7-1-2014

DEFINING THE COMPLEXITIES OF MEF2 FUNCTION AND REGULATION IN DROSOPHILA

TyAnna L. Lovato

TyAnna L. Lovato
Candidate

Biology
Department

This dissertation is approved, and it is acceptable in quality and form for publication:

Approved by the Dissertation Committee:

Richard M. Cripps, Chairperson

Stephen A. Stricker

Coenraad Adema

Helen J. Hathaway

DEFINING THE COMPLEXITIES OF *MEF2* FUNCTION AND REGULATION IN *DROSOPHILA*

By

TYANNA L. LOVATO

Bachelor of Science, University of New Mexico, 1998

DISSERTATION

Submitted in Partial Fulfillment of the
Requirements for the Degree of

Doctor of Philosophy

Biology

The University of New Mexico
Albuquerque, New Mexico

July, 2014

ACKNOWLEDGMENTS

I would like to thank my advisor, Richard Cripps for his guidance, flexibility and support. I would like to thank my committee members, Steve Stricker, Coen Adema and Helen Hathaway for their time, positive feedback and encouragement. I would also like to thank the entire Cripps' Lab for moral and technical support.

On a personal note, I would like to thank my husband Jason and my two sons Jaylen and Devin for all of their love and understanding. I would like to thank my brother Ryan for allowing me to complain. And lastly, I would like to thank all of my extended family for their encouragement and support.

DEFINING THE COMPLEXITIES OF *MEF2* FUNCTION AND REGULATION IN *DROSOPHILA*

By

TyAnna L. Lovato

Bachelor of Science, Biology, University of New Mexico

DISSERTATION

Submitted in Partial Fulfillment of the
Requirements for the Degree of

Doctor of Philosophy

Biology

The University of New Mexico
Albuquerque, New Mexico

July, 2014

ABSTRACT

The development of cardiac and somatic muscle is coordinated by many factors that are highly conserved across species. Mutations in the coding or regulatory sequences of a gene can alter the structure, function or levels of expression of the protein it encodes. We have identified the genetic mechanism underlying the temperature sensitive mutation in MEF2 (myocyte enhancer factor 2), which is required for the differentiation of all muscle lineages. In addition, we have identified a regulatory element upstream of the *Mef2* gene that directs its expression from stage fourteen through the end of embryogenesis which is activated by the NK-homeobox transcription factor Tinman and the GATA transcription factor Pannier. Additionally, we have found that Tinman and Pannier work in collaboration with MEF2 to activate the entire cardiac program. Finally, we have purified cardiac cells from *Drosophila* embryos and have paired-end sequenced their mRNA in order to identify novel factors that contribute to heart development.

TABLE OF CONTENTS

LIST OF FIGURES

CHAPTER 1

CHAPTER 2

INTRODUCTION

In this study, we utilize the model organism *Drosophila melanogaster*. Its life cycle from embryo to adult spans a short 10 days and the framework for its development is laid out during the first 24 hours of embryogenesis. Although *Drosophila* is only distantly related to vertebrates, its heart development is remarkably similar to early vertebrate heart development. For example, the *Drosophila* heart is initially specified in two rows of cells that lie on either side of the embryo. These cells migrate dorsally to eventually form a closed linear tube that is capable of taking in hemolymph through inflow tracts and pumping it anteriorly through the animal to circulate nutrients.

Human heart development begins much the same way. Bilateral rows of cells migrate to form a linear tube that lies in the ventral plane of the animal. This tube is capable of beating at approximately 21 days of development. While the vertebrate tube goes on to form a more complicated structure, this initial linear tube contains inflow tracts and the capability of pumping oxygen-containing blood throughout the animal (Bodmer and Frasch, 2010). Not only is this early structure similar to the linear tube in *Drosophila*, but the factors that direct its formation are also highly conserved. The realization that extensive similarities exist began with the discovery of the homeobox transcription factor Tinman in *Drosophila* and its homologue in vertebrates Nkx2-5 (Bodmer, 1993; Lints et al., 1993). The function of Tinman/Nkx2-5 proteins are not identical in the two organisms, but both are absolutely required for the formation of the dorsal vessel or vertebrate heart, and

since this initial discovery an entire homologous regulatory network has been discovered.

In fact, a core of cardiac transcription factors exists in all modern organisms that possess a heart suggesting that the common ancestor of bilaterians contained a primitive heart that was created through this conserved toolkit. As the heart increased in complexity, many of the genes encoding this core of regulatory factors duplicated and their expression patterns became more specialized to allow for the formation of increasingly complicated circulatory systems (Cripps and Olson, 2002; Olson, 2006).

For various reasons, *Drosophila* has been a very useful model for understanding the interactions and hierarchy of this core of cardiac regulators. First, with few exceptions, the genome of *Drosophila* has not undergone extensive duplication. For example, null-mutations in any of the core factors in *Drosophila* result in very distinct cardiac phenotypes whereas in vertebrates, a null-mutation might not have an effect due to compensation from a duplicated gene with overlapping function. Second, as a result of revolutionary sequencing techniques, documentation by Online Mendelian Inheritance in Man (OMIM) has determined that of the 1682 human genes that have been identified to have a mutation that causes disease, 74% have a homologous gene in *Drosophila.* Additionally, one third of these homologies have been show to have functional conservation (Bier and Bodmer, 2004). More specifically, mutations in many of the genes identified in *Drosophila* heart development have been shown to cause coronary heart disease in humans (Schott et al., 1998; Basson et al., 1997; Garg

et al., 2003). Researchers would not have readily identified these genes in humans without the groundwork that had been laid in studies of *Drosophila* and other models. Finally, the *Drosophila* community consists of a massive collaboration among scientists whose cooperation has expedited the resolution of many important questions. No other model organism possesses such vast libraries of mutants, transgenics, RNAi lines, deletion kits, databases and community support.

My studies are specifically focused on the regulatory genes that direct heart development. In particular, I am interested in the mechanisms by which regulatory genes themselves are regulated. Understanding the specific mechanisms that control their regulatory gene actions at a basic level can provide insight into the developmental problems that might arise in higher organisms.

In our present study, we critically examine the transcription factor MEF2. The MADS box transcription factor MEF2 (Myocyte Enhancer Factor 2) is required for the differentiation of all muscle lineages during *Drosophila* embryogenesis (Cripps et al. 1999) and is expressed in the nervous system later in development (Schulz et al., 1996). *Mef2* contains numerous regulatory enhancers in the approximately 15 kb sequence 5' to the transcription start site that specifically direct expression in different muscle types (Schulz et al., 1996, Nguyen and Xu, 1998, Cripps et al. 1999). Not surprisingly, null mutants of such a widely expressed transcription factor fail to differentiate muscle properly and are not viable (Lilly et al. 1995). Vertebrates possess four Mef2 genes: A-D with

overlapping expression in all of the muscle lineages but also in the nervous system, the immune system, blood cells, neural crest cells, vascular endothelium and chondrocytes during bone formation (reviewed in Black and Cripps, 2010, Potthoff and Olson, 2007). In the vertebrate heart, knockouts of Mef2a and Mef2c are lethal due to severe cardiac defects (Lin et al., 1997; Naya et al., 2002). Mef2d does not appear to be important for heart development and Mef2b is still being investigated (Reviewed in Black and Cripps, 2010).

The complex mechanisms underlying the function of a single transcription factor are examined from many different aspects in this manuscript. Single nucleotide mutations in the coding sequence of a regulatory protein can affect its function at many levels. For example, DNA binding ability can be impaired, interactions with cofactors and other regulatory genes might be altered and/or the ability to activate transcription might be affected. Additionally, mutations in non-coding, regulatory sequences can affect expression levels of a gene. *Mef2* is an example of this type of complexity. Nguyen and Xu (1998) first described more enhancers for *Mef2* than the number of tissues in which it is expressed. Gunthorpe et al. (1999) carried out rescue experiments of *Mef2* null mutants and discovered that each tissue required different expression levels of MEF2 protein to develop properly. Their findings provide an explanation for exactly why so many enhancers are required for precisely titering the correct level of MEF2 to allow proper morphogenesis. Clearly, understanding both the regulation and function of key regulatory genes such as *Mef2* is critical to defining the network of factors that control muscle development.

In the first chapter of this dissertation, I characterize the genetic changes underlying a temperature sensitive MEF2 mutant and identify critical residues in the MEF2 protein. I also used this characterization to predict the creation of new temperature sensitive combinations.

In the second chapter, I identified a cardiac enhancer of *Mef2,* which is novel in that it is expressed in all cardiac cell types at a late stage in embryogenesis, and demonstrate that it is activated by the cardiogenic factors Tinman and Pannier. Such findings support the idea that proper heart development requires the fine-tuning of additional MEF2 protein expression before hatching to the larval stage. The network of tissue specific co-factors and co-regulators MEF2 collaborates with further complicates the diversity of MEF2 expression and regulation. The NK-homeodomain factor Tinman and the GATA factor Pannier work to activate *Mef2* (Klinedinst and Bodmer, 2003, Hendren et al., 2007, Gajewski et al., 1999, Gajewski et al., 2001, Han and Olson, 2005) and in the second chapter I also demonstrate that these factors work with MEF2 protein to activate downstream genes in the cardiac program. This latter finding is a novel discovery in *Drosophila* but has been demonstrated in vertebrate cell culture with the conversion of mouse fibroblast cells into cardioblasts with the simple transfection of MEF2, Tinman and Pannier homologs (Ieda et al., 2010 and Ifkovits et al., 2014). Such findings demonstrate that this phenomenon holds true in *Drosophila* as well and will provide a convenient system for further investigation.

Despite our current understanding of how regulatory factors specify the developing heart, there are still thousands of predicted *Drosophila* genes that have yet to be characterized. Fortunately, recent technology has improved the prospects of sequencing mRNA from very small amounts of starting material. In the final chapter, I demonstrate how we have overcome the technical obstacle to purify 104 heart cells from the many thousands of cells present at the end of *Drosophila* embryogenesis. I then determine their transcriptome in order to identify novel factors involved in heart development.

CHAPTER 1

A Molecular Mechanism of Temperature Sensitivity for Mutations Affecting the *Drosophila* Muscle Regulator Myocyte Enhance Factor-2

From Genetics 183: 107-117

TyAnna L. Lovato, Melanie M. Adams, Phillip W. Baker and Richard M. Cripps

Abstract

Temperature sensitive (TS) mutants are a useful tool for elucidating gene function where a gene of interest is essential at multiple stages of development, however, the molecular mechanisms behind TS alleles vary from mutant to mutant. TS for the myogenic transcription factor Myocyte enhancer factor-2 (MEF2) in Drosophila arises in the heteroallelic combination of two *Mef2* alleles, *30-5* and *44-5*. We show that the *30-5* mutation affects the N-terminal DNA binding domain. This results in impaired DNA binding ability, where the *30-5* homozygotes display activation of downstream target genes and initiation of the myogenic program but cannot survive to adulthood. The *44-5* mutation deletes a downstream splice acceptor site, retaining intronic sequence in the mature transcript, and resulting in a severely truncated protein that is unable to activate MEF2 targets. *44-5* homozygotes consequently show severely impaired myogenesis. We propose that in heteroallelic mutants at the permissive temperature, 30-5/44-5 heterodimers can form and they have a sufficiently stable interaction with DNA to activate myogenic gene expression; at the restrictive temperature, 44-5 homodimers displace 30-5/44-5 heterodimers from target genes, thus acting in a dominant-negative manner. To test this model, we studied additional *Mef2* alleles for their ability to complement the *30-5* allele. An allele affecting the DNA binding domain failed to complement *30-5* at any temperature, whereas two alleles affecting downstream residues also showed temperature-dependent complementation. Thus, by combining one MEF2 isoform having weakened DNA binding ability with a second truncated MEF2

mutant that has lost its activation ability, a TS form of intragenic complementation can be generated. These findings will provide new insight and guidance into the functions of dimeric proteins, and how they might be engineered to generate TS combinations.

Introduction

Temperature sensitive (TS) mutants have a noted history in defining gene function. Using TS alleles, geneticists are able to turn genes off at a specific time of their choosing and then observe the phenotypic effect of their genetic manipulation. TS mutants have been used for such diverse studies as identifying yeast genes critical in the various stages of the cell cycle (Hartwell et al. 1970), and in metazoans to dissect the requirements for genes which function at multiple developmental stages (see for example COX and BAYLIES 2005). This powerful tool has proven especially useful in teasing out and categorizing genes which function at multiple stages to support organism viability.

Defining the molecular mechanisms behind temperature-sensitivity enables us to more thoroughly understand the synthesis, folding, collaboration and function of a protein of interest. Sadler and Novick (1964) were one of the first groups to categorize TS mutants by the mechanism of their dysfunction, using the bacteriophage system (Sadler AND Novick 1964). The TL (thermolabile) class of mutants demonstrated a mutant phenotype when grown at a restrictive temperature or when shifted to a restrictive temperature during a later stage of development or adulthood. The mutant phenotype in these types of TS mutants was usually due to the destabilization of the encoded protein, and

subsequent loss of protein function as a result of increased temperature. This could be due to decreased melting temperature from the loss of an hydrophobic amino acid, or due to decreased ability to interact with DNA or other proteins as dimers or multimers.

Another category of TS mutants, TSF (temperature sensitive folding) only showed a phenotype if incubated at the restrictive temperature during synthesis. If they were shifted to the restrictive temperature at a later stage, the protein did not lose its conformation and no phenotype was apparent. These mutants were recessive and did not demonstrate intragenic complementation as did the thermolabile types. It was postulated that many mutants in this class affect the formation, folding or initial integration of the encoded protein into larger complexes (Gordon and King 1993; Edgar and Lielausis 1963). Clearly, the mechanism of TS for a particular protein is strongly dependent upon its function as part of a multimeric or macromolecular complex, and upon its half-life in the cell.

TS mutants have generally been isolated by screening through large numbers of organisms that have undergone random mutagenesis. More recently, scientists have looked for a way to engineer TS proteins. One approach was to introduce charged amino acids into the internal region of the CcdB toxin encoded by the F' plasmid (Chakshusmathi *et al*. 2004). CcdB is a good candidate for TS screening because when transformed into *Escherichia coli*, it is lethal when functional. Internally charged amino acids served to successfully destabilize the protein at the restrictive temperature in many cases; however it was not always

fully effective. Clearly, the engineering of new TS mutants will benefit from further characterization of the molecular basis of existing TS alleles.

We have recently described the isolation of a temperature-sensitive combination of alleles for the muscle transcription factor Myocyte enhancer factor 2 (MEF2) (Baker et al. 2005). Vertebrates possess four copies of MEF2, all of which contain a highly conserved, N-terminal MADS domain [MCM1, Agamous, Deficiens, serum response factor (SRF)] spanning amino acids 1 through 57, and a 29 amino acid MEF2 domain immediately adjacent (reviewed in Black and Olson 1998; Black and Cripps 2009). Both mutagenesis studies and biophysical determination of the solution structure of MEF2A, have clearly shown that the MADS domain is required for DNA binding; also, a portion of the MADS domain (residues 35-50) along with the adjacent MEF2 domain, is required for proper dimerization (Molkentin *et al.* 1996a; Santelli and Richmond, 2000; Huang *et al.* 2000).

Upon dimerization and DNA binding, MEF2 factors are potent activators of target structural genes, both directly and in complex with other factors (Molkentin *et al.* 1995; Hamamori *et al.* 1997; Kelly Tanaka *et al*. 2008). Much of this transcriptional activation ability of MEF2 proteins arises from the function of residues located C-terminal to the MEF2 domain, which are both required and sufficient for MEF2 transcriptional activity (Martin *et al*., 1994; Wong *et al*., 1995; Molkentin *et al.*, 1995). Clearly, MEF2 is a highly modular protein with distinct functional domains.

Drosophila contains a single *Mef2* gene which is expressed in muscles at all stages of development (Nguyen *et al*. 1994; Lilly *et al*, 1994; Taylor *et al*, 1995; Baker *et al*. 2005). *Mef2* null mutants do not survive beyond embryogenesis and display a profound lack of differentiated muscle (Bour *et al*., 1995; Lilly *et al.,* 1995), making this system a useful model for the study of MEF2 protein function. In this paper, we use this Drosophila system to define a molecular mechanism for the temperature sensitivity of a subset of *Mef2* mutant alleles. We show that a mutation affecting the MADS domain (allele *30-5*) can complement a mutation affecting C-terminal residues (allele *44-5)*, but that this complementation is thermolabile. We provide evidence that the thermolability arises from competition between the attenuated DNA binding ability of the 30-5 isoform, which can function effectively at the permissive temperature; and the strong DNA binding ability of the 44-5 isoform, which acts as a dominant-negative factor at the restrictive temperature. Our analyses of several additional *Mef2* mutants confirm this model. These studies provide important new insight into critical regions of the MEF2 factor, and more broadly contribute to our understanding of the molecular basis of TS mutants.

Materials and Methods

Drosophila stocks and crosses

Drosophila stocks were grown on Carpenter's medium (Carpenter 1950) at the indicated temperatures. *Mef2* mutant stocks were obtained from Elliot Goldstein (Arizona State University; Goldstein *et al.* 2001) and balanced over a *CyO, Cy Kr-GFP* balancer chromosome (Casso *et al.* 1999) to visualize

homozygous or heteroallelic mutants at any stage of development via the lack of GFP expression. Viability studies were achieved by crossing *CyO, Cy Kr-GFP / Mef2*^{x} with *CyO, Cy Kr-GFP / Mef2*^{y}. We counted the number of heteroallelic escaper adults and the total number of progeny. Viability counts reflect the percentage of heteroallelic escapers detected, as a proportion of the number of such escapers expected if their viability were normal.

Immunohistochemistry

Embryos were collected and stained as described by Patel (1994). Primary antibodies were rabbit anti-Myosin heavy-chain (1:250; (Kiehart and Feghali 1986), mouse anti –GFP (1:250; Invitrogen Corp.), and rat anti-Tropomyosin (1:250; Peckham *et al*., 1992, Abcam Immunochemicals). Secondary detection was using the Vectastain Elite kit (Vector Laboratories) and diaminobenzidine (DAB) stain. Samples were cleared in glycerol and mounted for photography with an Olympus BX51 photomicroscope using DIC optics. Images were collated using Adobe Photoshop.

DNA and RNA methods: To identify the lesions associated with each *Mef2* allele, RNA was purified from larvae of the genotype *CyO, Cy Kr-GFP / Mef2*^{x}. The RNA was purified using the Qiagen RNeasy Mini kit and RNase-free DNase set, and then subjected to reverse transcription followed by PCR, using Invitrogen Superscript III One-Step RT-PCR. Primers for the PCR were (5'-ATGGGCCGCAAAAAAATTCAAATATC-3') and (5'-CTATGTGCCCCATCCGCC-3'), which were designed to amplify the entire *Mef2* coding region. PCR products were cloned into the pGEM-T Easy vector (Promega Corp.), and twelve positive

clones for each genotype were sequenced in their entirety, to ensure that a cDNA arising from the mutant *Mef2* allele was sequenced several times. In all cases, we found several wild-type cDNA clones and several clones which consistently showed a departure from wild-type.

To confirm the sequence alterations observed in the cDNA as arising from the *Mef2* mutant allele, we also isolated DNA from heterozygotes and amplified the appropriate genomic region by PCR. Primers used to characterize *44-5* splice variants were (5'-GCTGGAGATGTCGAACG-3' and 5'-CTGCATATCCCACATCATCC-3'). For all mutants characterized, the lesion at the DNA level matched or accounted for the changes observed in the cDNAs.

Electrophoretic mobility shift assays

Wild-type MEF2 protein was generated from pSK-DMef2 (Lilly *et al*., 1995) using T3 RNA polymerase and the TnT Coupled Transcription/Translation system (Promega). For mutant MEF2 proteins, the appropriate sequence change was introduced into pSK-DMef2 using the Invitrogen Gene-Tailor Site Directed Mutagenesis Kit, and confirmed by sequencing. Protein was then generated as described for wild-type. In order to use equivalent amounts of protein for wild-type and mutant shifts, parallel reactions were set up so that wild-type and mutant proteins were generated in the presence of either radioactive 35S-methionine, or non-radioactive methionine. Next, radioactive reactions were analyzed by SDS-PAGE and the radioactive bands corresponding to full-length MEF2 protein were quantified using the Cyclone Packard Phosphorimager. Based upon the efficiency of MEF2 protein generation as determined in the

radioactive reactions, the volume of non-radioactive MEF2 lysate was correspondingly adjusted in the gel shift reactions to ensure that equal amounts of protein were used. Differences in the volumes used of programmed lysate were corrected by addition of unprogrammed lysate.

Electrophoretic mobility shift assays were set up according to standard methods (Sambrook *et al.,*1998). Reactions were incubated in water baths at either 18ºC or 29ºC, and the gels were run at a constant temperature of 18ºC or 29ºC by circulating water of the appropriate temperature through the electrophoresis rig. After running, gels were dried and radioactive bands were quantified using the Packard Cyclone Phosphorimager. Data shown represent an average of at least two reactions for each mutant compared to wild-type at each temperature.

Cell culture and transfections

For a wild-type expression plasmid we used pPac-Pl-Mef2 (KELLY TANAKA *et al*., 2008). Mutant expression plasmids were generated by sub-cloning the mutant cDNAs into pPac-Pl. The reporter plasmid used was the -593/+2 *Actin 57B-lacZ* construct described by Kelly *et al.* (2002). Transfections and analyses were as described in Kelly Tanaka *et al.* (2008). All experiments were carried out in triplicate, and the average fold activation for each experiment is reported.

Results

Temperature dependent phenotypes of *30-5* and *44-5* alleles

In order to define the basis of the *30-5/44-5* temperature-sensitivity, we first studied the extent of embryonic muscle differentiation in homozygotes for each mutant *Mef2* allele. These studies were performed at 18ºC and 29ºC, and were compared to both wild-type and the heteroallelic combination that we had previously described (Baker *et al.*, 2005).

At stage 16, wild-type embryos stained for the accumulation of Tropomyosin showed extensive skeletal muscle differentiation at both temperatures (Figure 1 A, B); and the heteroallelic combination showed relatively normal patterning of skeletal muscles at the permissive temperature, but a strong hypomorphic phenotype at the restrictive temperature (Figure 1C, D).

For *30-5* homozygotes, muscle differentiation appeared relatively normal at both temperatures (Figure 1 E, F), suggesting that this allele on its own has significant wild-type function, and does not show a strong temperature dependence in its activity. Given the relatively normal pattern of muscles in *30-5* animals, we determined if they showed survival to adulthood at either 18ºC or 29ºC. We observed that *30-5* homozygotes never showed viability to the adult stage, although this might arise from second-site mutations that might have accumulated on the *30-5* chromosome. To address this possibility, we tested the viability of mutants heteroallelic for *30-5* and the *Mef2* null allele *P544* (Lilly *et al*., 1995). We found in this instance that 13% of *30-5/P544* mutants survived to adulthood at the permissive temperature (12 observed out of 89 expected) but none survived at the restrictive temperature (0 observed out of 238 expected). We conclude that the *30-5* allele is a mild hypomorphic mutation, which shows a

slight attenuation in its function when the temperature is raised. However, this very mild temperature-dependent effect does not account for the profound TS observed in *30-5/44-5* mutants. For the *44-5* allele, homozygotes showed severe defects in muscle development at both temperatures (Figure 1 G, H). Again, there did not appear to be a strong effect upon the severity of the embryonic phenotype caused by raising the mutants at different temperatures.

Overall we concluded that these two alleles, while individually not strongly temperature-dependent in their activity, somehow interacted to generate a temperature-dependent effect in heteroallelic combinations. It is interesting to note that at the permissive temperature the *30-5/44-5* combination appears phenotypically similar to the *30-5* homozygotes, whereas at the restrictive temperature the heteroallelic combination appears similar to the *44-5* mutants. Thus we reasoned that the *44-5* allele has an antimorphic effect (*i.e.*, it interferes with *30-5* function), but only at the restrictive temperature, and only in combination with *30-5* mutants.

Sequence analysis of *30-5*

Our next goal was to understand the molecular mechanism underlying this temperature sensitive effect. We isolated RNA from *30-5/+* heterozygotes and carried out an RT-PCR, amplifying *Mef2* cDNA using primers flanking the coding region. We cloned and sequenced the resulting products and discovered that the *30-5* gene contains a G to A point mutation in the MADS box (Figure 2A) which results in an alanine to threonine conversion at amino acid 32 of the predicted protein (Figure 2C). The same mutation was observed in clones of genomic DNA

isolated from *30-5/+* heterozygotes. Given that *30-5* homozygotes develop a fairly normal pattern of skeletal muscles, we hypothesized that *30-5* mutants are capable of forming MEF2 dimers and activating downstream target genes, albeit with reduced efficiency.

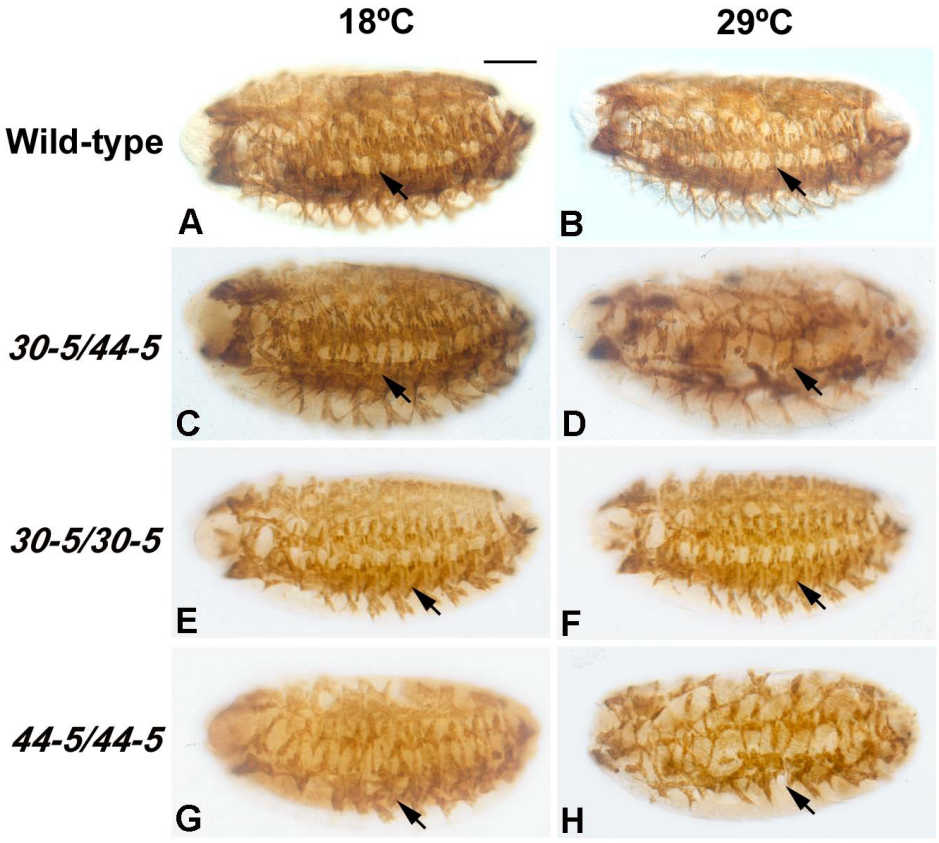


Figure 1
Skeletal muscle development in wild type and mutant *Mef2* alleles. Stage 16 embryos were stained with anti-Tropomyosin (A, B, E-H) or anti-Myosin Heavy-chain (C, D) to visualize skeletal muscle patterning and differentiation. Embryos were raised at the permissive temperature (18^{o}C) and the restrictive temperature (29^{o}C) to compare phenotypes at the two temperatures. (A-B) Wild type embryos develop normally at both the permissive and restrictive temperatures. In normal development, the three lateral transverse muscles 1-3 (LT1-3) are present in each segment (arrows). (C) The heteroallelic combination of the *30-5* allele and the *44-5* allele is phenotypically similar to wild type at 18^{o}C. Many segments still contain LT1-3 (arrow). (D) The *30-5/44-5* mutant is severely hypomorphic at 29^{o}C. (E-F) The *30-5* homozygous mutants closely resemble wild type at both temperatures, but are not viable to adulthood. (G-E) The *44-5* homozygous mutants are severely hypomorphic at both temperatures, and LT muscle numbers are greatly reduced (arrow). Note that the *44-5* homozygotes most closely resemble the *30-5/44-5* combination raised at 29ºC.

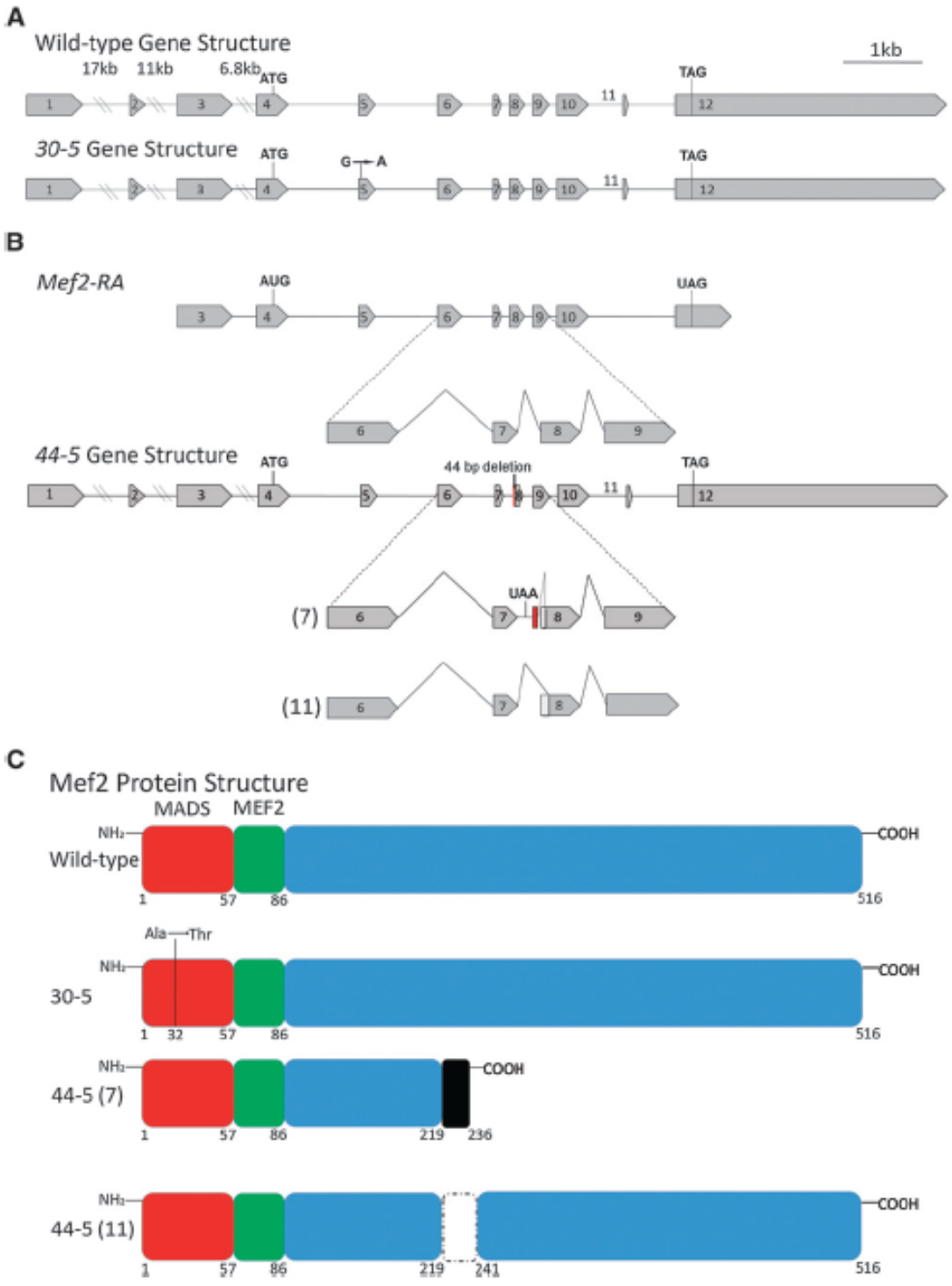

A
Wild-type Gene Structure
17kb
11kb
6.8kb
ATG
TAG
1kb
30-5 Gene Structure
ATG
G→A
TAG
B
Mef2-RA
AUG
UAG
44-5 Gene Structure
ATG
44 bp deletion
TAG
(7)
UAA
(11)
C
Mef2 Protein Structure
MADS
MEF2
Wild-type
NH₂
COOH
1
57
86
516
Ala→Thr
30-5
32
44-5 (7)
219
236
44-5 (11)
241

Figure 2
Comparison of gene structure, splicing patterns and protein structure in *Mef2* wild-type and mutant alleles. (A) Gene structure of wild-type and *30-5. 30-5* contains a G to A point mutation in exon 5. (B) The wild-type, *Mef2-RA* isoform and a blow-up of its splicing pattern between exons 6 through 9 compared to the 44-5 gene structure and its two alternative splicing patterns. The gene contains an additional 9 bp of nucleotides following intron 7 and 44 bp of nucleotides deleted from exon 8. The *44-5 (7)* isoform retains intron 7 and splices into the 45th bp of exon 8. The *44-5 (11)* isoform splices into the 52nd bp of exon 8. (C) Wild-type MEF2 contains a 57 amino acid MADS DNA-binding domain, a 29 amino acid MEF2 DNA-binding and dimerization domain. 30-5 contains an alanine to threonine substitution at amino acid 32. The black shaded box of 44-5 (7) is translated intronic sequence and contains a premature stop codon at amino acid 236. 44-5 (11) is missing amino acids 219 through 241.

The 30-5 protein demonstrates decreased DNA binding ability

To test this hypothesis, we performed an electrophoretic mobility shift assay (EMSA) using a radioactively labeled *Actin 57B* MEF2 site as a probe. The reaction was carried out at both 18ºC and 29ºC, and equal amounts of wild-type and 30-5 protein were used for binding reactions (see Materials and Methods for details). The results of the EMSA demonstrated that the 30-5 isoform consistently showed a reduced ability to bind to the probe relative to wild-type MEF2 protein. However, this reduction in DNA binding was not affected by temperature (Figure 3A, B). In cotransfection assays 30-5 was able to activate *Act57B* in cell culture at both temperatures equally well (data not shown). These experiments, along with the phenotypic characterization of terminal differentiation in *30-5* mutants, support the idea that the 30-5 isoform in isolation is capable of activating the muscle differentiation pathway. However, due to its decreased DNA binding ability, the pathway is not activated robustly enough in vivo for the animal to survive to adulthood. Furthermore, the temperature sensitivity must arise from an interaction of the 30-5 isoform with the 44-5 isoform.

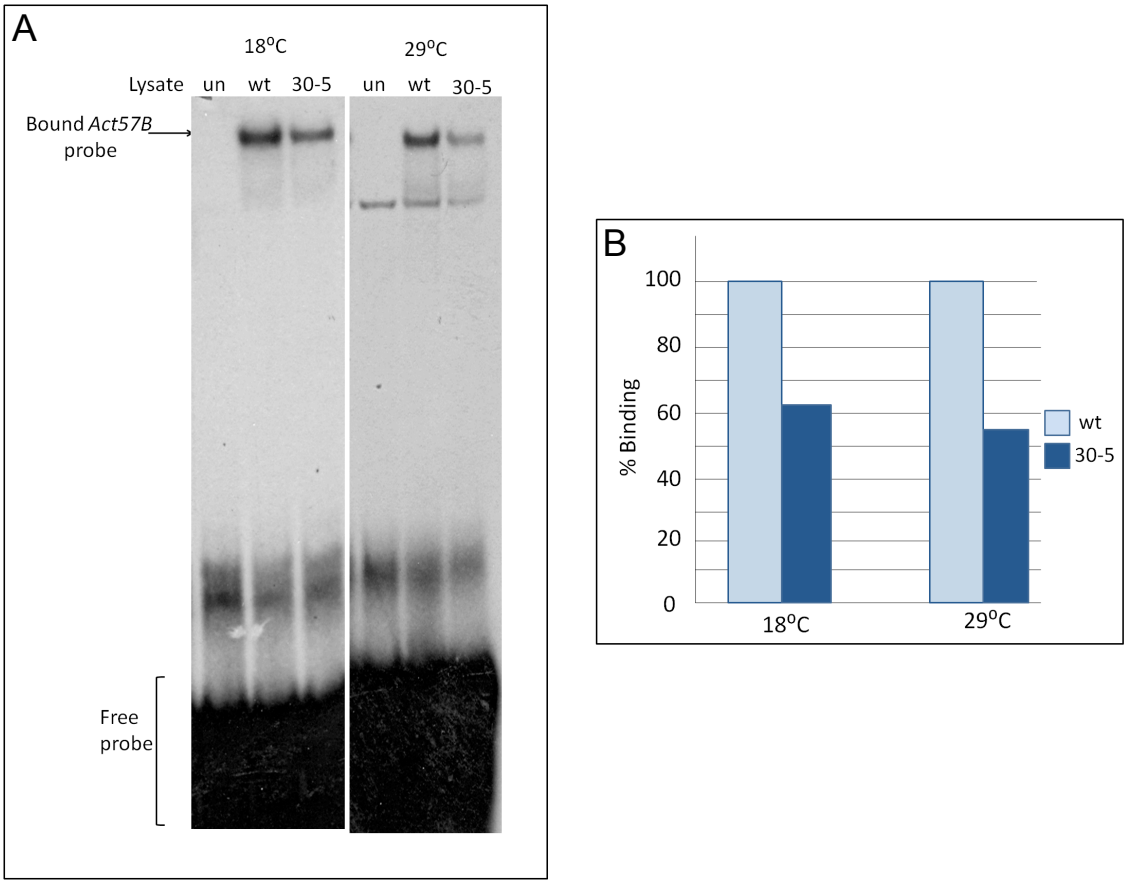


Figure 3
DNA binding ability of the 30-5 isoform at 18ºC and 29ºC. (A) 30-5 and radioactively labeled probe from the enhancer region of *Act57B* were able to form complexes that moved considerably slower than free probe through a polyacrylamide gel; however the band intensity was significantly less intense when compared to wild-type regardless of temperature. An additional non-specific band is seen at 29ºC in the unprogrammed lysate as well as the experimental lanes. (B) Percent binding of 30-5 compared to wild-type (100%) quantified by band intensity (see materials and methods).

Sequence analysis and functional characterization of 44-5 protein isoforms

To gain further insight into this mechanism we next isolated RNA from *44-5/+* heterozygotes and carried out an RT-PCR with primers flanking the *Mef2* coding region. We then cloned and sequenced the resulting *Mef2* product as performed previously for 30-5. We found two distinct mutant transcripts which we refer to as *44-5 (7) and 44-5 (11)* (Figure 2 B, C), both of which lacked the

normal downstream splice acceptor site of exon 8. The 44-5 (7) transcript was missing the first 44 bases of exon 8 and contained extra nucleotides that did not match the *Mef2* coding sequence, but did match sequence from intron 7. The 44-5 (11) transcript lacked the first 51 bases of exon 8. We then analyzed *Mef2* genomic clones and discovered that in the *44-5* allele there was a 44 bp deletion which removed the exon 8 splice acceptor site and a portion of exon 8 coding sequence; all but the last 3bp of intron 7 were retained.

The effects of these altered transcripts upon the encoded MEF2 protein are potentially profound (Figure 2C). The *44-5* (11) splice variation results in a 22 amino acid internal deletion from the final protein structure; and the 44-5 (7) splice variant encodes a polypeptide with a large C-terminal deletion. Both of these isoforms contain mutations in a region that has not been fully characterized but which may affect the activation domain of MEF2. We therefore predicted that that 44-5 protein isoforms would be able to bind to DNA but we were uncertain if the binding ability of the severely truncated 44-5 (7) isoform would be reduced at the restrictive temperature. We found that each of the 44-5 (7) and 44-5 (11) isoforms bound to the *Act57B* MEF2 site at 18ºC and 29ºC extremely well (Figure 4A, B). In fact, the 44-5 (11) isoform consistently bound to DNA even more effectively that wild-type MEF2.

Since DNA binding ability of the 44-5 isoforms was not reduced, but terminal differentiation of *44-5* homozygotes was severely affected at both temperatures, we concluded that the 44-5 isoforms must not be able to activate downstream target genes. We tested this using co-transfection assays (Figure

4C), and found that the ability of the 44-5 (7) isoform to activate downstream targets was severely reduced at both temperatures. However, the 44-5 (11) isoform was capable of activating transcription as effectively as wild-type MEF2 at both temperatures.

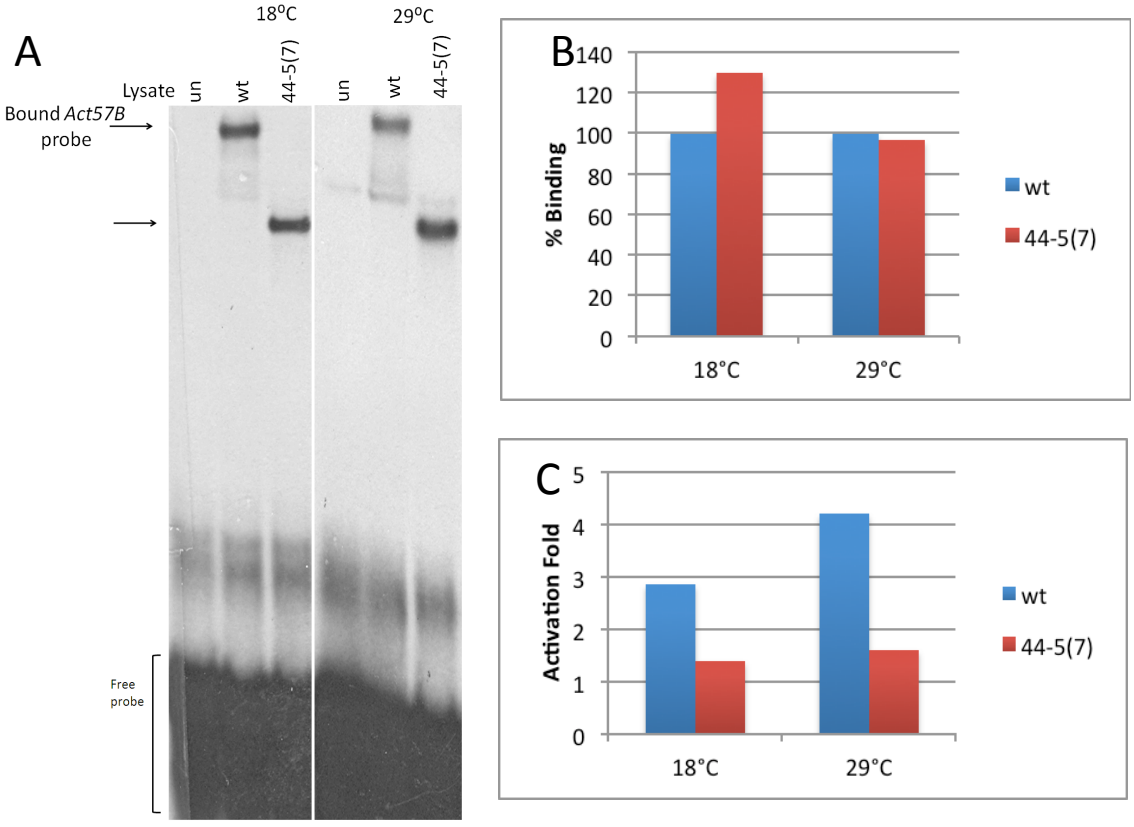


Figure 4
DNA binding and activation ability of the 44-5 isoforms. (A) 44-5 (7)/(11) and radioactively labeled probe from the enhancer region of *Act57B* were able to form complexes that moved considerably slower than free probe through a polyacrylamide gel; the 44-5 (7) isoform complex is considerably smaller, as expected since the protein is severely truncated. (B) Percent binding of 44-5 (7) and (11) compared to wild-type (100%) quantified by band intensity. Both isoforms were able to bind to DNA as well as, if not better than, wild-type at both temperatures. (C) Activation of an *Act57B* enhancer fused to a lac-z reporter gene by either wild-type MEF2, 44-5 (7) or (11) isoforms in cell culture. Wild-type MEF2's ability to activate a downstream target gene increases at 29ºC. 44-5 (7)'s ability to activate is severely retarded at both temperatures. 44-5(11) is able to activate better than wild-type.

The observation that one of the 44-5 mutant isoforms was capable of activating a canonical MEF2 target gene was inconsistent with our demonstration

of a severe muscle phenotype in *44-5* homozygotes. We therefore postulated that there might be differential expression of the two mutant isoforms. To test this, we performed another RT-PCR on *44-5/+* heterozygotes, using a primer pair which would generate characteristic products for each of the mutant and wild-type isoforms. This RT-PCR was carried out on mRNA isolated from animals raised at either 18ºC or 29ºC. We found that, at both temperatures, the *44-5 (7)* isoform was by far the predominant of the two mutant transcripts (Figure 5). This result supported the severity of the phenotypic data in *44-5* homozygotes at both temperatures, since the predominant transcript *44-5 (7)* binds downstream targets very well but is incapable of transcriptional activation.

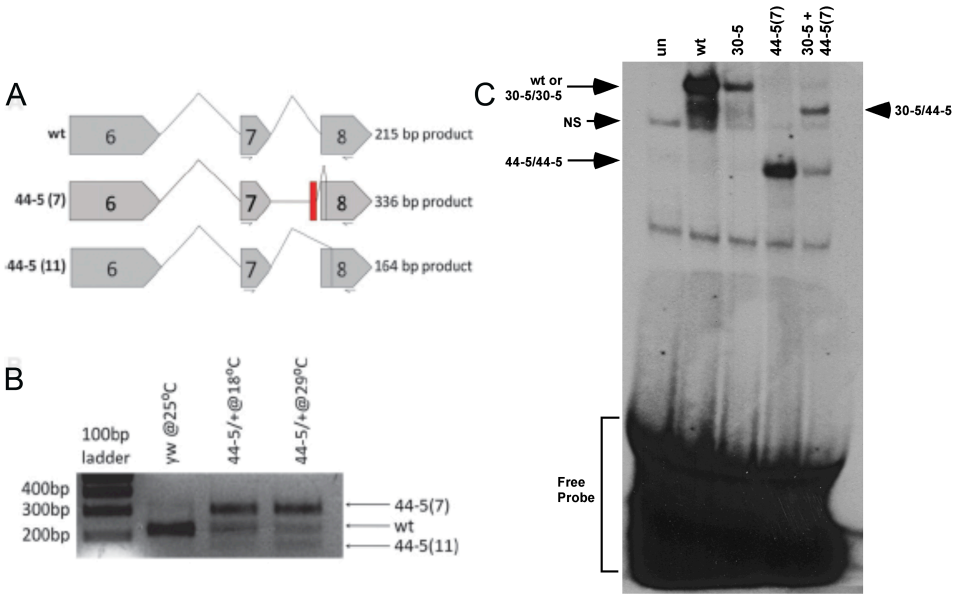


Figure 5
Quantification of MEF2 transcript isoforms at 18ºC and 29ºC. (A) Diagram of wild-type and the two 44-5 isoforms. 44-5 (7) retains intron sequence and is therefore larger than wild-type. 44-5 (11) is missing 51bp of exon 8 and is therefore smaller than wild-type (B) RT-PCR of wild-type and 44-5/+ heterozygotes. 44-5 (7) is the dominant transcript in heterozygotes at both temperatures. (C) Electromobility shift assay with a radioactively labeled MEF2 target. The first lane contains only unprogrammed lysate. The second land has wt MEF2 protein added. A robust band is seen. In lane 3, the 30-5 mutant protein binding is significantly reduced. The truncated 44-5(7) mutant forms a smaller shift but the band intensity is robust. In the last lane a combination of the two mutant proteins results in an intermediate shift from the heteroallelic dimer and less robust shifts from the two homodimers.

A model to explain temperature dependency of *30-5/44-5* mutants

Through our analysis of the *30-5/44-5* heteroallelic phenotypes, we have observed that the TS effect shows a phenotype that is closer to the homozygous *30-5* mutant at the permissive temperature and closer to the *44-5* mutant at the restrictive temperature. Having established that the 30-5 isoform has a defect in DNA binding, and that the predominant 44-5 isoform can bind DNA but fails to activate transcription, we proposed the following model to explain the differential survival of *30-5/44-5* heteroallelic animals when raised at different temperatures.

The 30-5/44-5 heterodimer can bind DNA and activate downstream genes, due to the interaction of the 44-5 DNA binding ability and the 30-5 activation ability. This positive interaction occurs successfully at the permissive temperature, and is sufficient to rescue the lethality of *30-5* homozygotes. At the restrictive temperature, where molecular interactions are more restrictive to low-affinity binding than at the lower temperature, the 44-5 homodimers out-compete the 30-5/44-5 heterodimers for DNA binding. The superior binding of 44-5 to DNA at the restrictive temperature results in a more severely mutant phenotype, since 44-5 is incapable of initiating the myogenic program.

Analyses of additional *Mef2* alleles support the model

To test this model we sought to determine if other *Mef2* alleles with mutations in defined regions of the polypeptide would show a temperature-dependent rescue of the *30-5* mutant phenotype. We predicted that an allele with a mutation affecting the DNA binding domain would not be TS when combined with *30-5*, whereas alleles encoding proteins with normal DNA binding domains

and mutated C-terminal regions would demonstrate a TS effect similar to that observed with the *30-5/44-5* crosses.

For our analysis, we chose three *Mef2* alleles isolated by GOLDSTEIN et al. (2001) which demonstrated severely mutant phenotypes as homozygotes. The *26-7* homozygotes were functionally null for muscle development (Figure 6A), whereas the *26-49* and *66-5* alleles showed severely hypomorphic phenotypes (Figure 6B, C). Due to the severity of the mutations, these alleles were good candidates for a test of TS amelioration of the phenotype when combined with the *30-5* allele.

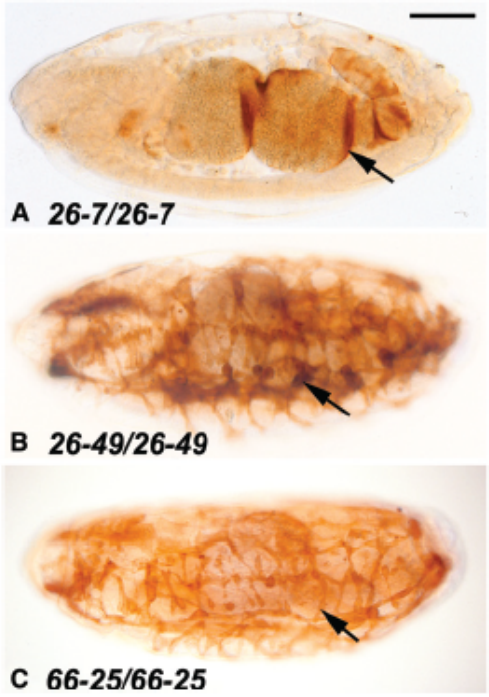


Figure 6
Skeletal muscle development in three new *Mef2* mutant alleles. (A) The *26-7* homozygous mutant is a null phenotype—no skeletal muscle is present. (B-C) The *26-49* and *66-25* homozygous mutant phenotypes are severely hypomorphic.

We performed sequence analyses of the three alleles, to determine if the mutation affected either the DNA binding and dimerization region of the protein,

or a more C-terminal region. We first analyzed the coding region of each allele through cDNA sequence obtained from isolated RNA, and then confirmed these

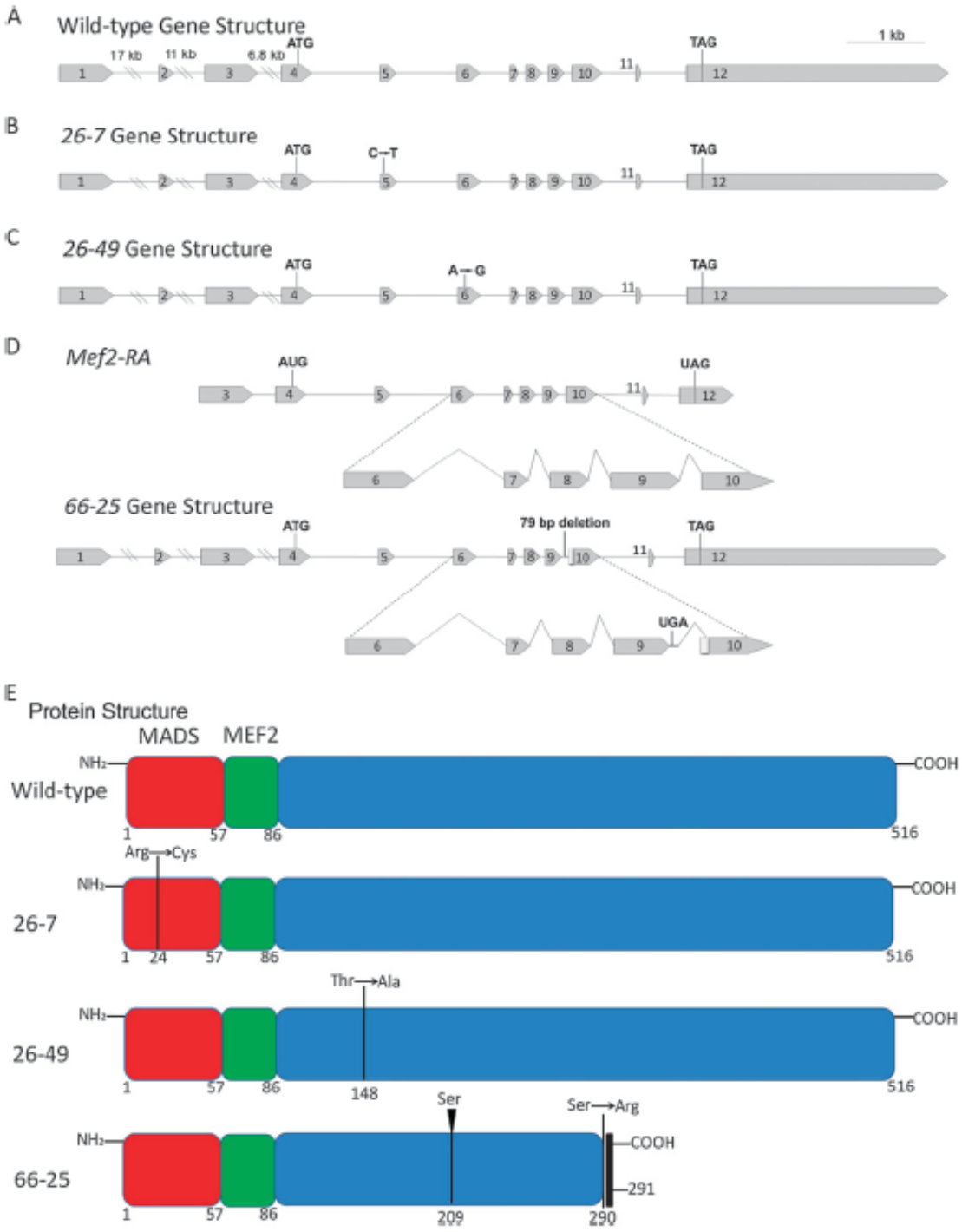


Figure 7
Analysis of gene and predicted protein structure for three new *Mef2* mutant alleles. (A) Wild type gene structure contains 12 exons. The ATG start codon is in the 4th exon and the TAG stop codon is in the 12th exon. (B) The *26-7* allele has a point mutation from cytosine to thymine in exon 5, encoding an amino acid change from arginine to cysteine at amino acid 24. (C) The *26-49* allele has a point mutation from adenine to guanine in exon 6, encoding an amino acid change from threonine to alanine at amino acid 148. (D) The *66-25* mutation is a 79 base pair deletion after exon 9, which deletes the splice acceptor site and causes the gene to read into the intron and splice into the 10th exon downstream to where it would normally splice in wild type. The result is to encode a stop codon one amino acid after the 9th exon. The normal splicing pattern is shown in the *Mef2-RA* structure. (E) Predicted protein structure for wild-type and three new *Mef2* mutant alleles as indicated: the *26-7* mutation encodes an amino acid change from arginine to cysteine at amino acid 24, which occurs in the DNA-binding MADS domain; the *26-49* mutation encodes an amino acid change from threonine to alanine at amino acid 148, which occurs in the downstream C-terminal region; the *66-25* mutation encodes an insertion of serine at amino acid 209 and a premature stop arising from read-through of intron 9.

sequence alterations through analyses of genomic DNA sequence (see Materials and Methods for details). For two of the alleles, the mutation was a point mutation in the coding region: when compared to wild-type (Figure 7A) the *26-7* mutation affected the MADS box, encoding a missense mutation at amino acid 24 from Arg to Cys, CGC to UGC (Figure 7B, E); interestingly, this is the same point change as was observed for the 26-6 allele, also isolated by GOLDSTEIN et al. (2001) and analyzed by NGUYEN *et al.* (2002); the *26-49* mutation was outside of the MADS and MEF2 domains, encoding a missense mutation at amino acid 148 from Thr to Ala, ACA to ACG (Figure 7C, E). For the *66-25* allele, our analysis of the cDNA revealed two alterations: there awas an insertion of a Serine codonat amino acid 209 of the coding sequence; more severely, there was an insertion of 17bp, introducing a UGA stop codon, after the ninth exon (Figure 7D). We then analyzed genomic DNA from *66-25* and determined that the mutation was a 79-bp deletion of the ninth intron, which removes the

splice acceptor AG di-nucleotide of exon 10. The result of this mutation is a truncation of the encoded protein (Figure 7E). This mutation was also outside of the MADS and MEF2 domains, encoding a truncated protein.

To determine if a temperature sensitive effect was observed when we crossed the three new alleles with *30-5*, we used immunohistochemistry to assess Tropomyosin accumulation in heteroallelic embryos raised at permissive or restrictive temperatures. For *30-5/26-7* there was a slight amelioration of the *26-7* homozygous phenotype, but the effect was not TS: at both temperatures the mutants still showed severely affected muscle development (Figure 8A, B). By

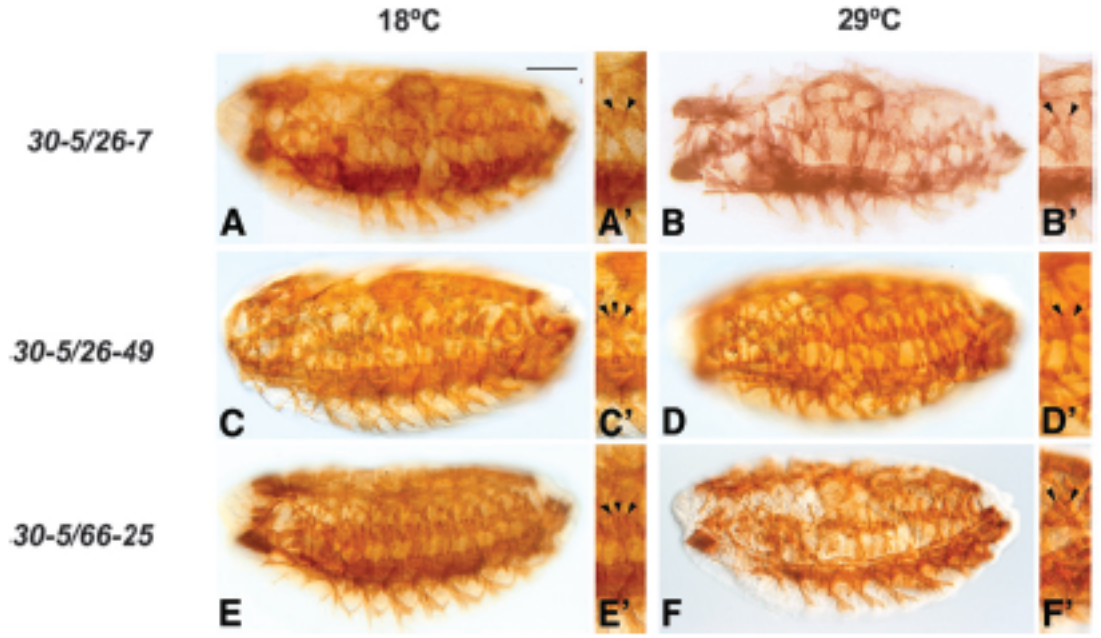


Figure 8
Heteroallelic crosses of new *Mef2* alleles with the *30-5* allele at 18°C and 29°C. (A-B) The *26-7/30-5* crosses were severely mutant at both temperatures. The showed very low survivorship at 18°C (9%) and no survivorship at 29°C. (A'-B') At both temperatures many segments contained only two lateral transverse (LT) muscles, as opposed to three LT muscles in wild type (arrowheads). (C-D) The *26-49/30-5* crosses were temperature sensitive: at 18°C the phenotype was closer to wild-type (C), with three LT muscles per segment (C'), and survivorship was 67%. At 29°C the phenotype was more mutant (D), with some segments having only two LT muscles (D'), and survivorship was only 14%. (E-F) The *66-25/30-5* crosses were also temperature sensitive, with similar phenotypes to those of the *26-49/30-5* crosses. At 18°C the phenotype is closer to wild type (E), with three LT muscles per segment (E') and survivorship of 49%. At 29°C the phenotype was more mutant (F), with some segments having only two LT muscles (F') and survivorship of only 31%.

contrast there was a modest temperature-dependent rescue of both of the C-terminal region mutants, *26-49* and *66-25*, when combined with *30-5* (Figure 8C mutants, yet the phenotype was more severe at 29^{o}C (Figure 8C-F). F). In both cases, significant muscle fibers were observed in heteroallelic

To document this temperature-sensitivity in a more quantitative manner, we studied the survival to adulthood of the heteroallelic mutant combinations studied above. The original TS combination of *30-5/44-5* showed a dramatic difference in viability at the two temperatures, with 58% of embryos surviving to adulthood at the permissive temperature and 0% surviving at the restrictive temperature (Baker *et al*., 2005). For the *30-5/26-7* combination of two MADS domain mutant alleles, the survival rate was 9% (31 observed out of 132 expected) at the permissive temperature and 0% (0 observed, 170 expected) at the restrictive temperature. When *30-5* was combined with the C-terminal region mutants the viability was more clearly TS: for the *30-5/26-49* cross, the survival rate was 67% (134 observed, 200 expected) at the permissive temperature and 14% (33 observed, 239 expected) at the restrictive temperature; for the *30-5/66-25* cross, the survival rates were 49% (93 observed, 189 expected) at the permissive temperature and 31% (61 observed, 196 expected) at the restrictive temperature.

Taken together, these findings further support the model for temperature sensitivity: heteroallelic combination of one *Mef2* mutant showing weakened DNA binding ability with a second *Mef2* mutant showing weakened activation ability

can generate a TS effect. Nevertheless, we note that the severity of the TS effect can still vary between different alleles in the same class.

Discussion

MEF2 is a widely expressed, multi-functional dimeric protein that works in collaboration with many other factors to activate and repress a variety of developmental processes at multiple stages of development (reviewed in Black and Cripps, 2009; Potthoff and Olson 2007). We have previously exploited the temperature-sensitivity of the *30-5/44-5* combination to elucidate a role for MEF2 in adult myogenesis, where we showed that a strong reduction in MEF2 function could nevertheless support several aspects of adult muscle development (Baker et al. 2005). In this paper we have characterized the molecular mechanism behind this TS combination protein, in an effort to elucidate more information about the dynamics of temperature sensitivity and the mechanisms of MEF2 function.

Current models of temperature sensitivity take into account mutations that affect the proper formation or folding of a protein when synthesized at a restrictive temperature, and mutations that destabilize a protein when shifted to a restrictive temperature (Sadler and Novick, 1964). Mutations that are capable of destabilizing a protein can be found in DNA binding domains, protein-protein interaction domains or internal hydrophobic regions of a protein (Sundberg and Davis 1997, Smith *et al.*1980, Gordon and King 1993, Edgar and Lielausis, 1963). In addition, the temperature-sensitivity of homodimeric proteins has received attention. In several instances, intragenic complementation has been

observed for dimeric or multimeric proteins, and the classical literature has found that in several cases such complementation is TS (Fincham 1966). This relationship was underlined by Sundberg and Davis (1997), who showed that mutations affecting different functional regions of the *S. cerevisiae* spindle protein could show effective TS intragenic complementation.

The *30-5/44-5* TS findings that we describe fit the latter model effectively. The combination of the *30-5* DNA-binding mutant with *44-5* or with either one of two additional C-terminal mutants demonstrated a TS effect, whereas combination of *30-5* with another allele affecting the DNA binding domain did not show effective rescue and temperature-sensitivity. We note that the severity of the various TS effects that we have observed is still the most striking with *30-5/44-5*. This might be attributed to a less severe disruption of the activation domain in the C-terminal mutants *26-49* or *66-25*. Indeed, the 26-49 isoform contains only a single amino acid substitution in the C-terminus, but still produces a full-length protein. The 66-25 mutant protein isoform is truncated like that of 44-5, however it encodes approximately 71 more amino acids in the C-terminal region. As homozygotes, these mutants are severe, lethal hypomorphs, but in combination they complement well and demonstrate a TS effect.

Our findings for the temperature-sensitivity of *Mef2* allelic combinations could be applied to the analysis of other transcription factors. A combination of alleles for the Drosophila *twist* gene also shows TS in the heteroallelic arrangement (Thisse et al. 1997). Neither of these mutations affect the DNA

binding domain (Baylies and Bate, 1997), however they might nevertheless impact regions of the protein with distinct functions.

Regarding MEF2 functional domains, it is interesting to note the severity of the phenotype in *26-49* homozygotes. Its single amino acid substitution of Threonine to Alanine at amino acid 148 in the C-terminus suggests a critical role for Thr148 in normal MEF2 activity. One possible function for this residue is that it is a target of phosphorylation during development, and there is significant evidence from tissue culture studies that phosphorylation of mammalian MEF2 proteins is important to its function (Molkentin *et al.* 1996b, McDermott *et al*. 2002, Gulick and Zhu 2004). However this aspect of MEF2 biology has yet to be fully addressed using an in vivo system, and the *26-49* allele might be critical in this regard.

An additional regulatory domain appears to lie between amino acids 219-241 in the C-terminal region of MEF2. This region has a potentially inhibitory function based upon the increased transcriptional activation activity of the 44-5 (11) isoform that lacks these amino acids. Previous work using murine MEF2s has identified a repression domain located in exon 9 (Gulick and Zhu 2004). Alternative splice variants lacking this domain have an increased ability to activate downstream MEF2 targets in cell culture as was seen in the cell culture assay for 44-5 (11). While the region identified by Gulick and Zhu (2004) probably does not correspond to amino acids 219-241, these finding nevertheless confirm that there is much still to be learned regarding the function of the MEF2 C-terminal region.

ACKNOWLEDGEMENTS

I would like to thank Phill Baker for the many crosses of *Mef2* mutants that led to the identification of the temperature sensitive phenotype. I would like to thank Melanie Adams for the embryo stains as well as the characterization of the 26-7. 26-49 and 66-25 alleles. I am very grateful to the individuals who supplied reagents essential to this work: Elliot Goldstein and Daniel Kiehart. This work was supported by a grant from the NIH (GM61738) to RMC. MMA was supported by a MARC fellowship from the NIH (5T34GM008751). I acknowledge technical support from the Department of Biology's Molecular Biology Facility, supported by NIH grant number 1P20RR18754 from the Institute Development Award (IDeA) Program of the National Center for Research Resources.

CHAPTER 2

Tinman and Pannier activate and collaborate with MEF2 to promote heart cell fate

Abstract

Transcription of the MADS domain transcription factor MEF2 (Myocyte Enhancer Factor 2) is regulated by numerous and overlapping enhancers which tightly control its expression throughout the mesoderm. To understand how Mef2 expression is controlled in the heart, we identified a late stage *Mef2* cardiac enhancer that is active in all heart cells beginning at stage 14 of embryonic development. This enhancer is regulated by the NK-homeobox transcription factor Tinman and the GATA transcription factor Pannier through both direct and indirect interactions with the enhancer. Since Tinman, Pannier and MEF2 are evolutionarily conserved from *Drosophila* to vertebrates (Cripps and Olson, 2002) and since their vertebrate homologs can convert mouse fibroblast cells to cardiomyocytes in different activator cocktails (Ieda et al., 2010 and Ifkovits et al., 2014), we tested whether their overexpression in vivo could ectopically activate known cardiac marker genes. We found that mesodermal overexpression of Tinman and Pannier resulted in approximately 20% of embryos with ectopic *Hand* and *Sulphonylurea receptor (Sur)* expression, and 100% of embryos expressed ectopic *Wingblister* expression in the posterior of the heart. By simply adding MEF2 alongside Tinman and Pannier, a dramatic expansion of *Hand* and the *Sulfonylurea receptor* was seen in almost 100% of the embryos stained. These results demonstrate the ability to initiate ectopic cardiac fate in vivo, with the combination of only three members of the conserved *Drosophila* cardiac transcription network and provides an opportunity for this genetic model system to be used to dissect the mechanisms of cardiac specification.

Introduction

Many developmentally important genes contain multiple enhancers with overlapping activity. Duplicate enhancers were first found when looking for targets of Dorsal in the *Drosophila* embryo. The genes *brinker* and *sog,* for which Dorsal-dependent enhancers had already been described, were found to have secondary enhancers in distant locations, one of which was in the intron of another gene (Hong et al. 2008). In another example, duplicate enhancers in the gene *snail* were shown to work equally well alone when they were the only regulatory element in a *snail* mutant background. At elevated temperatures however, there was a reduction in snail expression and disruptions in gastrulation (Perry et al., 2010) suggesting that the redundancy provides developmental insurance for the embryo during stressful environmental situations.

Given the importance of *Mef2* expression to heart muscle differentiation, it might be predicted that *Mef2* contains multiple enhancers active in the developing cardiac mesoderm. In *Drosophila*, the heart is comprised of two distinct cells types that can be characterized by their mutually exclusive expression of the NK-homeobomain transcription factor Tinman or the orphan steroid hormone receptor Seven-up. Whereas Tin and Svp cell types perform distinct function in the mature organ, the cells are still contractile, and express similar groups of contractile protein isoforms (Molina and Cripps, 2001; Ponzielli et al., 2002; Zhang and Bernstein, 2001). Accordingly, both cardiac cell types express *Mef2*, which is required for muscle protein gene expression in all contractile heart cells (Lilly et al., 1995, Bour et al., 1995).

Three enhancers have been identified that regulate *Mef2* transcription in the heart. The first enhancer is active early, at stage 11 of embryogenesis, and drives expression in Tinman-expressing cells (Gajewski et al. 1997, Cripps et al. 1999). This enhancer is activated by Tinman and the zinc finger transcription factor Pannier, which maintain enhancer activity through the end of embryogenesis. A second enhancer drives *Mef2* expression in Seven-up expressing cells (Gajewski et al., 2000), and a third enhancer becomes active in the somatic mesoderm and both cardiac cell types at stage 14 of embryogenesis (Nguyen and Xu, 1998). A portion of this enhancer responsible for somatic mesodermal expression was found to be regulated by a member of the Gli superfamily of transcription factors called Lame duck (Duan et al., 2001).

MEF2 is an essential activator of genes required for differentiation (Bour et al., 1995, Lily et al., 1995, Ranganayakulu et al., 1995). Given the importance of MEF2 expression for normal formation of such a wide range of tissues and structures, it is not surprising that its regulation is so complex, where expression of Mef2 in mesodermal tissues is regulated by an array of cis-regulatory elements (Schulz et al 1997; Nguyen and Xu, 1998).

In this chapter, we identify the transcription factors that regulate the cardiac expression of a third *Mef2* cardiac enhancer. The regulation of this enhancer is unique in that it is active in Tin plus Svp cell types of the heart whereas the previous enhancers were active either in the Tinman expressing cardiac cells or the Seven-up expressing cardiac cells. Our data demonstrate that the enhancer is activated directly by Tinman binding to an essential site, and

indirectly by Pannier, presumably through interacting with Tinman. In addition, we demonstrate that Tinman, Pannier and MEF2 can collaborate to activate markers of cardiac differentiation in the mesoderm.

Materials and Methods

Generation of promoter-*lacZ* constructs

The 345bp enhancer was PCR-amplified with the forward primer (5-'CCTCTCTTTTGGCAGAAAGTCG-3') and reverse primer (5'-AAACTCATCTCCACGCCACTGC-3'). It was cloned into the vector pLacZattb and injected into flies using phiC31 integrase at the landing site 86Fb (Bischof et al., 2007). Mutation of the Tinman and Lameduck consensus sequences was carried out by PCR site directed mutagenesis (Horton, 1993). Primers for each of the constructs were designed to contain an *EcoRI* site in place of the consensus binding sequences. Primary PCR amplification products using the original forward primer/reverse mutated primer and the original reverse primer/mutated forward primer were used as templates in a secondary PCR amplification of the full-length enhancer. The forward Tinman mutant primer was (5'GAGTCGAAAT*GAATTC*GCTGAACTGAACTTC3') and the reverse was (5'GAAGTCAGTTCAGC*GAATTC*ATTTCGACTC3'). The forward Lameduck mutant primer was (5'TTTGAATGAGATTTATGAAA*GAATTC*AAAACATCATC'3) and the reverse was (5'- GATGATGTTTT*GAATTC*TTTCATAAATCTCATTCAAA-3'). *EcoRI* sites are in italics. Generation of transgenic flies carrying the mutated enhancers was carried out as previously described.

Immunohistochemistry

Embryos were collected and fixed according to Patel et al. (1987). We used a primary antibody against β-galactosidase from Promega (Madison, WI) at a concentration of 1:400 and a rabbit MEF2s antibody (Lilly et al. 1995) at a concentration of 1:1000. For non-fluorescent stains, we used the Vectastain Elite Kit (Vector Laboratories, Burlingame, CA). For fluorescent stains, we used Alexa 488 anti-mouse and 568 anti-rabbit secondary antibodies (Molecular Probes, Eugene, OR).

In situ hybridization

Embryos were prepared according to Lècuyer et al. (2008) until the hybridization step after which the Watakebe et al. (2010) protocol for labeling and hybridization was followed. The following primers were used to amplify portions of each transcript from yw embryo cDNA, which was subsequently cloned into the pGEM-T Easy vector (Promega A1360). Sequence orientation was determined by sequencing. Plasmids were then linearized at the 5' end of the transcript and the appropriate RNA polymerase (either T7 or SP6) was used to generate an anti-sense FITC RNA probe to each transcript according the protocol in Watakebe.

Transcript	Forward primer 5'-3'	Reverse primer 5'-3'
Hand	ATGTTTAAGAATTCCGTTGCC	CGTGCGGCCCTTGGTCG
Sur	CCGCCATTTCGTGTGTTTGT	GTGGTTGCCTCATAGTGCCT
wb	CCAAACGGCGTCTACAGGAT	GATGACTGGCGTGCTTTTCC

Electrophoretic mobility shift assay

Complementary DNA oligonucleotides were ordered from Sigma-Aldrich (St. Louis, MO) to generate double stranded DNA molecules with 5' GG overhangs. The oligos were radioactively labeled with 32P-dCTP (Perkin Elmer, Waltham, MA) using Klenow enzyme (New England Biolabs, Beverly, MA). The sequences tested were Tin1 5'-GG-GAGTCGAAATCACTTGAGCTGAACTGA-3' and Pnr1 5'GG-TTGCATAATTGATACCACCGCAGA-3'. Tinman protein was synthesized in vitro from the pBSK plasmid (Bodmer, 1990) using T3 polymerase in the Promega TNT Coupled Reticulocyte Lysate System (Promega, Madison, WI).

Cell culture co-transfection assay

Tinman and Pannier cDNAs were cloned into the pPacPl plasmid and the late stage enhancer was cloned into CHAB. Transfections were carried out with TransIT Transfection Reagent (Mirus, Madison, WI) according to the manufacturers directions. Experiments were carried out in triplicate and the average activation fold was calculated.

Fly stocks and crosses

The 69B-gal4 line and *UAS-pannier* lines were obtained from the Bloomington Stock Center. Ryan et al. (2007) generated the UAS-tinman line. The *UAS-tinman* and *UAS-pannier* transgenes are both on chromosome two. So, the *UAS-tinman;pannier* line was generated by recombination. We added our enhancer line to the third chromosome of this stock using standard genetic techniques.

Results

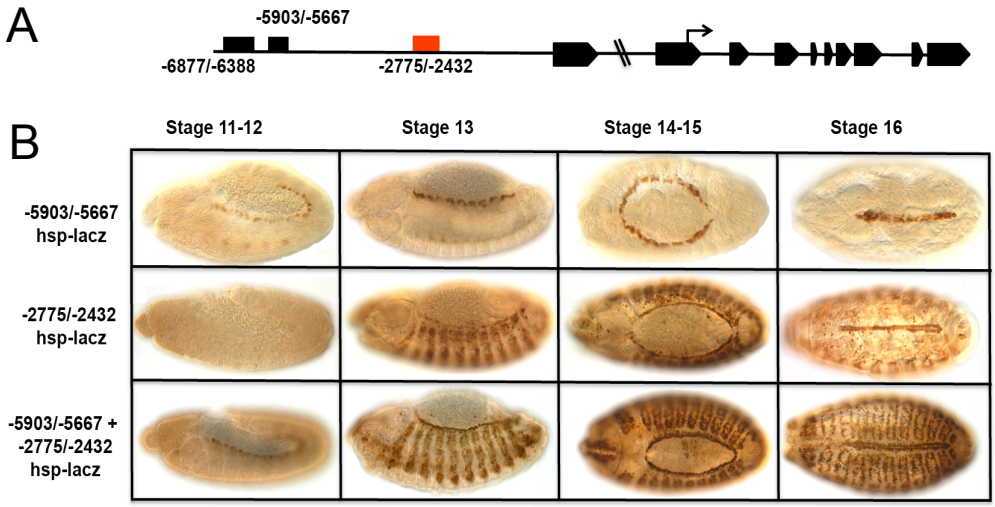
Mef2 contains a cardiac enhancer expressed in both cell types of the heart

We have worked to isolate the cardiac specific portion of the large late stage enhancer first identified by Nguyen and Xu (1998) which lies more proximal to the transcription start site of the *Mef2* gene than the previously identified cardiac enhancers (Figure 1A). We generated PCR fragments from this region, cloned them into a plasmid containing a *lacZ* reporter gene and generated transgenic flies that contained the Mef2-lacZ constructs. Embryos from these lines were stained with an antibody against β-Galactosidase to visualize activity of the enhancer. The smallest enhancer fragment with complete activity in both cardiac cell types lies at -2432/-2775. This 345 bp fragment also contained the 170bp mesodermal enhancer characterized by Duan et al (2001) and attempts to separate the cardiac and somatic enhancer activities resulted in loss of activity from both the cardiac cells and somatic mesoderm (data not shown). This observation indicates that certain enhancer sequences are utilized in both cardiac and skeletal muscle tissues. The enhancer becomes active at stage 14 and remains active until the end of embryogenesis (Figure 1B, row 2).

The -5903/-5667 enhancer becomes active early in development when the cardiac cells are just becoming specified at stage 11. Its expression becomes restricted to the Tinman expressing cells by stage 13 and remains active until the end of embryogenesis (Figure 1B, row 1). We were unable to reproduce the previously published Svp-cell specific enhancer activity of the -6877/-6388 sequence. However, when we fused our late stage enhancer to the early Tinman

specific enhancer, we observed that the fused enhancer completely recapitulated MEF2 expression in the heart from the earliest stage to the end of embryogenesis in both cell types (Figure 1B, row 3).

To gain insight into how expression of *Mef2* is regulated via this enhancer, we analyzed the sequence for known transcription factor binding sites. We observed consensus sequences for Tinman and Pannier (Figure 1C). As these two factors have already been shown to activate *Mef2* in the heart via the --5903/-5667 enhancer, we tested their ability to activate the enhancer in vitro, as well as tested their binding ability to the identified consensus sites.



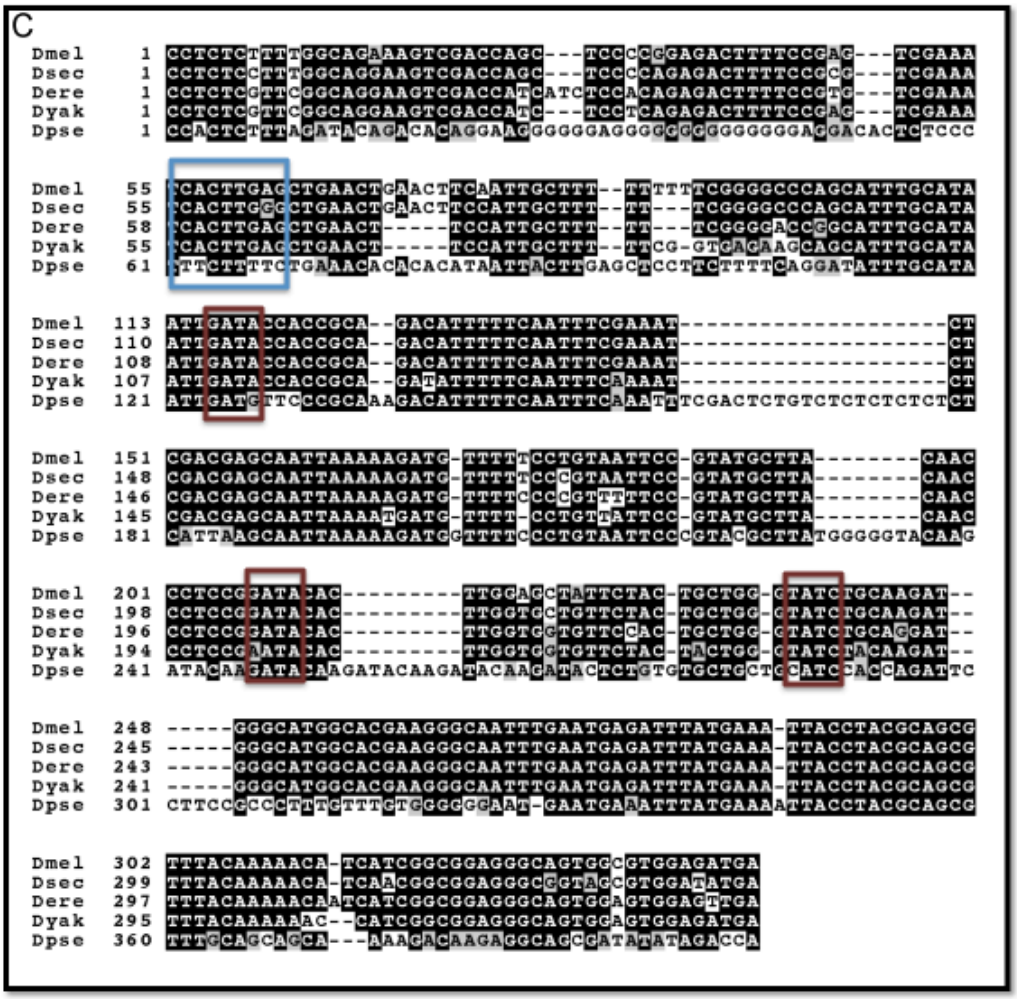


Figure 1: Identification of a *Mef2* Cardiac Enhancer
A: Cartoon of the *Mef2* gene and its cardiac enhancers in the 5' upstream region. The most distal green box refers to the Seven-up specific enhancer. The middle green box represents the Tinman-specific enhancer and the yellow box is the late stage cardiac and somatic mesodermal enhancer.
B: Activity of the late stage enhancer fused to a *lacZ* reporter. The embryos are stained for an antibody against β-Galactosidase. The enhancer does not become active until stage 14 of embryogenesis. It is active in the somatic mesoderm and both cell types of the heart.
C: Alignment of the shadow enhancer sequence with four *Drosophila* species. A conserved Tinman binding site is marked by a blue boxe and Pannier sites are marked with red boxes.

Tinman and Pannier activate the enhancer in vitro and Tinman is capable of binding to its consensus sequence within the enhancer

To test the abiity of candidate factors to regulate the Mef2 enhancer, we transfected *Drosophila* S2 cells with the enhancer fused to a *lacZ* reporter gene

along with plasmids containing the cDNAs of either Tinman, Pannier or both. After incubation for 48hr, cell lysates were prepared and reporter activity was determined using a quantitative ßGal assay. There was moderate activation of the *Mef2-lacZ* construct with Tinman, while Pannier on its own was unable to significantly activate the enhancer. When Tinman and Pannier were combined however, activation was more than additive suggesting that the two might work synergistically to activate *Mef2* in the heart (Figure 2A).

Next we tested the ability of each factor to bind to the Mef2 enhancer in electrophoretic mobility shift assays using in vitro translated proteins and radioactively labeled DNA. Tinman bound to the consensus site within the enhancer robustly, as visualized by the presence of a protein plus probe complex in the assay (Figure 2B). The interaction with Tin was specific, because cold wild-type competitor was able to compete away binding, while cold mutant competitor (that had the consensus site replaced with an EcoRI site) was unable to reduce binding. In the Pnr binding assay, a non-specific band was detected which was competed away by both wild-type and mutant competitors suggesting the Pnr does not bind to this enhancer sequence.

The Tinman binding site is required for enhancer activity in vivo

Having demonstrated that Tin could bind to the consensus site in the enhancer, we next determined if the site was required for enhancer activity. Using site-directed mutagenesis, we replaced the Tin consensus site within the context of the full-length enhancer, fused the mutated enhancer to a *lacZ* reporter

and generated additional transgenic flies. When we analyzed reporter expression in transgenic embryos, we noted that *LacZ* reporter activity was slightly reduced

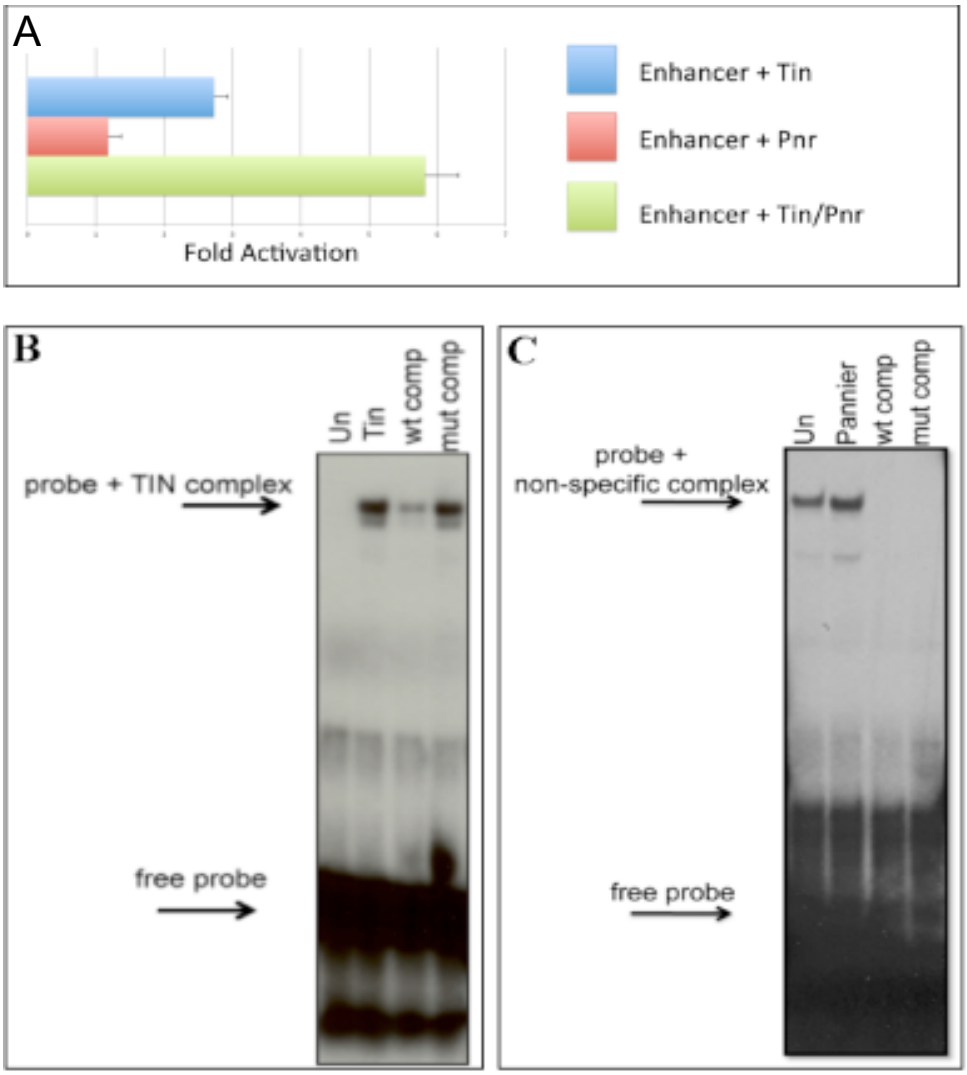


Figure 2: Tinman and Pannier activate the enhancer in vitro and Tinman is capable of binding to its consensus sequence within the enhancer
A: Activation of the late stage *Mef2* enhancer fused to a *lacZ* reporter gene in S2 cells by Tinman, Pannier or Tinman and Pannier. Tinman was able to activate the reporter moderately while activation by Pannier was not significant above negative controls. When combined, activation of the reporter was increased significantly above Tinman alone.
B: Electrophoretic mobility shift assay to determine if Tinman could bind to its consensus site. Free probe had a high mobility when combined with unprogrammed lysate (Un). A complex was formed in the presence of Tinman (Tin), which was competed by 300X excess of nonradioactive wild-type sequence (wt comp) but not by 300X excess of nonradioactive mutant sequence (mut comp).
C: Electrophoretic mobility shift assay to determine if Pannier could bind to its consensus site. Free probe had a non-specific complex formed with the unprogrammed lysate (un). A complex was formed with the addition of Pannier (Pnr), which ran slightly faster than the non-specific band. The band was competed by both 300X excess of nonradioactive wild-type sequence (wt comp) and 300X excess mutated sequence (mut comp).

in the somatic mesodermal cells, but still present. However, reporter expression was completely lost from all cells of the heart (Figure 3). This loss of enhancer activity was apparent at early and late stages of cardiogenesis, suggesting that Tin is a direct activator of this enhancer in vivo during the embryonic stage.

Ectopic expression of Tinman and Pannier results in expansion of enhancer activity

To further test the hypothesis that Tin and Pnr are direct and positive activators of Mef2 expression via this enhancer, we determined if ectopic expression of Tin could expand the activity of our enhancer. To achieve this, we initially generated embryos carrying the ectodermal and nervous system Gal4 driver *69Bgal4*, plus *UAS-tin* and the *Mef2* enhancer-*lacZ*. This combination drives expression of Tin in the ectoderm and ventral nerve cord. If Tin is a direct activator of the enhancer, we predict that these embryos should show lacZ expression in the ectoderm and/or nerve cord. Since the mesodermal expression of the enhancer is quite robust, we directed our attention to the ventral nerve cord to look for expansion of expression of the enhancer but failed to see any reporter activity in this tissue. We hypothesized that there might be expansion of the enhancer in other areas of the ectoderm that might be difficult to discern given the intensity of the mesodermal activity. To address this, we used site-directed mutagenesis to mutate the Lame duck consensus sequence, which is responsible for the somatic mesodermal expression. We generated additional transgenic flies carrying this mutated construct.

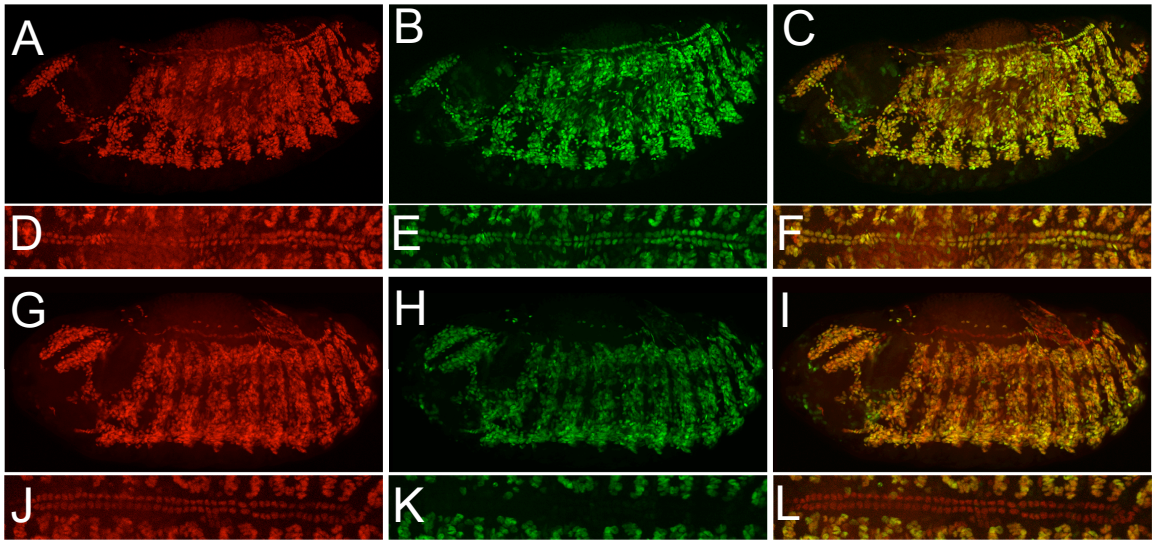


Figure 3: Mutation of the Tinman consensus sites results in loss of activity from the cardiac cells
A-C: Stage 14 late stage enhancer *lacZ* embryos. D-F: The same embryos at stage 16. A,D: Antibody stain against Mef2. Expression can be seen in all cells of the heart and throughout the somatic mesoderm. B,E: Antibody stain against β-Galactosidase. Activity of the late stage enhancer is almost identical to that of Mef2 expression. In B, the enhancer is just becoming active in the heart cells and a few cells are lacking activity but by stage 16 (E) all cardiac cells show activity. C: Merge of A and B. F: Merge of D and E.
G-I: Stage 14 late stage enhancer with the Tinman consensus site mutated. J-L: The same embryos at stage 16. G,J: Antibody stain against Mef2. Expression can again be seen in all cells of the heart and throughout the somatic mesoderm. H,K: Antibody stain against β-Galactosidase. Activity of the mutated enhancer is completely lost from all of the cardiac cells and is reduced in the somatic mesoderm. I: Merge of G and H. L: Merge of J and K. Arrow heads point to MEF2 positive cardiac cells, arrows point to the same cells lacking β-Galactosidase.

In these embryos, *LacZ* expression was reduced in the somatic mesoderm as expected, with some patchy pockets of expression remaining. Strangely, at stage 16, the cardiac cell activity of the enhancer became inconsistent with random groups of cells losing expression (Figure 4B compared to 3B). This made some sense because when we tried to remove that last 170 bp of sequence from the enhancer, we lost all activity suggesting that 3' sequences are critical to enhancer activity. Despite the reduced cardiac activity at the later stage, we proceeded with the overexpression studies.

We repeated the ectopic Tin expression experiment with this new transgenic line, and saw significant ectopic expression of lacZ in the ectoderm demonstrating that Tin could activate enhancer-*lacZ* activity in vivo outside of the mesoderm (Figure 4H). Since Pnr and Tin collaborated to activate the *Mef2-lacZ* in tissue culture cells, we also tested whether Tinman required Pannier for more robust ectopic activation of the enhancer. We generated embryos carrying *UAS-tin;UAS-pnr*;*69gal4* and our new enhancer line. Ectopic expression was again observed in the ectoderm. However, this time expression was much more robust (Figure 4K). In addition, we observed ectodermal expression of *Mef2* in these embryos (Figure 4J), which further supports the collaboration of these two factors in activating *Mef2* through this enhancer.

Tinman and Pannier activate *hand and wb* transcription

Tinman and Pannier working together to activate our enhancer was consistent with numerous accounts in the literature of their collaboration and cross-regulation (Klinedinst and Bodmer, 2003, Gajewski et al., 1999, Gajewski et al., 2001, Han and Olson, 2005). Additionally, recent reports in vertebrates have demonstrated that the Tinman, Pannier and MEF2 homologs collaborate to convert mouse fibroblast cells to cardiac cells (Ieda et al., 2010 and Ifkovits et al., 2014). We wondered if we could use the *Drosophila* in vivo system in order to investigate this phenomenon further. First we tested our system by analyzing the

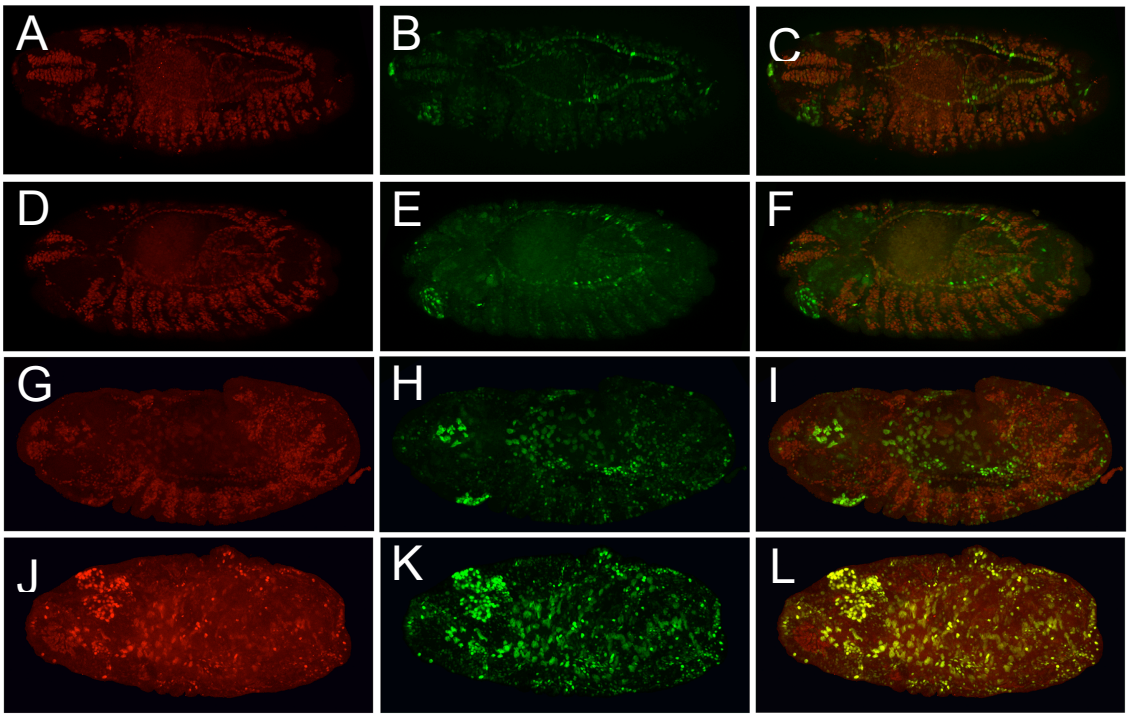


Figure 4: Ectopic expression of the Tinman and Pannier transcription factors results expansion of enhancer activity.
A-C: Stage 14 late stage enhancer *lacZ* with the Lame duck consensus site mutated. A: Antibody stain against MEF2. Expression can be seen in all cells of the heart and throughout the somatic mesoderm. B: Antibody stain against β-Galactosidase. At stage 14, activity of the enhancer can be seen in all cells of the heart but the somatic mesodermal stain is reduced. C: Merge of A and B.
D-F: Stage 14 69BGAL4;UAS-pnr;LMD mutant enhancer. D: Antibody stain against MEF2 showing normal muscle patterns. E: Antibody stain against β-Galactosidase. Enhancer activity is similar to B. F: Merge of D and E.
G-I: Early stage 69BGAL4;UAS-tin; LMD mutant enhancer. G: Antibody against MEF2 showing normal muscle patterns. H: Activity of the enhancer is significantly expanded in the ectoderm. I: Merge of H and I.
J-L: Early stage 69BGAL4;UAS-tin;pnr; LMD mutant enhancer. J: Antibody against MEF2 showing expression in the ectoderm. K: Activity of the enhancer is more robustly activated in the ectoderm when compared to H and co-localizes with the MEF2 expression seen in J.

effects of only Tinman and Pannier over-expression. It had been documented that these factors activate *Hand* and *Sur* (Klinedinst and Bodmer, 2003, Hendren et al., 2007, Akasaka et al., 2006) therefore we over-expressed *tinman* and *pannier* and assessed whether these factors were able to activate transcription of *Hand* and *Sur.* In approximately 20% of the embryos, we observed ectopic accumulation of *Hand* or *Sur* transcripts (Figure 5D, 5E). We also tested for

expression of *wingblister (wb)*, a gene whose enhancer has not yet been characterized, and interestingly found that in 100% of the embryos, *wb* expression was expanded, but only in the posterior of the heart (Figure 5F). These studies demonstrated that cardiac marker gene expression could be modestly expanded upon over-expression of *tinman* plus *pannier.*

Tinman, Pannier and MEF2 work in collaboration to dramatically activate the cardiac program

We next investigated the effect upon marker gene expression of addition of the third factor used in vertebrate conversion experiments, namely MEF2. With the addition of MEF2 to embryos over-expressing Tin and Pnr, the patterns of *Hand* and *Sur* transcripts were dramatically expanded in 100% of embryos stained (Figure 6A, B, D) while *wingblister* transcripts lost their normal pattern of expression. We interpret these results to indicate that MEF2 can potentiate the cardiogenic effects of expression of *tinman* and *pannier* in the mesoderm. To determine if these three factors could induce cardiac fate outside of the mesoderm, our last experiment was to test whether or not Tin, Pnr and MEF2 were potent enough to activate the cardiac program in the ectoderm, using the *69B-gal4* driver line. We found that for the majority of embryos stained, there was no expansion of cardiac markers into ectodermal tissues. However, in a small percentage of embryos, we could see ectopic expression of *Hand* and *Sur* but not *wingblister* (Figure 7A-D). These results suggested that conversion of

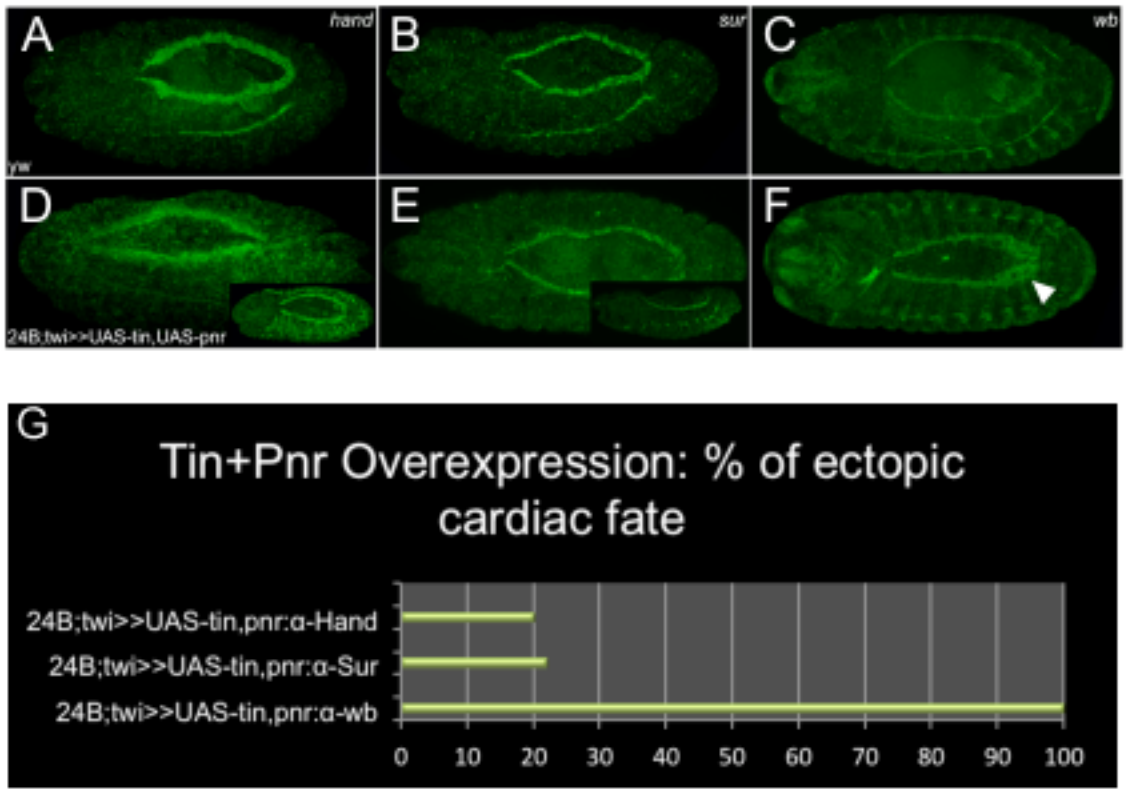


Figure 5: Over-expression of Tinman and Pannier in the mesoderm results in expansion of the expression of three cardiac factors.
A-C: Wild-type expression of *hand, sur and wb* transcripts.
D-F: 24B;twi>>UAS-tin,UAS-pnr embryos. D,E: *hand* and *sur* transcripts were expanded in ~20% of embryos. F: 100% of *wb* transcripts were expanded specifically in the posterior of the heart (arrowhead).
G: Quantification of ectopic cardiac fate.

to a cardiac fate requires some threshold level of activation by the converting factors that may or may not be met when utilizing an ectodermal driver.

Discussion

We have identified a *Mef2* enhancer that becomes active late in embryogenesis in all cells of the heart. We have shown that its activity is dependent upon a single Tin consensus binding site which is capable of binding to the Tin protein in vivo. Tin can activate the enhancer in cell culture and when

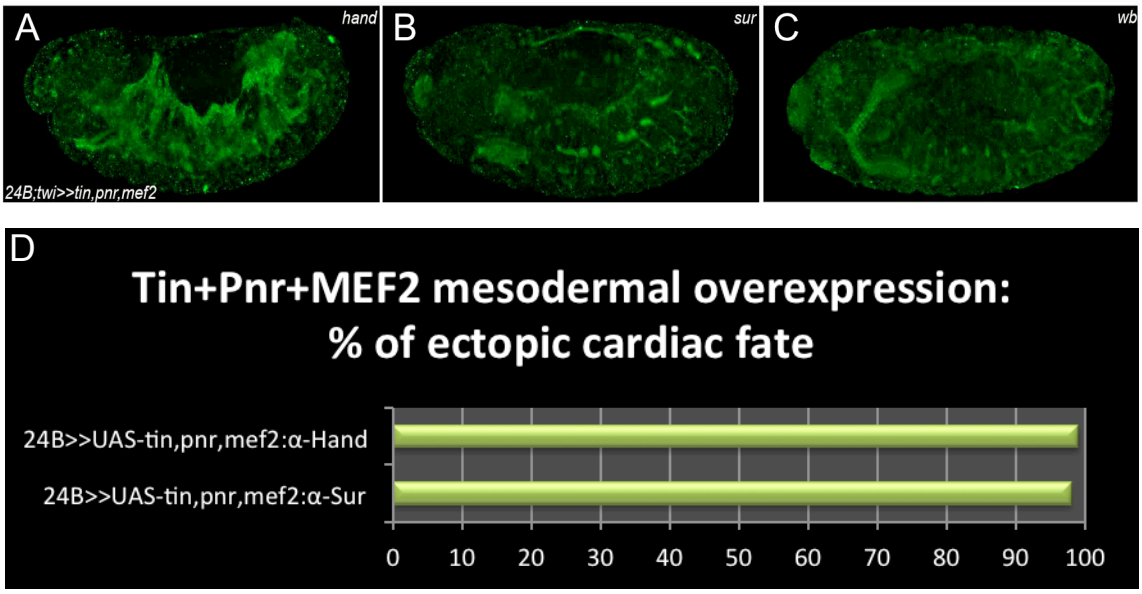


Figure 6: Over-expression of Tin, Pnr and MEF2 in the mesoderm results in a dramatic expansion of cardiac factors
A-C: 24B;twi>>tin,pnr,mef2 embryos. A,B: *hand and sur* transcripts are significantly expanded. C: *wb* transcripts have lost their normal pattern of expression. We were unable to determine if the levels were significantly different.
D: Quantification of ectopic cardiac fate.

the GATA transcription factor Pnr is added, there was synergistic activation of the enhancer. While we have shown that Pnr is unable to physically bind to the *Mef2* enhancer in vivo, Pannier has been shown previously to bind directly to Tin and synergistically activate the early cardiac enhancer of *Mef2 (*Gajewski et al., 2001). Ectopic expression of Tin in the ectoderm results in expansion of our enhancer, while combined ectopic expression of Tin and Pnr dramatically expands enhancer activity as well as MEF2 expression in the ectoderm. These

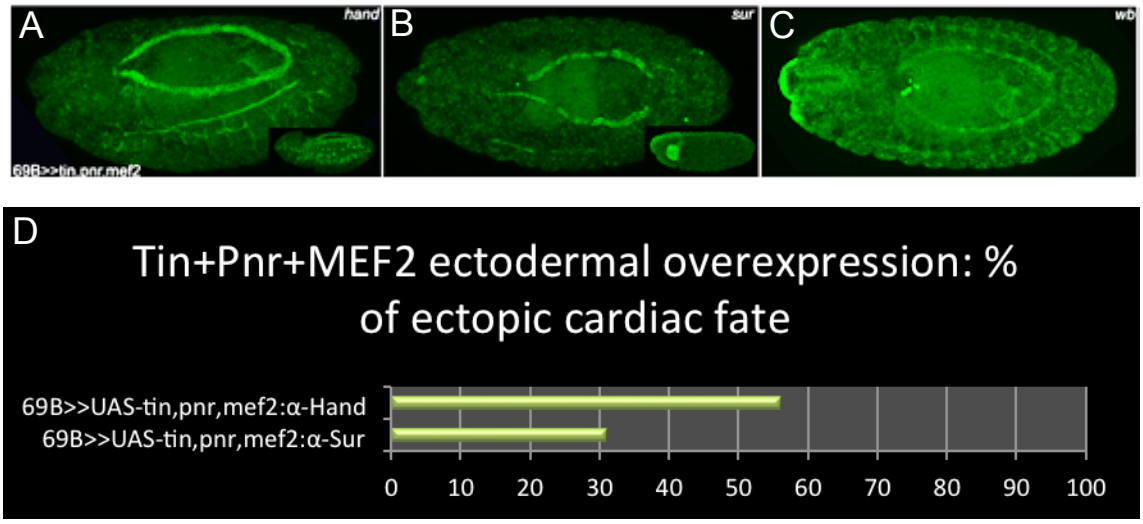


Figure 7: Ectopic expression of Tin, Pnr and MEF2 in the ectoderm results in a partial expansion of cardiac factors
A-C: 69B>>tin,pnr,mef2 embryos. A-C: The majority of embryos had normal transcripts but a small portion displayed expanded transcript expression in various patterns.
D: Quantification of ectopic cardiac fate.

results demonstrate the power of these two factors to regulate *Mef2* and the importance of maintaining MEF2 at high levels in the heart through the end of embryogenesis by having duplicate enhancers with similar activity.

Given the ability of Tin and Pnr to activate other cardiac factors such as *Hand and Sur*, we tested their ability to activate an additional cardiac factor *wingblister,* whose enhancer has not yet been identified. Interestingly, expansion of wb transcripts was observed but only in the posterior portion of the heart. This suggests that *wingblister* is differentially regulated along the anterior/posterior axis of the embryo and that a third factor, expressed in the posterior of the heart, is required in addition to Tin and Pnr to activate *wingblister* expression. A likely candidate is the Hox gene *abdominal* A, which has been shown to be required for

the expression of heart specific factors in the dorsal vessel (Lovato et al., 2002, Ponzielli et al., 2002).

Vertebrate conversion studies demonstrating the potency of Tin, Pnr and MEF2 homologs to generate cardioblasts from fibroblasts led us to hypothesize that they might work together the same way in *Drosophila* to activate the cardiac program more broadly. In fact, the addition of MEF2 to the over-expression studies presented here resulted in significantly expanded expression of both *Hand* and *Sur* transcripts in the mesoderm. Tin, Pnr and MEF2 were also able to activate *Hand* and *Sur* transcripts in the ectoderm, which is more difficult to do since it is a completely different germ layer and not the normal environment in which these factors would be active. The combination of Tinman, Pannier and MEF2 is clearly a more potent activator of two previously characterized cardiac enhancers. The *Mef2* gene itself possesses distinct enhancers that are both activated by Tin and Pnr. Tin has been shown to be an activator of Pnr in the heart early in development and conversely, Pnr has been shown to be required for maintaining proper levels of Tin in the heart (Klinedinst and Bodmer, 2003). This complicated cross-regulation is in line with previous work with MEF2 rescue experiments in which differing levels of MEF2 expression rescued different structures in the embryo (Gunthorpe et al., 1999). We hypothesize that multiple enhancers and overlapping regulation fine tune the levels of expression of regulatory and structural genes specifically to the structure being formed.

ACKNOWLEDGEMENTS

I would like to acknowledge Cheryl Sensibaugh for the identification of the minimal late stage enhancer. I also thank Melanie Martinez for work on the mutated Lame Duck consensus binding site enhancer as well as additional mutations within the enhancer for Pannier and Tinman sites. I would finally like to thank Kirstie Swingle for carrying out the cell culture assays with Tinman and Pannier.

Chapter 3

Identification of Novel Heart Genes in *Drosophila*

Abstract

The genes that are involved in the formation of the heart are highly conserved between *Drosophila* and vertebrates (Reviewed in Cripps and Olson, 2002). Despite the complete sequencing of the genomes of both of these organisms, we have yet to identify all of the genes that are specific to heart development and function. Utilizing a new technique (Bryantsev and Cripps, 2012), I generated a transgenic line, which expresses GFP in the heart, and used Fluorescent Activated Cell Sorting (FACS) to isolate heart cells from disrupted embryos. From 100,000 cells collected, I isolated RNA from these cells and control GFP-negative cells, and utilized RNAseq with the goal of identifying novel factors involved in the formation of the heart. I was able to identify approximately 1000 genes that were enriched in the heart, many of which have not yet been characterized, and confirmed a small selection of new genes for their enrichment in cardiac cells by RT-PCR. These studies provide a catalog of genes that can now be tested in the *Drosophila* system for roles in heart development. Since 34% of the identified genes are conserved in mammals, these studies have the potential to provide new insight into cardiac developmental mechanisms of higher animals.

Introduction

RNA sequencing is rapidly becoming a method of choice for identifying patterns of gene expression within a certain RNA population. RNA sequencing depends upon the efficient isolation of RNA and then subjecting the RNA to high-throughput sequencing in order to generate tens of millions of reads per sample. When these reads are mapped back to the annotated genome, the data can be use to identify gene structures, to identify novel genes, to define alternative splicing patterns of known genes, and to calculate the relative levels of gene expression between tissues (Reviewed in Wang et al., 2009, Brown et al., 2014).

Drosophila would be an ideal model organism to carry out such experiments due to its high genetic conservation to vertebrates and reduced incidence of gene duplications, which make knockout studies more enlightening. However, isolating specific cells from specific tissues during development, particularly during embryogenesis, has proved quite challenging. In particular, the heart comprises only 104 cells out of the many thousand cells in the embryo at the end of embryogenesis. To isolate this small number of cardiac cells, a new method is available that utilizes the Gal4-UAS system to drive expression of GFP in the heart (Salmand et al., 2011, Bryantsev and Cripps, 2012) and then use fluorescence activated cell sorting to extract these cells. We have followed this protocol with the ultimate goal of utilizing RNA-sequencing technology in order to identify genes expressed in the heart that have not yet been identified or characterized. I identified over 1,000 genes that were enriched in cardiac cells compared to non-cardiac cells, and confirmed that a subset of these cells have

enriched expression in the cardiac cells through RT-PCR or BDGP in situ hybridization. These studies provide a small database of potential cardiac genes that have not yet been characterized for further investigation.

Materials and Methods

Drosophila stocks

The *Sur-Gal4;UAS-2X GFP* transgenic line was generated from available laboratory stocks using standard genetic techniques.

Sample preparation for FACS

Embryos were collected and processed for FACS according to the methods described by Bryantsev and Cripps (2012). The processed cells were sorted with a Legacy MoFlo sorter (Beckman Coulter). Three modifications were made to the published protocol. Instead of elastase treatment, trypsin was used at 0.5 mg/ml for 10 minutes at room temperature. We wanted to see if trypsin was more efficient at dissociating cells. However, elastase and trypsin seem to work equally well. The Seecof saline was replaced with standard sheath fluid (BioSure Flow Cytometry Sheath Solution Preservative Free 8x concentrate, Cat. No. 1027). In the original protocol, the sorting solution used by the facility was not compatible with cell viability. However, this new sorting solution seemed to work well in other experiments and thus was used in the protocol. Finally, we used Vybrant DyeCycle Ruby stain (Life Technologies, V-10309) to stain nuclei owing to technical difficulties with the UV laser that read the Hoechst stain used previously.

RNA extraction for RNA-sequencing and RT-PCR

RNA was purified using the Qiagen RNeasy Mini Kit (#79254). An additional DNase treatment was carried out according to the handbook protocol (Qiagen 79254). cDNA was synthesized with Superscript II Reverse Transcriptase Kit (Invitrogen #11708) using 30 ng of RNA for each sample. The following primers were used for analysis:

Transcript	Forward primer 5'-3'	Reverse primer 5'-3'
Act 5C	AAGGATCGCTTGTCTGGG	GTATATCATATATATCTCATGTGG
Sur	CCTTGCTATCGGCGCTTCT	CGTGTCATTCAAGAGCCAAGGT
Tl	GTGGACGCCAGGATACAAGGA	TCGGAAGCGGCCTTTAGTC
Hand	AAAAGGAGAGAAGGCGAACC	CGACTGACCGGCTTAAGTTC
CG14857	GCACGTCATGGATTCGACC	TTACACACCCAGTTCAGCTCT
CG11085	AGAACAGATCGAAGGCGCAA	CGATAGGCAGGACTCCACAC
CG34371	GTCATCTGGTATCGCCAGGG	CAGCACCTGAAAGGAGTCGT

Analysis of RNA-sequencing results

Otogenetics (Norcross, GA) carried out polyA cDNA preparation, Illumina library preparation and HiSeq2000 paired-end sequencing on both our negative control and GFP-positive total RNA samples. They provided us with a basic bioinformatics analysis that included RMS-normalized RPKM values (Reads per kilobase per million reads). We calculated the percentage of reads that came from the GFP positive cells and used 80% as a cutoff for enrichment.

Results

Experimental Setup for FACS

The original *Sur-Gal4;UAS-2X GFP* transgenic line generated by Bryantsev and Cripps (2012) was lost. However, the individual homozygous lines for *Sur-Gal4* and *UAS-2X GFP* were still available. My first attempt to generate a homozygous line was unsuccessful. We hypothesized that the small number of flies selected early in the crosses to generate the line underwent a mutation that rendered them infertile. Upon repeating the cross, I was able to generate a viable homozygous line, which expressed GFP in the cardiac cells (Figure 1A).

Once generated, we propagated large numbers of the line and collected their embryos every two hours before storing them at 4 degrees C until a full day's worth of plates were collected. The embryos were then aged for 10 hours to reach approximately stage 13 of development according to the previously described protocol. Unfortunately, when we analyzed the embryos, there was a high incidence of lethality, which we believed to arise from the cold treatment. Thus, we deviated from the original protocol and collected embryos every two hours and aged them sequentially. It was important to not go beyond 10 hours of development for two reasons. First, we hoped to identify genes that might be involved in the specification of the heart and thus wanted to limit the purified cell population to cardiac precursors. Second, at later stages of development with this transgenic line, GFP is robustly expressed in the large salivary glands of the embryo, and we wanted to avoid collecting these cells (Figure 1B).

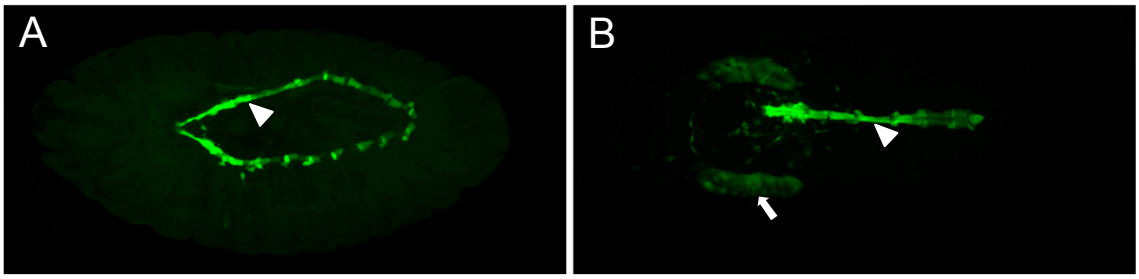


Figure 1: GFP Expression in the *Sur>>Gal4;UAS 2XGFP* Transgenic Line
A-B: Embryos stained with an antibody against the GFP protein. A is a stage 14 embryo that has not yet completed dorsal closure. B is stage a 16 embryo with a completely formed heart tube. Arrowheads point to heart cells and the arrow points to the salivary gland.

Embryos were then prepared for FACS as previously described with three small modifications (see Materials and Methods). The sorting facility was having technical difficulties with the laser reading the Hoechst emission and so we used the Vybrant DyeCycle Ruby to stain live nuclei. Cells were then selected for GFP and single Ruby positive nuclei. We avoided high intensity Ruby stained cells that suggested multiple cells or clumps to reduce the chance of co-isolating non-GFP cells that might have attached to a GFP positive cell (Figure 2).

We obtained approximately 10,000-20,000 cells per sort of both GFP positive cells as well as GFP-negative cells to serve as a negative control. Since the GFP positive cells were the limiting cells, we collected an equal number of non-GFP cells as a negative control. After six sorting appointments, all of the cells that had been collected were pooled directly into RNA extraction lysis buffer. This resulted in approximately 500 ng of cardiac cell and control RNA. A small amount of RNA from each sample was used in RT-PCR experiments in order to confirm the enrichment of known cardiac genes. Amplification of a cytoplasmic actin transcript, *Act5C*, showed equal intensity across both samples,

indicating that approximately equal amounts of cDNA had been generated for each sample. By contrast, transcripts of the cardiac genes *Hand, Sur,* and *Toll 1* transcripts were all enriched in the cardiac sample (Figure 3). With these data in hand, we felt confident enough to proceed with RNA sequencing of our collected RNA.

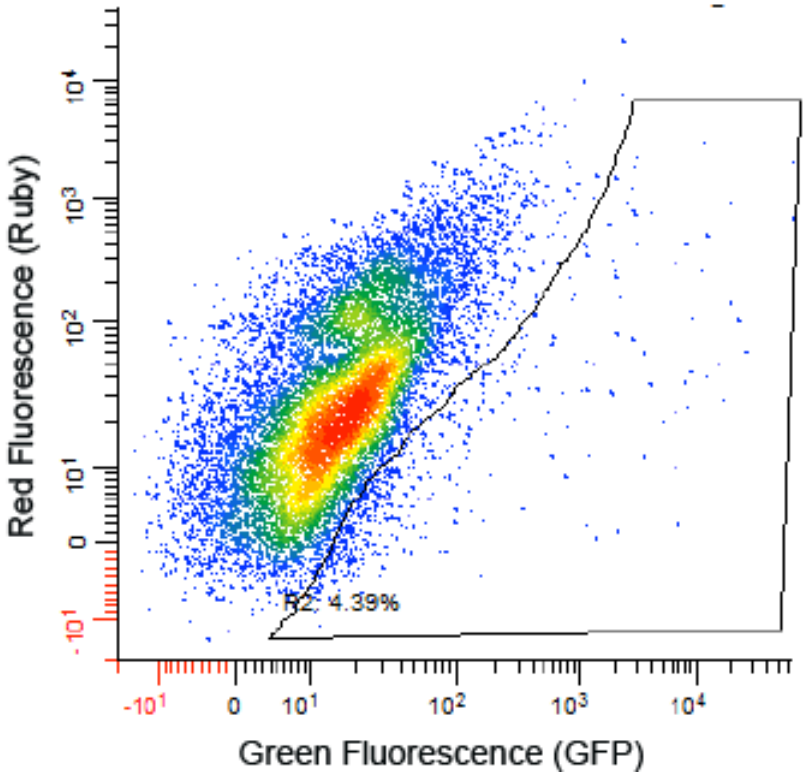


Figure 2: Isolation of cells co-expressing GFP and Ruby Vybrant Dye
Boxed cells indicate our target cells co-expressing GFP and Ruby stain. Cells expressing both stains at a very high density are thought to be clumps of cells and were avoided to reduce the collection of non-GFP cells that might be adherent to GFP positive cells.

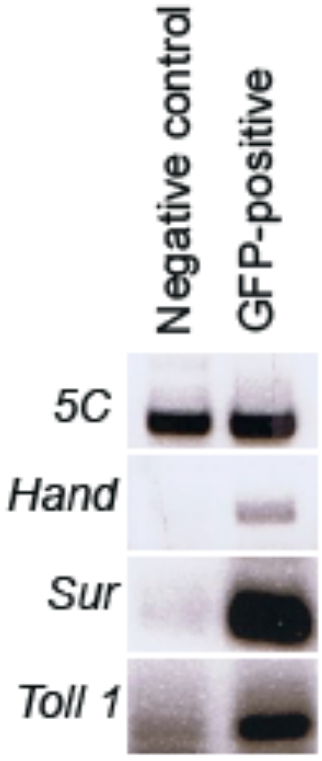


Figure 3: Semi-quantitative RT-PCR of known cardiac genes
5C transcripts were expressed equally in the non-GFP and GFP-positive cells and was used as an equal loading control. *Hand, Sur and Toll 1* transcripts were all enriched in the GFP-positive cells.

RNA sequencing of cardiac samples reveals enrichment for cardiac genes.

Sequencing was carried out by Otogenetics on an Illumina platform. Despite a very small amount of starting material, Otogenetics obtained 16 and 20 million confidently mapped reads for the negative control and GFP expressing cells respectively (Figure 4). Otogenetics supplied us with a basic bioinformatics analysis, which included a map alignment to the *Drosophila* genome and expression analysis in the form of RPKM (reads per kilobase per million reads), RMS-normalized read counts, as well as a Z-score for each read to measure its distance from the total mean of counts for each sample. This allowed us to

Negative Control		Sur>>GFP
Mapped reads	17,092,173	21,385,425
mapped confidently	16,049,459	20,024,133
mapped repetitively	1,042,714	1,361,292
Reads not mapped	9,964,609	12,485,811
no mapping	7,399,985	9,228,018
low quality	2,169,477	2,785,808
ribosomal RNA	221,348	228,424

Figure 4: Read details from RNA-sequencing
Otogenetics obtained 16 million confidently mapped reads from the negative control sample and 20 million from the GFP-positive sample. Each sample had similar numbers of reads that could not be mapped as well as low quality reads and ribosomal RNA contamination.

compare reads of shorter transcripts with very long ones (reads per kilobase) as well as compare reads between samples (per million reads). Of 11,000 total genes that were mapped to the genome, 984 were enriched in the GFP-positive cells defined by at least 80% or more of the total transcripts being expressed from the GFP positive cells compared to the negative control. This percentage could result from a relatively low number of RPKMs, as long as at least 80% came from the experimental group.

To assess the validity of our data, we first determined if known cardiac genes were enriched in our sample (Figure 5). Some genes expressed in the heart are also expressed in other tissues. For instance, *Mef2* is expressed in

Transcript	% Expressed in GFP ++	GFP++ RPKM	GFP - RPKM
Sur	99%	7418	87
tin	95%	3623	157
Hand	98%	5466	98
pnr	87%	763	99
H15	99%	166	1
tup	87%	1952	235
apt	85%	1.55E+04	2254
Nep4	98%	2893	57
prc	97%	2.78E+04	619
Tl	80%	1.54E+04	3334

Figure 5: Enriched Cardiac Transcripts in *Sur>>GFP Cells*
Known cardiac transcript expression in GFP positive cells and the percent of transcrpts being expressed in cardioblasts as compared to the negative control. Corresponding RPKM values are listed from both samples demonstrating the variance of RPKMs in known enriched transcripts.

all cells of the heart but is also very strongly expressed in the somatic mesoderm. Thus, we selected for our analysis heart transcripts that have only moderate expression in other tissues. We chose nine strongly expressed cardiac factors and found that they were all enriched in our GFP positive cells. However, one transcript, the pericardial cell marker *pericardin (prc),* was also enriched in our

GFP sorted cells, suggesting some contamination of cardiac cells with the closely associated pericardial cells.

Selection of potential cardiac genes

We selected five candidates from our pool of enriched, GFP positive genes (approximately 1000) that have not been characterized. We chose genes that had at least 500 RPKMs and 80% enrichment in our GFP positive sample. We also narrowed candidates down by utilizing the large amount of information contained on the Flybase website that might suggest the candidate is a legitimate protein-coding gene that is expressed in the embryo. Some examples of the information we looked at included the modENCODE temporal expression pattern that confirmed the gene's expression during embryogenesis, the predicted gene structure as well as predicted protein functional domains such as protein or DNA-binding. We designed primers for each for RT-PCR analysis. We amplified the transcripts from *Sur>>GFP* cell cDNA alongside the negative control cDNA, to determine if there was an enriched level of the candidate in the sample. In each case there was a clear enrichment of the candidate gene in the GFP positive cDNA (Figure 6). The results of these RT-PCR experiments confirmed the validity of the RNA sequencing data.

To determine if the genes that we identified show expression in cardiac tissue in vivo, we are in the process of carrying out in situ hybridization assays for each of the candidates validated by the RT-PCR experiments. In the meantime, we also studied data generated by the BDGP (Berkeley Drosophila Genome Project), which has assembled a database of high throughput in situ

hybridizations of *Drosophila* genes. They have tested approximately 7,500 genes so far. From our top candidate genes, two have been analyzed by BDGP, and both of these genes are expressed in the developing heart for (Figure 7).

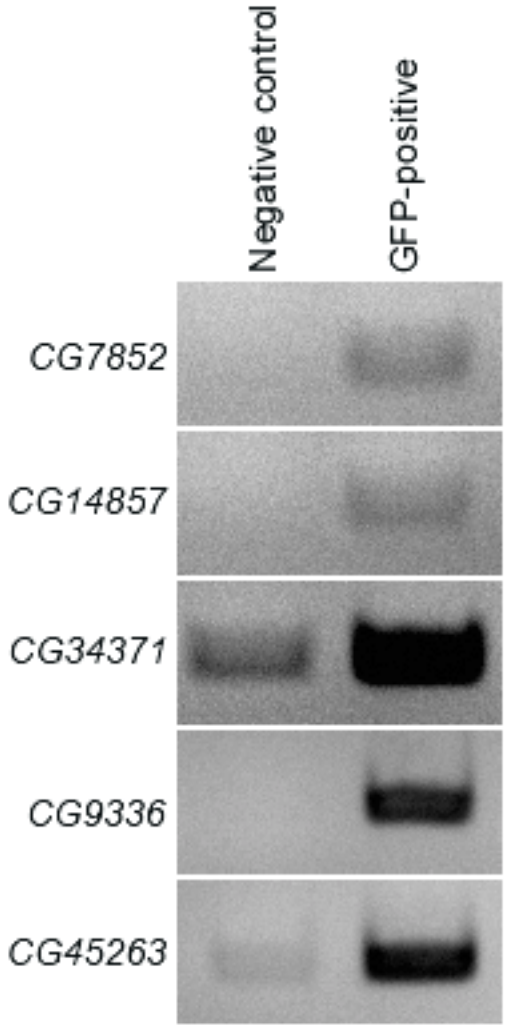


Figure 6: Semi-quantitative pcr of candidate heart genes
Each candidate gene was amplified from negative control and *GFP* positive cDNA. Each candidate expressed a much stronger band than the negative control if it had a band at all. The same cDNA was used from Figure 3.

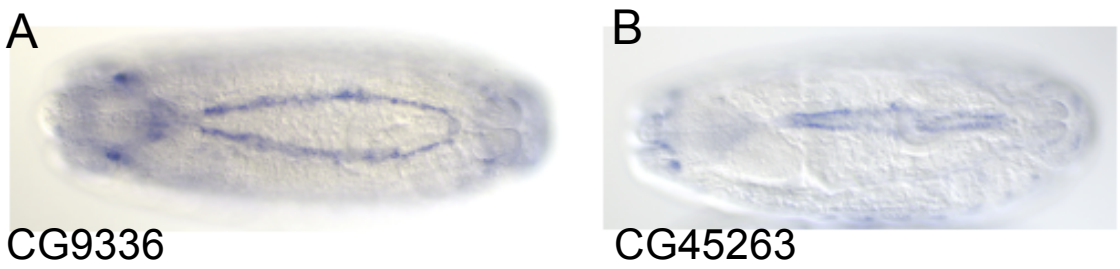


Figure 7: BDGP in situ hybridization of two candidate heart genes
A-B: High throughput in situ hybridization of transcripts from the CG9336 and CG45263 genes. Expression can be seen very specifically in the heart cells.

Discussion

We used FACS to isolate the small number of cells in the *Drosophila* embryo that comprise the cardiac mesoderm. We obtained sufficient RNA from these cells to carry out RNA sequencing, and we worked to validate experimentally the results we have obtained. We have analyzed the sequenced transcripts and looked for enrichment in known cardiac specific transcripts. We found that there was enrichment in expected cardiac factors but we also observed enrichment in the closely associated pericardial marker *pericardin.* This finding may result from expression of *prc* in the heart cells at an early stage; more likely, some pericardial cells adhere to the cardiac population during purification. It should be possible to reduce contamination with pericardial cells by a slightly prolonged treatment with proteases, to ensure thorough separation of cells. Nevertheless, our results showed strong enrichment for known cardiac factors in the sorted population. In addition, there is some evidence of salivary gland contamination in our enriched transcripts, which was identified by BDGP in situ hybridization data. While we took care not to include older embryos, it is possible they were contained in our sample. One way of avoiding this

contamination in the future will be to use two different cardiac markers stained with fluorescent tags. In this way we can select for cells that are only positive for both emissions, which, will reduce the chance of obtaining contaminants.

RNA sequencing of specific tissues can provide insight into the particular state of a tissue at a given developmental state or environmental stressor. While *Drosophila* is normally an ideal model organism for most studies, RNA sequencing of embryonic cardioblasts has proved challenging due to the small percentage of cardiac cells present at the end of embryogenesis. We feel that despite the likelihood of some contamination, we are beginning to generate a list of cardioblast enriched genes that warrant experimental investigation. Even if only a small percentage of our candidates turn out to be cardiac specific, it will provide a considerable addition to the very small database of known cardiac factors. Despite the limited number of known cardiac genes to date, many cardiac diseases have been attributed to mutations in them. For example in the vertebrate Tin homolog (NKX2-5), a dominant mutation has been mapped that leads to malformed cardiac structures and irregular conduction (Schott et al., 1998). Mutations in the vertebrate homologs for the cardiac transcription factors H15 (TBX5) and Pnr (GATA4) have both been shown to cause septal defects (Basson et al.,1997; Garg et al., 2003). As mentioned previously, 74% of human diseases attributed to a mutation in a specific gene, have a *Drosophila* homolog (Online Mendelian Inheritance in Man (OMIM)). Identification of previously unknown members of the cardiac program could prove useful in identifying the underlying causes of cardiac disease and defects.

ACKNOWLEDGEMENTS

I would like to thank Anton Bryantsev for all of his technical assistance and advice for this project. I would also like to thank all of the Cripps lab members for helping me at odd hours to collect and age the embryos so that I didn't have to sleep at the lab. I couldn't have done it without the support.

LITERATURE CITED

Akasaka, T., Klinedinst, S., Ocorr, K., Bustamante, E.L., Kim, S.K. and Bodmer, R. (2006). The ATP-sensitive potassium (KATP) channel-encoded dSUR gene is required for *Drosophila* heart function and is regulated by Tinman. *Proc. Natl. Acad. Sci. U.S.A.* 103(32):11999-12004.

Baker, P., Kelly Tanaka, K.K., Klitgord, N. and Cripps R.M. (2005). Adult myogenesis in *Drosophila melanogaster* can proceed independently of myocyte enhancer actor-2. *Genetics.* 170:1747-1759.

Basson, C.T., Bachinsky, D.R., Lin, R.C., Levi, T., Elkins, J.A., Soults, J., Grayzei, D., Krompouzou, E., Traill, T.A., Leblanc-Straceski, J., Renault, B., Kucherlapati, R., Seidman, J.G. and Seidman, C.E. (1997). Mutations in human TBX5 cause limb and cardiac malformation in Holt-Oram syndrome. *Nat Genetics* 15(1):30-4.

Baylies, M.K., and Bate, M., (1996). Twist: A myogenic switch in *Drosophila*. *Science.* 272:1481-1484.

Bier, E. and Bodmer, R. (2004). *Drosophila*, an emerging model for cardiac disease. *Gene* 342, 1-11.

Bischof, J., Maeda, R.K., Hediger, M., Karch, F. and Basler, K. (2007). An optimized Transgenesis system for Drosophila using germ-line-specific phiC31 integrases. *Proc Natl Acad Sci*. 104, 3312-3317.

Black, B.L. and Cripps, R.M. (2010) Myocyte Enhancer Factor 2 Transcription Factors in Heart Development and Disease. In: Nadia Rosenthal and Richard P. Harvey, editors, Heart Development and Regeneration Volume 2. Oxford: Academic Press, 2010 pp.
673-699.

Black, B.L. and Olson, E.N. (1998). Transcriptional control of muscle development by myocyte enhancer factor-2 (MEF2) proteins. *Annu Rev Cell Dev Biol.* 14:167-96.

Bodmer, R. (1993). The gene *tinman* is required for the specification of the heart and visceral muscles in *Drosophila. Development* 118, 719-729.

Bodmer, R. and Frasch, M. “Development and Aging of the *Drosophila* Heart.” *Heart Development and Regeneration Volume 1.* Ed. Nadia Rosenthal and Richard P. Harvey. London: Elsevier, 2010.

Bour, B.A., Obrien, M.A., Lockwood, W.L., Goldstein, E.S., and Bodmer, R. (1995). *Drosophila* MEF2, a transcription factor that is essential for myogenesis. *Genes and Development.* 9:730-741.

Brown, J.B. Boley, N., Eisman, R., May G.E., Stoiber, M.H., Duff, M.O., Booth, B.W., Wen, J., Park, S., Suzuki, A.M., Wan, K.H., Yu, C., Zhang, D., Carlson, J.W., Cherbas, L., Eads, B.D., Miller, D., Mockaitis, K., Roberts, J., Davis, C.A., Frise, E., Hammonds, A.S., Olson, S., Shenker,S. and Sturgill, D., Samsonova, A.A., Weiszmann, R., Robinson, G., Hernandez, J., Andrews, J., Bickel, P.J., Carninci, P., Cherbas, P., Gingeras, T.R., Hoskins, R.A., Kaufman, T.C., Lai, E.C., Oliver, B., Perrimon, N., Graveley, B.R. and Celniker, S.E. (2014). Diversity and dynamics of the *Drosophila* transcriptome. *Nature* doi:10.1038/nature12962.

Bryantsev, A.L. and Cripps, R.M. (2012). Purification of cardiac cells from *Drosophila* embryos. *Methods.* 56:44-59.,

Carpenter, J.M. (1950). “A new semisynthetic food medium for *Drosophila*.” *Drosophila Inform. Serv.* 24:96-97.

Casso, D., Ramirez-Weber, F.-A. and Kornberg, T.B. (1999). GFP-tagged balancer chromosomes for *Drosophila melanogaster. Mechanisms of Development.* **88:** 229-232.

Chakshusmathi, G., Mondal, K., Lakshmi, G.S., Singh, G., Roy, A., Babu Ch, R., Madhusudhanan, S. and Varadarajan, R. (2004). Design of temperature-sensitive mutants solely from amino acid sequence. *Proceedings of the National Academy of Sciences.* 101(21):7925-7930.

Cox, V.T. and Baylies, M.K. (2005). Specification of individual Slouch muscle progenitors in *Drosophila* requires sequential Wingless signaling. *Development.* 132: 713-724.

Cripps, R.M., Zhao, B. and Olson, E.N. (1999). Transcription of the Myogenic Regulatory Gene *Mef2* in Cardiac, Somatic, and Visceral Muscle Cell Lineages is Regulated by a Tinman-Dependent Core Enhancer. *Developmental Biology* 215, 420-430).

Cripps, R.M. and Olson, E.N. (2002). Control of Cardiac Development by an Evolutionarily Conserved Transcriptional Network. *Developmental Biology* 246, 14-28.

Duan, H., Skeath, J.B. and Nguyen, H.T. (2001) *Drosophila* Lame duck, a novel member of the Gli superfamily, acts as a key regulator of myogenesis by

controlling fusion-competent myoblast development. *Development* 128, 4489-4500.

Edgar, R.S. and Lielausis, I. (1963). “Temperature-Sensitive Mutants of Bacteriophage T4D: Their Isolation and Genetic Characterization.” *Genetics.* 49:649-662.

Fincham, J.R.S. (1966). Genetic Complementation. New York: W.A. Benjamin, Inc.

Garg, V., Kathiriya, I.S., Barnes, R., Schluterman, M.K., King, I.N., Butler, C.A., Rothrock, C.R., Eapen, R.S., Hirayama-Yamada, K., Joo, K., Matsuoka, R., Cohen, J.C. and Srivastava, D. (2003). GATA4 mutations cause human congenital heart defects and reveal an interaction with TBX5. *Nature* 424(6947):443-7.

Goldstein, E.S., Treadway, S.L., Stephenson, A.E., Gramstad, G.D., Keilty, A., Kirsch, L., Imperial, M., Guest, S., Hudson, S.G., LaBell, A.A., O’Day, M., Duncan, C., Tallman, M., Cattelino, A. and Lim, J. (2001). A genetic analysis of the cytological region 46C-F containing the *Drosophila melanogaster* homolog of the jun proto-oncogene. *Molecular and General Genomics.* 266:695-700.

Gordon, C.L. and King, J. (1993). Genetic Propeties of Temperature-Sensitive Folding Mutants of the Coat Protein o Phage P22. *Genetics.* 136:427-438.

Gulick, T. and Zhu, B. (2004). Phosphorylation and alternative pre-mRNA splicing converge to regulate myocyte enhancer factor 2C activity. *Mol Cell Biol.* 24(18): 8264-75.

Gajewski, K., Kim, Y, Lee Y.M., Olson, E.N. and Schulz, R.A. (1997) D-mef2 is a target for Tinman activation during Drosophila heart development. *Embo J* 16(3): 515-22.

Gajewski, K., Choi, C.Y., Kim, Y. and Schulz, R.A.(2000) Genetically distinct cardial cells within the *Drosophila* heart. *Genesis* 28(1): 36-43.

Gajewski, K., Zhang, Q., Choi, C.Y., Fossett, N., Dang, A.. Kim, Y.H., Kim, Y. and Schulz, R.A. (2001). Pannier is a transcriptional target and partner of Tinman during *Drosophila* cardiogenesis. *Developmental Biology* 2, 425-436.

Hamamori, Y., Hung-Yi, WU, Sartorelli, V. and Kedes, L. (1997). The basic domain of myogenic basic helix-loop-helix (bHLH) proteins is the novel target for direct inhibition by another bHLH protein, Twist. *Molecular and Cellular Biology.* 17(No.11): 6563-6573.

Han, Z. and Olson, E.N. (2005). *Hand* is a direct target of Tinman and GATA factors during *Drosophila* cardiogenesis and hematopoiesis. *Development* 132, 3525-3536.

Hanh, N.T. and Xu, X. (1998). *Drosophila mef2* Expression during Mesoderm Development Is Controlled by a Complex Array of cis-Acting Regulatory Modules. *Developmental Biology* 204, 550-566.

Hartwell, L.H., Culotti, J., and Reid, B. (1970). Genetic Control of the Cell-Division Cycle in Yeast, I. Detection of Mutants. *Proceedings of the National Academy of Sciences.* 66(No2):352-359.

Hendren, J.D., Shah, A.P., Arguelles, A.M. and Cripps, R.M. (2007). Cardiac expression of the *Drosophila* Sulphonylurea receptor gene is regulated by an intron enhancer dependent upon the NK homeodomain factor Tinman. *Mechanisms of Development* 134(6):416-26.

Hong, J.W., Hendrix, D.A. and Levine, M.S. (2008). Shadow enhancers as a source of Evolutionary novelty. *Science* 321: 1314.

Huang, K., Louis, J.M., Donaldson, L., Lim, F.L., Sharrocks, A.D. and Clore, G.M. (2000). Solution structure of the MEF2A-DNA complex: structural basis for the modulation of DNA bending and specificity by MADS-box transcription factors. *Embo j.* 19:2615-28.

Kelly, K.K., Meadows, S.M. and Cripps, R.M. (2002). *Drosophila* MEF2 is a direct regulator of Actin 57B transcription in cardiac, skeletal, and visceral muscle lineages. *Mechanisms of Development.* 110(1-2):39-50.

Kiehart, D.P. and Feghali, R. (1986). Cytoplasmic myosin from *Drosophila melanogaster. Journal of Cell Biology.* 103:1517-1525.

Klinedinst, S. and Bodmer, R. (2003). Gata factor Pannier is required to establish competence for heart progenitor formation. *Development* 130, 3027-3038.

Lècuyer, E., Parthasarathy, N. and Krause, H.M. (2008). Fluorescent in situ hybridization protocols in *Drosophila* embryos and tissues. *Methods in Molecular Biology* 420:289-302.

Lilly, B., Galewsky, S., Firulli, A.B., Schulz, R.A. and Olson, E.N. (1994). D-MEF2: A MADS box transcription factor expressed in differentiating mesoderm and muscle cell lineages during *Drosophila* embryogenesis. *Proceedings of the National Academy of Sciences USA.* 91:5662-5666.

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.* 267:688-693.

Lints, T.J., Parsons, L.M., Hartley, L., Lyons, I. and Harvey, R.P. (1993). *Nkx-2-5*: a novel murine homeobox gene expressed in early heart progenitor cells and their myogenic descendants. *Devolpment.* 119, 419-431.

Lovato, T.L., Nguyen, T.P., Molina, M.R. and Cripps, R.M. (2002). The Hox gene *abdominal-A* specifies heart cell fate in the *Drosophila* dorsal vessel.

Martin, J.F., Miano, J.M., Hustad, C.M., Copeland, N.G., Jenkins, N.A. and Olson, E.N. (1994). A *Mef2* gene that generates a muscle-specific isoform via alternative mRNA splicing. *Mol Cell biol.* 14; 1647-56. *Development* 129, 5019-5027.

Molina MR and **Cripps RM.** (2001) Ostia, the inflow tracts of the Drosophila heart, develop from a genetically distinct subset of cardial cells. *Mech. Dev.* **109:** 51-59.

Molkentin, J.D., Black, B.L., Martin, J.F. and Olson E.N. (1995). Cooperative activation of muscle gene expression by MEF2 and myogenic bHLH proteins. *Cell*. 83:1125-36.

Molkentin, J.D., Black, B.L., Martin, J.F. and Olson, E.N. (1996). Mutational analysis of the DNA binding, dimerization, and transcriptional activation domains of MEF2C. *Molecular and Cellular Biology.* 16(No.6):2627-2636.

Molkentin, J.D., Li, L. and Olson, E.N. (1996). Phosphorylation of the MADS-Box transcription factor MEF2C enhances its DNA binding activity. *J Biol Chem* 271:17199-204.

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 from the National Academy of Sciences USA.* 91:7520-7524.

Nguyen, T., Wang, J.B. 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.* 70:438-446.

Nguyen, H.T., and Xu, X. (1998). *Drosophila* mef2 expression during mesoderm Development is controlled by a complex array of cis-acting regulatory modules. *Dev. Biol.* 204, 550-566.

Olson, E.N. (2006). Gene Regulatory Networks in the Evolution and Development of the Heart. *Science* 313, 1922-1927.

Patel, N.(1994). Imaging neuronal subsets and other cell types in whole-mount *Drosophila* embryos and larvae using antibody probes. In "*Methods in Cell Biology* (Vol.44)*. Drosophila melanogaster:* Practical Uses in Cell Biology" (L.S.B. Goldstein and E. Fyrberg, eds) Academic Press, New York.

Perry, M.W., Boettiger, A.N., Bothma, J.P. and Levine, M. (2010). Shadow Enhancers Foster Robustness of *Drosophila* Gastrulation. *Current Biology* 20, 1562-1567.

Ponzielli, R., Astier, M., Chartier, A., Gallet, A., Therond, P. and Semeriva, M. (2002). Heart tube patterning in *Drosophila* requires integration of axial and segmental information provided by the Bithorax Complex genes and hedgehog signaling. *Development* 129(19):4509-4521.

Potthoff, M.J. and Olson, E.N. (2007). MEF2: a central regulator of diverse developmental programs. *Development.* 134(23):4131-40.

Peckham, M., Cripps, R.M., White, D.C.S. ad Bullard, B. (1992). Mechanics and protein content of insect flight muscles. *J. Exp. Biol.* 168:57-76.

Ryan, K.M., Hendren, J.D., Helander, L.A. and Cripps, R.M. (2007). The NK homeodomain transcription factor Tinman is a direct activator of seven-up in the Drosophila dorsal vessel. *Developmental Biology*. 302(2):694-702.

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*. *Dev Biol.* 171(1): 169-181.

Sadler, J.R. and Novick, A. (1964). "The Properties of Repressor and the Kinetics of its Action." *Journal of Molecular Biology.* 12:304-327.

Salmand, P-A., Iche-Torres, M and Perrin, L. (2011). Tissue-specific cell sorting from *Drosophila* embryos. *Methods and Technical Advances. Fly* 5:3, 261-265.

Santelli, E. and Richmond, T.J. (2000). Crystal Structure of MEF2A Core Bound to DNA at 1.5 Å Resolution. *J. Mol. Biol.* 297:437-449.

Schott, J.J., Benson, D.W., Basson, C.T., Pease, W., Silberbach, G.M., Moak, J.P., Maron, B.J., Seidman, C.E. and Seidman, J.G. (1998). Congentital

heart disease caused by mutations in the transcription factor NKX2-5.*Science* 281(5373):108-11.

Schulz, R.A., Chromey, C., Lu, M.F., Zhao, B., and Olson, E.N. (1996). Expression of the D-MEF2 transcription in the *Drosophila* brain suggests a role in neuronal cell differentiation. *Oncogene* 12, 1827-1831

Smith, D.H., Berget, P.B. and King, J. (1980). Temperature-sensitive mutants blocked in the folding or subunit assembly of the bacteriophage P22 tail-spike protein I. Fine-structure mapping. *Genetics.* 2:331-52.

Sundberg, H.A. and Davis, T.N. (1997). A mutational analysis identifies three functional regions of the spindle pole component Spc110p in Saccharomyces cerevisiae. *Mol Biol Cell.* 12:2575-90.

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 primordial and differentiated cells of the embryonic somatic, visceral and heart musculature. *Mechanisms of Development.* 50:29-41.

Thisse, B., Messal, M.E. and Perrin-Schmitt, F. (1987). The *twist* gene: Isolation of a *Drosophila* zygotic gene necessary for the establishment of dorsoventral pattern. *Nucleic Acids Res.*

Wang, Z, Gerstein, M, and Snyde, M. (2009). RNA-Seq: a revolutionary tool for transcriptomics. *Nature Reviews Genetics* 10, 57-63.

Watakabe, A., Komatsu, Y., Ohsawa, S. and Yamamori, T. (2010). Fluorescent in situ hybridization technique for cell type identification and characterization in the central nervous system. *Methods* 52(4): 367-74.

Zhang, S. and Bernstein, S.I. (2001). Spatially and temporally regulated expression of myosin heavy chain alternative exons during *Drosophila* embryogenesis. *Mech Dev* 101(1-2):35-45.