An open-source platform for biological-image analysis. *Nature Methods* 9(7), pp. 676-682. doi: 10.1038/nmeth.2019

Schnorrer, F. et al. 2010. Systematic genetic analysis of muscle morphogenesis and function in Drosophila. *Nature* 464(7286), pp. 287-291. doi: 10.1038/nature08799

Schulman, V. K., Dobi, K. C. and Baylies, M. K. 2015. Morphogenesis of the somatic musculature in Drosophila melanogaster. *Wiley Interdisciplinary Reviews: Developmental Biology* 4(4), pp. 313-334. doi: 10.1002/wdev.180

Schwartz, S., Truglio, M., Scott, M. J. and Fitzsimons, H. L. 2016. Long-term memory in Drosophila is influenced by histone deacetylase HDAC4 interacting with SUMO-Conjugating enzyme UBC9. *Genetics* 203(3), pp. 1249-1264. doi: 10.1534/genetics.115.183194

Shih, J., Hodge, R. and Andrade-Navarro, M. A. 2015. Comparison of inter- and intraspecies variation in humans and fruit flies. *Genomics Data* 3, pp. 49-54. doi: 10.1016/j.gdata.2014.11.010

Shlyueva, D., Stampfel, G. and Stark, A. 2014. Transcriptional enhancers: From properties to genome-wide predictions. *Nature Reviews Genetics* 15(4), pp. 272-286. doi: 10.1038/nrg3682

Shokri, L. et al. 2019. A Comprehensive Drosophila melanogaster Transcription Factor Interactome. *Cell Reports* 27(3), pp. 955-970.e957. doi: 10.1016/j.celrep.2019.03.071

Shore, P. and Sharrocks, A. D. 1995. The MADS-Box Family of Transcription Factors. *European Journal of Biochemistry* 229(1), pp. 1-13. doi: 10.1111/j.1432-1033.1995.tb20430.x

Shvedunova, M. and Akhtar, A. 2022. Modulation of cellular processes by histone and non-histone protein acetylation. *Nature Reviews Molecular Cell Biology* 23(5), pp. 329-349. doi: 10.1038/s41580-021-00441-y

Singh, A. J., Ramsey, S. A., Filtz, T. M. and Kioussi, C. 2018. Differential gene regulatory networks in development and disease. *Cellular and Molecular Life Sciences* 75(6), pp. 1013-1025. doi: 10.1007/s00018-017-2679-6

Soler, C., Han, J. and Taylor, M. V. 2012. The conserved transcription factor Mef2 has multiple roles in adult Drosophila musculature formation. *Development* 139(7), pp. 1270-1275. doi: 10.1242/dev.077875

Soler, C. and Taylor, M. V. 2009. The Him gene inhibits the development of Drosophila flight muscles during metamorphosis. *Mechanisms of Development* 126(7), pp. 595-603. doi: 10.1016/j.mod.2009.03.003

Spitz, F. and Furlong, E. E. M. 2012. Transcription factors: From enhancer binding to developmental control. *Nature Reviews Genetics* 13(9), pp. 613-626. doi: 10.1038/nrg3207

Spletter, M. L. et al. 2018. A transcriptomics resource reveals a transcriptional transition during ordered sarcomere morphogenesis in flight muscle. *eLife* 7, doi: 10.7554/eLife.34058

Stapleton, M. et al. 2002. A Drosophila full-length cDNA resource. *Genome biology* 3(12), doi: https://doi.org/10.1186/gb-2002-3-12-research0080

Stathopoulos, A. and Levine, M. 2004. Whole-genome analysis of Drosophila gastrulation. *Current Opinion in Genetics and Development* 14(5), pp. 477-484. doi: 10.1016/j.gde.2004.07.004

Stocker, H. and Gallant, P. 2008. Getting started : an overview on raising and handling Drosophila. *Methods in molecular biology (Clifton, N.J.)* 420, pp. 27-44. doi: 10.1007/978-1-59745-583-1_2

Sudarsan, V., Anant, S., Guptan, P., Vijayraghavan, K. and Skaer, H. 2001. Myoblast Diversification and Ectodermal Signaling in Drosophila. *Developmental Cell* 1(6), pp. 829-839. doi: 10.1016/S1534-5807(01)00089-2

Tanaka, K. K., Bryantsev, A. L. and Cripps, R. M. 2008. Myocyte enhancer factor 2 and chorion factor 2 collaborate in activation of the myogenic program in Drosophila. *Molecular and Cellular Biology* 28(5), pp. 1616-1629. doi: 10.1128/MCB.01169-07

Taylor, M. V., Beatty, K. E., Hunter, H. K. and Baylies, M. K. 1995. Drosophila MEF2 is regulated by twist and is expressed in both the primordia and differentiated cells of the embryonic somatic, visceral and heart musculature. *Mechanisms of Development* 50(1), pp. 29-41. doi: 10.1016/0925-4773(94)00323-F

Taylor, M. V. and Hughes, S. M. 2017. Mef2 and the skeletal muscle differentiation program. *Seminars in Cell and Developmental Biology* 72, pp. 33-44. doi: 10.1016/j.semcdb.2017.11.020

Thibault, S. T. et al. 2004. A complementary transposon tool kit for Drosophila melanogaster using P and piggyBac. *Nature Genetics* 36(3), pp. 283-287. doi: 10.1038/ng1314

Tixier, V., Bataillé, L. and Jagla, K. 2010. Diversification of muscle types: Recent insights from Drosophila. *Experimental Cell Research* 316(18), pp. 3019-3027. doi: 10.1016/j.yexcr.2010.07.013

Tolwinski, N. S. 2017. Introduction: Drosophila-A model system for developmental biology. *Journal of Developmental Biology* 5(3), doi: 10.3390/jdb5030009

Tripathi, B. K. and Irvine, K. D. 2022. The wing imaginal disc. *Genetics* 220(4), doi: 10.1093/genetics/iyac020

Vega, R. B. et al. 2004. Histone deacetylase 4 controls chondrocyte hypertrophy during skeletogenesis. *Cell* 119(4), pp. 555-566. doi: 10.1016/j.cell.2004.10.024

Venken, K. J. T. and Bellen, H. J. 2007. Transgenesis upgrades for Drosophila melanogaster. *Development* 134(20), pp. 3571-3584. doi: 10.1242/dev.005686

Vishal, K., Barajas Alonso, E., DeAguero, A. A., Waters, J. A., Chechenova, M. B. and Cripps, R. M. 2023. Phosphorylation of the Myogenic Factor Myocyte Enhancer Factor-2 Impacts Myogenesis In Vivo. *Molecular and Cellular Biology*, doi: 10.1080/10985549.2023.2198167

Vishal, K., Brooks, D. S., Bawa, S., Gameros, S., Stetsiv, M. and Geisbrecht, E. R. 2017. Adult muscle formation requires Drosophila moleskin for proliferation of wing disc-associated muscle precursors. *Genetics* 206(1), pp. 199-213. doi: 10.1534/genetics.116.193813

Wallberg, A. E., Pedersen, K., Lendahl, U. and Roeder, R. G. 2002. p300 and PCAF act cooperatively to mediate transcriptional activation from chromatin templates by Notch intracellular domains in vitro. *Molecular and Cellular Biology* 22(22), pp. 7812-7819. doi: 10.1128/MCB.22.22.7812-7819.2002

Wang, A. H. et al. 1999. HDAC4, a human histone deacetylase related to yeast HDA1, is a transcriptional corepressor. *Molecular and Cellular Biology* 19(11), pp. 7816-7827. doi: 10.1128/MCB.19.11.7816

Wang, A. H. and Yang, X. J. 2001. Histone deacetylase 4 possesses intrinsic nuclear import and export signals. *Molecular and Cellular Biology* 21(17), pp. 5992-6005. doi: 10.1128/MCB.21.17.5992-6005.2001

Wang, B. et al. 2011. A hormone-dependent module regulating energy balance. *Cell* 145(4), pp. 596-606. doi: 10.1016/j.cell.2011.04.013

Wang, D. Z., Renee Valdez, M., McAnally, J., Richardson, J. and Olson, E. N. 2001. The Mef2c gene is a direct transcriptional target of myogenic bHLH and MEF2 proteins during skeletal muscle development. *Development* 128(22), pp. 4623-4633.

Wang, J. Y. and Doudna, J. A. 2023. CRISPR technology: A decade of genome editing is only the beginning. *Science* 379(6629), doi: 10.1126/science.add8643

Wang, L., Tang, Y., Cole, P. A. and Marmorstein, R. 2008. Structure and chemistry of the p300/CBP and Rtt109 histone acetyltransferases: implications for histone acetyltransferase evolution and function. *Current Opinion in Structural Biology* 18(6), pp. 741-747. doi: 10.1016/j.sbi.2008.09.004

Wang, S., Shen, B., Ren, S., Zhao, Y., Zhang, S., Qu, J. and Liu, L. 2019. Implementation and application of FRET-FLIM technology. *Journal of Innovative Optical Health Sciences* 12(5), doi: 10.1142/S1793545819300106

Weitkunat, M. and Schnorrer, F. 2014. A guide to study Drosophila muscle biology. *Methods* 68(1), pp. 2-14. doi: 10.1016/j.ymeth.2014.02.037

Whyte, W. A. et al. 2013. Master transcription factors and mediator establish super-enhancers at key cell identity genes. *Cell* 153(2), pp. 307-319. doi: 10.1016/j.cell.2013.03.035

Wilczyski, B. and Furlong, E. E. M. 2010. Dynamic CRM occupancy reflects a temporal map of developmental progression. *Molecular Systems Biology* 6, doi: 10.1038/msb.2010.35

Wodarz, A., Hinz, U., Engelbert, M. and Knust, E. 1995. Expression of crumbs confers apical character on plasma membrane domains of ectodermal epithelia of drosophila. *Cell* 82(1), pp. 67-76. doi: 10.1016/0092-8674(95)90053-5

Wu, Z. et al. 2000. p38 and extracellular signal-regulated kinases regulate the myogenic program at multiple steps. *Molecular and Cellular Biology* 20(11), pp. 3951-3964. doi: 10.1128/MCB.20.11.3951-3964.2000

Xu, J. et al. 2022. Protein visualization and manipulation in Drosophila through the use of epitope tags recognized by nanobodies. *eLife* 11, doi: 10.7554/eLife.74326

Yang, X. J. and Seto, E. 2007. HATs and HDACs: From structure, function and regulation to novel strategies for therapy and prevention. *Oncogene* 26(37), pp. 5310-5318. doi: 10.1038/sj.onc.1210599

Yang, X. J. and Seto, E. 2008. The Rpd3/Hda1 family of lysine deacetylases: From bacteria and yeast to mice and men. *Nature Reviews Molecular Cell Biology* 9(3), pp. 206-218. doi: 10.1038/nrm2346

Zappia, M. P., de Castro, L., Ariss, M. M., Jefferson, H., Islam, A. B. M. M. K. and Frolov, M. V. 2020. A cell atlas of adult muscle precursors uncovers early events in fibre-type divergence in Drosophila. *EMBO Reports* 21(10), doi: 10.15252/embr.201949555

Zeremski, M., Stricker, J. R., Fischer, D., Zusman, S. B. and Cohen, D. 2003. Histone deacetylase dHDAC4 is involved in segmentation of the Drosophila embryo and is regulated by gap and pair-rule genes. *Genesis* 35(1), pp. 31-38. doi: 10.1002/gene.10159

Zhang, C. L., McKinsey, T. A., Chang, S., Antos, C. L., Hill, J. A. and Olson, E. N. 2002. Class II histone deacetylases act as signal-responsive repressors of cardiac hypertrophy. *Cell* 110(4), pp. 479-488. doi: 10.1016/S0092-8674(02)00861-9

Zhao, M. et al. 1999. Regulation of the MEF2 family of transcription factors by p38. *Molecular and Cellular Biology* 19(1), pp. 21-30. doi: 10.1128/MCB.19.1.21

Zia, A. and Rashid, S. 2021. Systematic transition modeling analysis in the MEF2B-DNA binding interface due to Y69H and K4E variants. *Journal of Molecular Graphics and Modelling* 108, doi: 10.1016/j.jmgm.2021.108009

Zirin, J., Bosch, J., Viswanatha, R., Mohr, S. E. and Perrimon, N. 2022. State-of-the-art CRISPR for in vivo and cell-based studies in Drosophila. *Trends in Genetics* 38(5), pp. 437-453. doi: 10.1016/j.tig.2021.11.006