The Dissertation Committee for Tara Lynn Rasmussen Certifies that this is the approved version of the following dissertation:

# An Intrinsic Requirement for *Smyd1* in Mouse Cardiac and Skeletal Muscle

Committee:

Phillip W. Tucker, Supervisor

Susan E. Bergeson

Scott W. Stevens

Jeffrey M. Gross

Edward M. Mills

# An Intrinsic Requirement for *Smyd1* in Mouse Cardiac and Skeletal Muscle

by

Tara Lynn Rasmussen, B.S.

Dissertation

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

Doctor of Philosophy

The University of Texas at Austin

May, 2008

# Dedication

To Paul D. Gottlieb, a great scientist and mentor, without whom this project could never have existed.

# Acknowledgements

I thank everyone who allowed this project to continue despite the loss of Dr. Gottlieb, especially Dr. Phillip Tucker. I thank all current and previous contributors to the project, including June Harriss, Dr. Inkyu Hwang, Dr. Rob Sims, Dr. Li Zhu, and Dr. Mark Brown, as well as our collaborators Eric Olson, Rhonda Bassel-Duby, James Richardson, Dillon Phan, Deepak Srivistava, Chong Park, and Stephanie Pierce. I also thank all past and current members of the Gottlieb and Tucker labs.

I would like to thank my parents for their love and support and my older siblings, Kirk and Chelsey, for being good role models and setting high standards of acheivement. I would like to thank my husband, Todd, for supporting, loving, and bearing with me through the last several years. All of my friends in Austin and around the country have been an integral part of my success. Without my friends or family I would not be in this position today.

# An Intrinsic Requirement for *Smyd1* in Mouse Cardiac and Skeletal Muscle

Publication No._____

Tara Lynn Rasmussen, PhD

The University of Texas at Austin, 2008

Supervisor: Phillip W. Tucker

*Smyd1* is the founder of a gene family whose members contain split SET and MYND Domains. *Smyd1* has several SET dependent lysine methyl-transferase substrates, including multiple histone lysines and at least one non-histone protein, *skNAC*. The MYND domain of *Smyd1* is required for protein interactions, such as that with *skNAC*. Conventional *Smyd1* knockouts die at E10 due to cardiac defects, including an enrichment of cardiac jelly, a decrease in trabeculation, and the loss of ventricular septation. *dHand*, a transcription factor specific for right ventricular development, and *Irx4*, a ventricle specific gene, are down regulated.

I have shown that an approximately one kb stretch of DNA sequence upstream of the muscle specific first exon of *Smyd1* is sufficient to drive expression of a reporter in transgenic mice. Cardiac specific expression is mediated by a proximal *Mef2* binding site whereas skeletal muscle expression is dependent on E-boxes. I have fully analyzed this

stretch of sequence via computational methods and made predictions on other potential regulatory factors.

Through the use of *Cre* mediated conditional knockouts, I have shown that the phenotype of the conventional knockout was not due to the introduction of the Neomycin cassette at the gene locus or due to cell non-autonomous effects on the heart. *Smyd1* is not only essential for cardiac septation, but throughout embryonic cardiac development, during embryonic skeletal muscle development, and in adult cardiac tissue. Conditionally deficient *Smyd1* embryonic hearts are less affected than conventional *Smyd1* knockouts, but are embryonically lethal and show poor trabeculation, cardiac hemorrhaging, and a pericardial edema. I detail that the *Nkx2.5-Cre* mediated *Smyd1* deletion phenocopies the *skNAC* conventional knockout and that both knockouts have similar changes in the expression levels of several genes. Furthermore, when *Smyd1* is conditionally removed from adult cardiac tissue, survival rates are diminished.

Surprisingly a skeletal muscle specific CKO of *Smyd1* mediated by *Myogenin-Cre* has resulted in perinatal lethality, with a visible phenotype as early as E15. Evident in the phenotype is a large edema between the epithelium and skeletal muscle, fewer myoblasts, decreased muscle mass, increased degenerating cells, and a potentially defective differentiation process.

# Table of Contents

# List of Tables

# List of Figures

# Chapter 1. Introduction and Background

## 1.1 *Smyd1* Discovery, Gene Layout, and Expression Pattern

During the process of studying the *CD8β* locus, a novel transcript was discovered that initiated at the same locus. This transcript belonged to a gene that initiated 200bp upstream of *CD8β* in the opposite orientation. This gene was entitled *Bop* for *CD8β* opposite and was later renamed *Smyd1* (SET and MYND Domains), the founder of a evolutionarily distinct family of genes containing a SET domain split by a MYND domain.

Initial studies determined that the mouse *Smyd1* encoded 3 alternatively spliced isoforms: *Smyd1a*, *Smyd1b*, and *Smyd1c* (Fig. 1.1). *Smyd1c* has a unique first exon (exon 1a) including both 5'UTR and coding sequence; *Smyd1c* also has a unique 3'UTR. *Smyd1a* and *b* exclude exon 1a and initiate at exon 1b, which is separated from exon1a by a 70kb intron. *Smyd1a* and *b* differ only by the small 39 base pair exon 5 that is included in *Smyd1a* and *c*, but excluded from *Smyd1b*. Transcripts from *Smyd1a* and *b* are restricted to cardiac and skeletal muscle in embryonic and adult tissue while *Smyd1c* is specific for thymocytes in adult and embryonic tissue.

Because the initial coding exons are distinct and separated by such a vast intron, it is logical that the regulatory regions are encoded by different sequences. Given the proximity to one another, *Smyd1c* is likely co-regulated with *CD8β*. It is expected that the regulatory region of the muscle specific isoforms functions independently of other genes and is adjacent to exon 1b. In other words, it is likely that motifs within this region are responsible for gene activation and repression in cardiac and skeletal muscle tissue.

**Fig 1.1 *Smyd1* Genomic Structure**

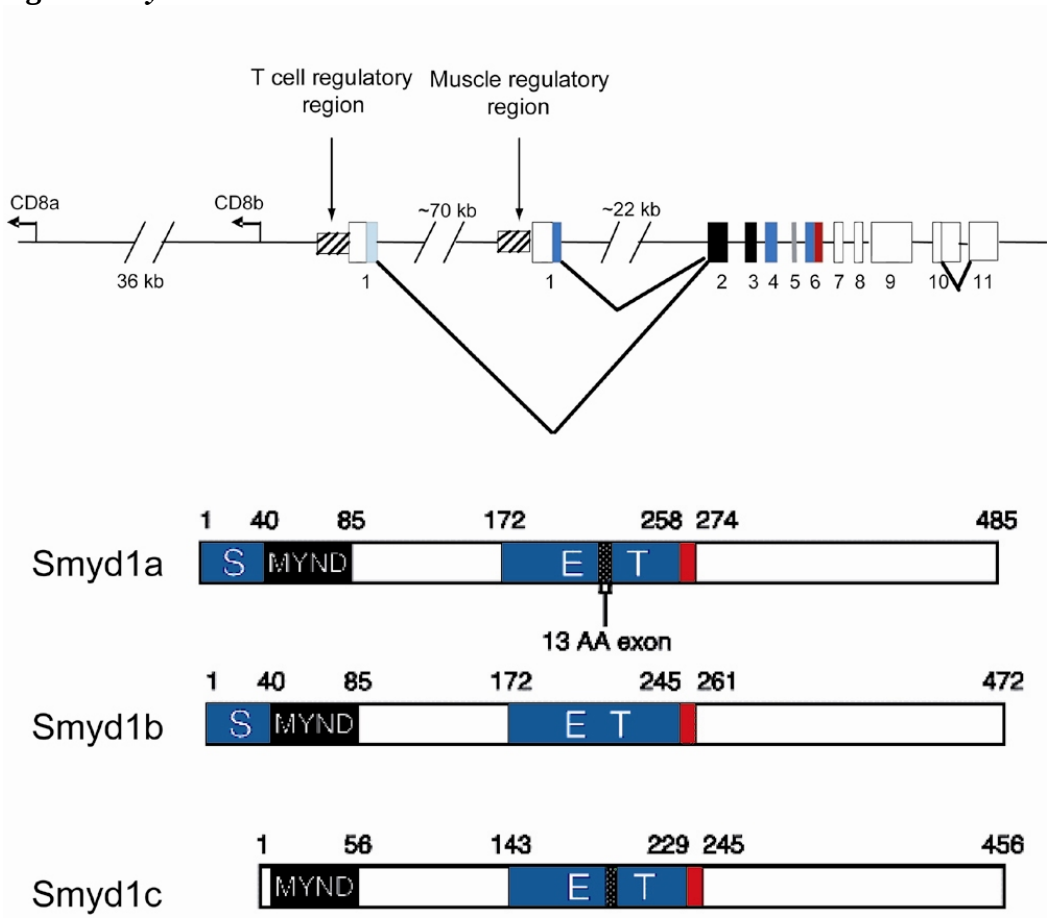


**Figure 1.1 *Smyd1* is a gene that encodes three isoforms.** The upper panel illustrates the genomic context of the *Smyd1* gene and the three characterized splice isoforms. The lower panel describes the three proteins. *Smyd1a* and *b* are expressed specifically in skeletal and cardiac muscle. *Smyd1c* has an alternative first exon and is expressed specifically in thymus. Both first exons are separated by a very large intron. The first exon for *Smyd1c* is directly adjacent to the *CD8β* gene. *Smyd1c* has a distinct 5' and 3' UTR.

## 1.2 Smyd Domains and their functions

*SET Domain*

Smyd defines a family of proteins that contain a split SET and a MYND domain. These domains are suggestive of Smyd function. Many SET domain containing proteins have been found to function as histone lysine methyl transferases, or HKMTases, such as *SUV39H1*, *G9a*, and *Clr4*. HKMTases can transfer one or more methyl groups onto specific lysine residues on histone tails. Five lysine residues have been discovered and characterized, including histone 3 lysine 4 (H3K4), histone 3 lysine 9 (H3K9), histone 3 lysine 27 (H3K27), histone 3 lysine 36 (H3K36), and histone 4 lysine 20 (H4K20). The histone core residue H3K79 has also been been characterized, although it is not present on histone tails. Methylation of H3K4, H3K36, and H3K79 are associated with transcriptional activation while methylation of H3K9 and H3K27 are associated with transcriptional repression [1]. H4K20 methylation has been implicated in DNA repair[1], DNA silencing [2], as well as the governing of the myogenic program [2].

*Smyd1* has been shown *in vitro* to methylate multiple histone substrates, including H3K4, a novel mark on H4, and at least one non-histone substrate, *skNAC* [3]. Several HKMTases have recently been shown to have non-histone targets. *Set 9*, which has long been known to methylate H3K4, *Smyd2*, which has recently been shown to methylate H3K36 [5] and H3K4 [4], and *Set8*, which is known to methylate H4K20 [5], have both recently been shown to methylate specific residues on *p53* [7-9]. However, it is still unusual for SET domains to have multiple histone residue and non-histone targets. Because the SET domain of Smyd proteins is interrupted by another domain, the specificity of the SET domain may be sacrificed causing Smyd proteins to have multiple specific histone targets or it may simply lead to an artificial promiscuity *in vitro*. Therefore, despite knowing 3 *in vitro* methylation targets for *Smyd1*, there are questions

as to which, if any, are true *in vivo* methylation targets and whether the most biologically significant methylation targets have been identified.

*MYND Domain*

The MYND domain has been very broadly implicated in protein-protein interactions through a PXLXP motif. *Smyd1* is most homologous to the MYND domain of *ETO* [10]. Like *ETO*, *Smyd1* binds directly to class I and class II HDACs, although this activity is independent of the MYND domain [6]. *Smyd1* associates with multiple repression complexes that include N-CoR and Sin3. Immunoprecipitation of N-CoR by *Smyd1* is dependent on the MYND domain [6]. *skNAC* is the most common *Smyd1* interacting partner pulled out of yeast 2-hybrid screens with both cardiac and skeletal muscle libraries [6, 7] and has been shown to bind specifically via a PXLXP motif to the MYND domain of *Smyd1* [8]. Through specific DNA binding at the promoter, *skNAC* has been reported to co-activate transcription of *Myoglobin*, a gene important for transporting and regulating oxygen levels in muscle tissue [9]. Other potentially interesting *Smyd1* interacting factors that bind through the MYND domain include *FKBP506(8)*, a regulator of *NFAT* and *calcineurin*; and *TRB3*,which is involved in insulin resistance [7].

*Protein Function*

*Gal4-Smyd1* fusion proteins repress transcriptional activity of a LexA-VP16 activated *Luciferase* reporter [10]. Repression activity is inhibited by mutations of the conserved cysteines in the MYND domain [6] or by addition of TSA, an HDAC inhibitor [10]. Although *Smyd1* is a transcriptional repressor when ectopically forced to bind DNA through *Gal-4* domains, it may act differently when it is recruited to DNA in the

endogenous context. When the reporter system is co-transfected with the *Smyd1-Gal4* fusion proteins and *EBNA2*, a PXLXP containing oncogene, *Smyd1b* activates expression of a *Luciferase* reporter (Personal Communication, R.J. Sims). While this experiment does not provide an endogenous context, it supports the hypothesis that *Smyd1* proteins will act differently in different contexts. Therefore, it is likely that when *Smyd1* is in specific complexes it acts to repress its target genes, but when found in other complexes, *Smyd1* may act to activate its target genes. This activity may be dependent on which histones are modified or which factors are bound to either the *Smyd1* protein or the DNA locus to which *Smyd1* is recruited. It may also be dependent on the specific expression patterns of co-activators or co-repressors. Dual roles could be part of a muscular checkpoint control mechanism; it could play a role in regulating chamber specific protein expression in the heart; or allow *Smyd1* to play different roles in temporal and spatial specific manners.

Little is known about the mechanism through which *Smyd1* functions, however, conserved SET and MYND domains, interaction studies, and reporter assays suggest that the 3 isoforms of *Smyd1* act as transcriptional factors by targeting histones or non-histone proteins for deacetylation and/or methylation. It seems likely that there are distinct roles and functional differences for *Smyd1a* and *Smyd1b*. In *Smyd1a*, the short 39 base pair exon 5 interrupts the C-terminal portion of the SET domain. Because *Smyd1b*, lacks exon 5 and more closely resembles a SET domain, most biochemical work thus far has focused on *Smyd1b*. However, when independently over expressed in C2C12, an immortal myoblast cell line, only *Smyd1b* methylates *skNAC* and effects *Myoglobin* transcript levels [3]. It is possible that these isoforms have differing roles in temporal and tissue specific contexts, but this has not been measured due to difficulty distinguishing both the transcripts and the proteins.

## 1.3 Important Players Regulating Cardiac and Skeletal Muscle Development and Disease

*Cardiac Development*

During development of all vertebrates, the heart is the first organ to undergo organogenesis and is essential for the proper development of all other organs. Every year 36,000 American infants, almost 1% of annual births, are expected to suffer from congenital cardiac defects [11]. Although almost as many people born with congenital cardiac defects survive at least until the age of 25, this group of birth defects leads to more infant deaths than any other birth defect [11].

The morphological changes in the developing heart are well understood. This is a complex process that requires precise timing of cell migration, myogenesis, morphogenesis and contractility. In murine heart development, the primary heart field forms a crescent during gastrulation that fuses into the linear heart tube at E8.0. At this point, the heart begins to beat and is essential for continued embryonic development. The proliferating secondary heart field migrates into the heart tube, causes it to loop to the right, and contributes to the right ventricle, inflow tract, and outflow tract [12-14]. The heart loop segments into atrial and ventricular chambers. After segmentation, the ventricular chambers must undergo compaction and trabeculation in order to maintain the necessary hemodynamic pressure while both atrial and ventricular chambers further septate to create the mature four chambered heart [15].

*Nkx2.5*, homologous to the drosophila gene *tinman,* and one of the earliest markers of cardiomyocyte populations, has been well studied and was the first putative regulator of the cardiogenic program [16]. However, while knockouts of *Nkx2.5* were striking, the effect was primarily on morphogenesis and the ventricular chambers [16]. It has recently been shown that *GATA-4* is sufficient to induce cardiomyogenesis [17] and

is hypothesized to work synergistically with *SMAD* proteins to activate *Nkx2.5* [18]. However, what is known about the molecular pathways governing cardiac development is not much more than a laundry list of critical transcription factors for distinct processes. The first known marker of cardiomyocytes is *MesP1*, which is first expressed E6.5 and is critical for proper cell migration into the cardiac crescent [19]. *Nkx2.5* is thought to play a role in regionalization, or chamber specification [20]. *Tbx5* and *Mlc2a* are essential for atrial development [21, 22]. *Mlc2v* specifically marks the ventricular myocardium [23]. *Isl1* is a known marker of the secondary heart field, which is thought to provide the majority of the cells for the right ventricle [24]. *Hand2*,or *dHand*, is essential for right ventricular development and *Hand1*, or *eHand*, primarily expressed in the left ventricle, is essential for cardiac looping [25]. Many different signaling pathways have been implicated in different compartmentalized aspects of cardiac development, including the *TGFβ/BMP* signaling cascade [12, 22, 26-29]. There are multitudes of individual genes known to play a role in cardiogenesis, but there is no clear consensus in the field on a hierarchical molecular pathway that regulates the complex process after the initial migration and fating of the early cardiac crescent.

Heterozygotes for *Smyd1* are completely viable and normal. However, homozygotes for the *Smyd1* knockout are embryonic lethal at E9.5, approximately halfway through gestation. At this stage, the null mutant has an enlarged left ventricular chamber with an excess of extracellular matrix (ECM) and is completely void of the right ventricle [10]. In the *Smyd1* null mice, expression of *Mlc2a*, *Mlc2v*, *Nkx 2.5*, and *Tbx5* transcripts were normal. However, *Hand2* expression is ablated and expression of the *Hand2* dependent gene, *Irx4*, is diminished [10].

*Skeletal Muscle Development*

Unlike the cardiac muscle, skeletal muscle is not necessary during development; skeletal muscle is necessary in order to breathe and feed at and after birth. In mice, skeletal muscle is also essential for locomotion necessary in order to compete with littermates during the nursing stage. Myogenic Regulatory Factors (MRFs), including *Myf5*, *Myf6/Mrf4*, *MyoD*, and *Myogenin (MyoG)*, have long been thought to direct and control muscle development. Surprisingly, however, single MRF mouse knockouts have had only marginal effects on myogenesis with one exception, the *MyoG* knockout, whose deficiency results in a normal number of myoblasts that are incapable of differentiating into myotubes [30, 31]. This argues that *MyoG* is essential for myogenic differentiation process.

Double knockouts of *MyoD* and *Myf5* have a large impact on myogenesis, suggesting that these two genes share redundant pathways and are overlapping in function. In this double knockout, mice are born, but fail to move or survive past birth. There are no detectable myoblasts or myoblast-like cells suggesting that *MyoD* and *Myf5* play a partially redundant role in cell fate determination and the commitment to a myogenic fate [32].

*Myf6* is expressed in a biphasic pattern including myotome restricted expression during embryonic day 9 through 11 and in all skeletal muscles from embryonic day 16 until after birth [33]. *Myf6* knockouts have not yielded straight forward data. However, it was discovered that alterations at the *Myf5* locus or *Myf6* locus have affects on both genes [34]. It has thereby been shown that the early phase of *Myf6* expression is important for skeletal muscle fate determination [34]. It is still thought that the second phase of *Myf6* expression has an important role in organizing and maintaining differentiated myotubes.

*Adult Cardiac Disease*

Even after development, the heart is one of the most vital organs. When the heart malfunctions other organs cannot continue normally. A malfunctioning heart's owner has decreased quality of life and life expectancy. One in three adults in the United States suffer from at least one type of cardiovascular disease ranging from hypertension to cardiac failure [11]. Therefore, the molecular signals that either induce cardiac disease or maintain healthy cardiac tissue are very intriguing subjects. Smyd1 has been correlated to cardiac ventricular diseases and to the maintenance of the ion gradients required for a normal heart beat.

A study performed on explanted diseased human hearts discovered that *Smyd1* is two or three fold over expressed in diseased left and right ventricles, respectively, compared to biopsies of the healthy transplanted hearts [35]. However, like many of the ion-channel genes analyzed in this study, expression of *Smyd1* was normal in explanted hearts that had been treated with a left ventricular assist device (LVAD) [35], which is a surgically implanted mechanical pump. This pump helps keep blood circulating, relieves hemodynamic pressure on the ventricle, and helps extend life expectancy for those on heart transplant waitlists. Therefore, *Smyd1* transcription is activated directly or indirectly by mechanical stress. *Smyd1* transcript levels return to normal in response to alleviation of that stress via LVAD implantation. However, it is unclear whether the role of *Smyd1* is to cause or overcome cardiac disease.

*Atrial and brain naturietic peptides* (*ANP* and *BNP*) are established markers for many cardiac diseases, particularly chronic heart failure and left ventricular dysfunction in human patients [36-39] and experimental rat models of myocardial infarction [36, 40]. Both are secreted in response to the mechanical stress of atrial or ventricular wall stretching [39]. *ANP* is thought to inhibit growth and proliferation [40] while *BNP* is

thought to induce myocardial relaxation [39]. Levels of *ANP* and *BNP* transcripts are directly correlated to the severity of heart failure [41]. Neither of these markers are returned to normal levels in LVAD treated patients [35].

In wildtype mice, *Kcnd2*, the gene encoding the Kv4.2 potassium channel α-subunit, is expressed differentially in the ventricular epicardium and endocardium resulting in a potassium gradient that is necessary for a stable heart beat. A recent study has shown that *Smyd1* and *Irx5* directly interact in order to regulate *Kcnd2* transcription [42]. When *Smyd1* is expressed in an *Irx5* null situation, *Kcnd2* is activated, but when Irx5 or Smyd1 are combined, activation of *Kcnd2* transcription is inhibited [42]. *Irx5* is endogenously expressed in an inverse gradient to *Kcnd2*, whereas *Smyd1* is expressed throughout the heart. This *Kcnd2* gradient could be caused and exaggerated if *Smyd1* activates *Kcnd2* in *Irx5* deficient zones and inhibits *Kcnd2* in *Irx5* positive zones. In mice null for *Irx5*, *Kcnd2* expression is equal in the ventricular epicardium and endocardium and tachycardia is more easily induced [42].

### 1.4 Rationale and Preview

In order to determine which pathways *Smyd1* is involved and to prove that *Smyd1a* and *b* were regulated by independent regulatory regions from *Smyd1c*, we analyzed the putative regulatory region for functionality and tissue specificity by using transgenic promoter-reporter studies. During the course of this study, I have analyzed the putative *Smyd1* regulatory regions for candidates that likely regulate *Smyd1* transcription. Some of these have been tested *in vivo* by using transgenic mutant promoter-reporter technologies. The others have been chronicled for use in later studies.

To determine what role *Smyd1* plays throughout cardiac development, during skeletal muscle development, and in adult maintenance of the two tissues, I have induced

the *Smyd1* knockout in each system independently using *Cre*-LoxP technologies. I have analyzed the phenotype in each tissue to determine likely pathways and processes with which *Smyd1* may be involved.

The data is reported in four chapters. Chapter 3 describes work done on the *Smyd1* promoter. Chapter 4 describes the three embryonic heart specific conditional knockouts. The fifth chapter discusses the embryonic skeletal muscle conditional knockout. Adult conditional knockouts were discussed in chapter six. Observations detailing the presence of novel alternative splice forms and the potential impact on previous data were discussed in a short seventh chapter. Conclusions drawn from each section were discussed within the relevant chapters. A discussion chapter was included to address concepts that were not contained within one specific chapter.

# Chapter 2. Materials and Methods

## 2.1 Computational Promoter Analysis

The regions immediately upstream of the *Smyd1* coding region from the mouse and human genomes were identified by blasting the coding sequences obtained from the National Center for BioInformatics (NCBI) for human and mouse *Smyd1* against the respective genome databases. The 3kb regions 5' to the first exon of the skeletal and cardiac muscle *Smyd1* isoform were identified in each genome, resulting in 2 sequences. These sequence were compared using 2-sequence blast program (NCBI) and Genomatix promoter software to identify conserved DNA regions and conserved protein binding sites that may affect transcriptional regulation of *Smyd1*.

## 2.2 Promoter Constructs

In order to obtain wildtype sequence from the *Smyd1* locus, Incyte Genomic, Inc. screened BAC libraries and provided three positive clones. One of these resulting clones, BAC 26242, contained the entire *Smyd1* locus. This was digested using XbaI and the resulting fragments were ligated into pBlueScript. From this XbaI library, clone BAC42-Xba9 was isolated and used as a template for creating wildtype promoter-reporter vectors for mouse transgenic expression.

The 0.5kb promoter construct was created by ligating a PCR product made using Vent polymerase, the BAC42-Xba9 template, and the following primers: mBOPPROF ctcagagtctgtcaggccttcg and mBop 0.5kb Rev ggagacactagggtgggg into the SmaI site of pBSSKAUG-bGalXSal, obtained from Ellen Lien and created by Doris Brown. The 1kb promoter construct was created in the same manner, except primer mBop 0.5kb Rev was replaced with the primer Bop Xba-9 Rev: gaactgtgagggtcacactc. The 3kb promoter

construct was created by digesting BAC42-Xba9 with BamHI and inserting the resulting 3.2kb fragment into BamHI digested 0.5kb construct and screening for orientation.

Mutations were created by PCR mutagenesis using the 1kb promoter as a template for the following sites: Myf5 c436a, a435g; Mef2 t399g, g404c; MyoD t733c; Mef2 a999g, a998g. These were all created by following the protocol provided with the Gene Editor Site Directed Mutagenesis Kit (Promega) and confirmed by DNA sequencing.

### 2.3 Transgenic Mice

Constructs were digested with XhoI and NotI to create a linear fragment of 4.9kb, 5.4kb, and 7.8kb, respectively. The wild type linear fragments were cleaned by passing over a Qiagen Miniprep column and injected into fertilized oocytes of B6C3F1 female mice. These were implanted into pseudopregnant ICR mice as previously described [43]. 24 founding lines were created.

### 2.4 X-gal Staining

Transgenic male mice were timed mated, as described below, to wildtype DB2 females. At various ages of development (8.5 through 15.5dpc), the embryos were dissected and LacZ expression was detected by staining with X-gal, as previously described [44]. Stained embryos were fixed in 4% paraformaldehyde at 4$^{o}$C overnight prior to paraffin sectioning, as previously described [44].

### 2.5 *Smyd1* Flox Vector and Targeting

Endogenous sequence was obtained from a Lambda FixII Vector 129SV Mouse Genomic Library (Stratagene). The targeting construct was designed in the Osdupdel vector, a gift from William A Kuziel, in order to introduce LoxP sites flanking the second

and the third exons of *Smyd1* followed by a neomycin cassette that is flanked by a third LoxP site. The targeting vector was electroporated into 129S6 ES cells. After selecting with G418 and gancyclovir, surviving clones were screened for homologous recombination by Southern analysis after digestion either with Bgl II and Sal I or Bgl II and Kpn I, and hybridization with probe 2 and probe 3, respectively.

ES clones showing correct targeting in both arms were injected into C57BL/6 blastocysts to create chimeric mice. The chimeric mice were mated to C57BL/6 females to create a germline knock-in referred to as $Smyd1^{KI}$. $Smyd1^{KI}$ mice, B6.129 hybrids, were crossed to ubiquitously expressing *EIIa-Cre* mice, B6.FVB hybrids, and progeny were screened for deletion of the neomycin cassette by Southern analysis and back-crossed in order to remove *EIIa-Cre*. Resulting $Smyd1^{Flox}$ mice lacking *EIIa-Cre* were crossed with several transgenic *Cre* lines including *Nkx2.5-Cre*, *MyoG-cre*, *aMHCcre*, *αMHC-CrePR1,* and *MLC-rtTA*, and *Tet-ON-Cre*.

**2.6 PCR genotyping**

At weaning, each mouse was given a unique ear punch and a short piece of tail (2-5mm) was biopsied. The tail tissue was cooked overnight at 55°C to 60°C in 200ul Tail PCR Buffer (50mM KCl, 10mM Tris (pH8.3), 2.5mM MgCl2, 0.1mg/ml gelatin, 0.45% v/v NP40, 0.45% v/v Tween 20) plus 2ul Proteinase K (10mg/mL), boiled for 5 minutes and centrifuged for 10 minutes. 3-5ul of this solution was used in each genotypic PCR reaction. PCR conditions were as follows: step 1) 94°C 5min, Step 2) 94°C 30s, Step 3) 62°C for 1min, Step 4) 72°C 1min, Step 5) Go To step 2- 29 times, Step 6) 72 °C 7 min, Step 7) 4°C for ever. The primers KO7S: cacatctttggtgtggtatggc and KO8AS: ctcacttgcgtcccagtacttg detect $Smyd1^{WT}$ (540bp) and $Smyd1^{Flox}$ (620bp). Alternatively, primers KO5AS: gcatacgcacatgtgctcgc and KO4S: tcatgagatgggcatgagcc detect $Smyd1^{WT}$

(432 bp) and *Smyd1*$^{Flox}$ (552 bp) (Fig. 2.1). Genotyping primers KO4S: tcatgagatgggcatgagcc and WAK65: gcccggttctttttgtcaagaccga detect *Smyd1*$^{KO}$ and yield a product predicted at 985bp, but that runs at approximately 850bp. Genotyping primers for *Cre* give a product at 268bp and are Cre-1: ggacatgttcagggatcgccaggcg and Cre2: gcataaccagtgaaacagcattgctg.

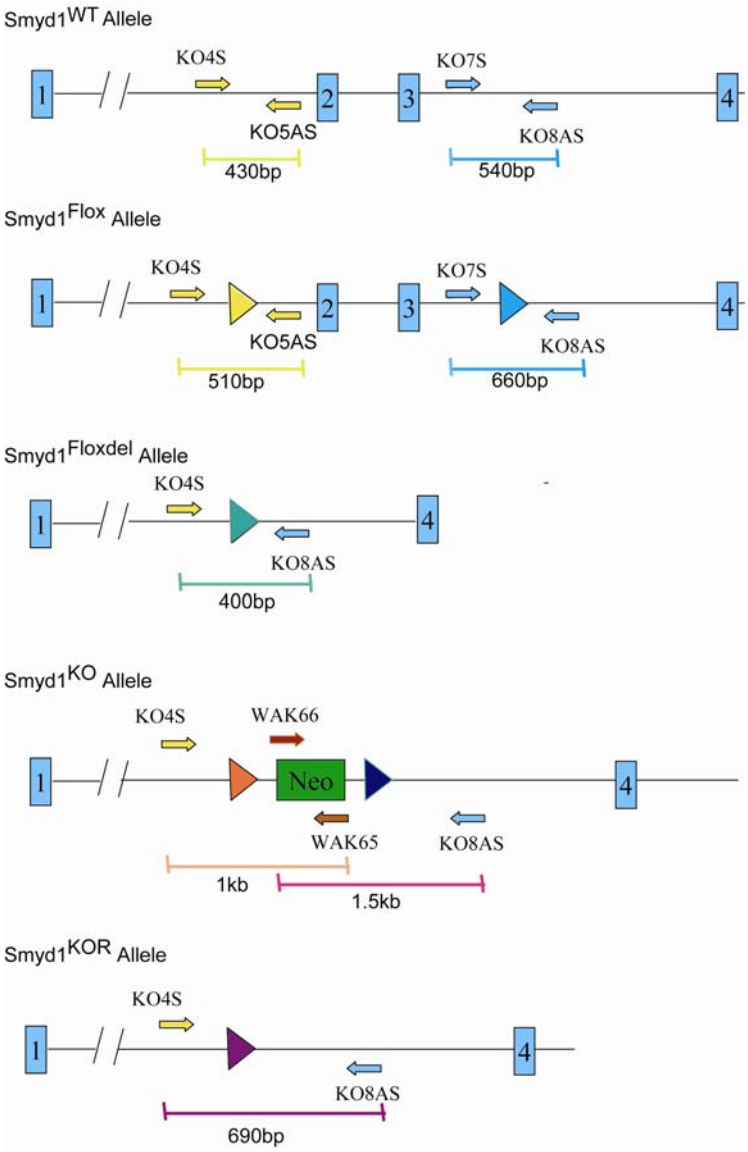


**Fig 2.1 PCR Genotyping Schematic.** Primers have been designed to flank the targeted knock-in LoxP sites in a manner that distinguishes several *Smyd1* alleles. KO5AS/KO4S and KO8AS/KO7S are two primer pairs that detect $Smyd1^{WT}$ and $Smyd1^{Flox}$ at varying sizes. Primers internal to the LoxP sites have been deleted from $Smyd1^{Floxdel}$, $Smyd1^{KO}$, and $Smyd1^{KOR}$. The external primers can be used to detect the recombination products, $Smyd1^{Floxdel}$ and $Smyd1^{KOR}$. Primers internal to the neomycin cassette are paired with endogenous primers to detect the $Smyd1^{KO}$allele.

## 2.7 Recombination Assays

Recombination in each cross was determined by PCR genotyping to determine tissue specificity and efficiency of deletion. Heterozygous mice (*Smyd1*$^{Flox/WT}$; XX-Cre$^{+/?}$) were sacrificed and tissues were harvested into eppendorf tubes. Tissue samples were incubated at 55-60$^{o}$C in 430ul Southern Tail PCR Buffer (50mM Tris pH 8.0, 100mM EDTA, 100mM NaCl, 1%SDS) and 70ul proteinase K (10mg/ml). After complete digestion of the tissue (16-48 hours), samples were phenol:chloroform extracted. Briefly, 500ul of phenol was added to each sample, vortexed for 30 seconds and centrifuged at full speed in a microcentrifuge for 5 minutes. The upper aqueous phase was removed and the above procedure was repeated with Phenol:isoamyl:Chloroform (25:1:24) and Chloroform. The third aqueous phase was removed and precipitated with 166ul of 30% PEG/1.5MNaCl at -80$^{o}$C for 15 minutes. After precipitation the samples were centrifuged at full speed in a microcentrifuge for 15 minutes, the supernatant was discarded, and the pellet was washed in 70% ethanol. This pellet was then dried and resuspended in 100ul of water at 37$^{o}$C.

0.1 to 5ul of sample were used for genomic PCR using the following program: step 1) 94$^{o}$C 5', Step 2) 94$^{o}$C 30s, Step 3) 62$^{o}$C for 1min, Step 4) 72$^{o}$C 1min, Step 5) Go To step 2- 29 times, Step 6) 72$^{o}$C 7 min, Step 7) 4$^{o}$C for ever. The primers KO7S: cacatctttggtgtggtatggc and KO8AS: ctcacttgcgtcccagtacttg detect *Smyd1*$^{WT}$ (540bp) and *Smyd1*$^{Flox}$ (620bp) and the primers KO4S: tcatgagatgggcatgagcc and KO5AS: gcatacgcacatgtgctcgc detect *Smyd1*$^{WT}$ (432 bp) and *Smyd1*$^{Flox}$ (552 bp). KO8AS and KO4S detect *Smyd1*$^{Floxdel}$ (395bp).

## 2.8 Timed Matings and Mouse Husbandry

Originally, mice were mated with the following non-gender specific scheme, *Smyd1*$^{KO/Flox}$ by *Smyd1*$^{Flox/WT}$; XX-*Cre*$^{+/-}$ to yield *Smyd1*$^{KO/Flox}$; XX-*Cre*$^{+/-}$. It was later determined that the following scheme was equally effective: *Smyd1*$^{Flox/WT}$; XX-*Cre*$^{+/-}$ x *Smyd1*$^{Flox/Flox}$ to obtain *Smyd1*$^{Flox/Flox}$; XX-*Cre*$^{+/-}$. At weaning, mice were given a unique ear punch and approximately 5mm of tail was biopsied for genotyping.

In order to determine phenotypes and time of embryonic lethality, female mice were housed with male mice. Females were checked for plugs daily. The observation of a plug was considered day 0.5. Females were sacrificed and ovaries were removed on the embryonic day of interest. Ovaries were placed in Phosphate Buffered Solution (PBS: 137mM NaCl, 2.7mM KCl, 10mM $Na_2HPO_4$, 2mm $KH_2PO_4$ (pH 7.4)) or DEPC treated PBS immediately after removal. Embryos were removed from the ovaries and dissected in PBS or DEPC PBS. Ages were confirmed by comparing to Rugh's Organogeny [45]. As well as harvesting tissues of interest, a piece of embryonic tail was harvested for genotyping. Embryonic genotyping was carried out in the same manner as described above, with the exception that tails were incubated in Tail PCR Buffer at 55-60$^{o}$C for 3 to 16 hours.

## 2.9 Drug Administration

*RU486*

RU486, or Mifepristone (Sigma), was suspended at the appropriate concentration in corn oil. This process was aided by heat and rotation in a Bambino hybridization oven. Doses of 1.25 to 10mg in a volume of 0.1 to 0.3ccs were injected intraperitoneally into adult mice daily for five sequential days.

*Doxycycline*

300mg of Doxycycline hydrochloride (Sigma) was mixed with 50g of ground mouse feed and 2.5g sucrose. Approximately 50-60mL of deionized water was added. The food was rolled into pellets and kept in the dark at 4$^{o}$C for up to a week. Pellets were also made with only ground mouse feed, sucrose, and water for use as a negative control. Mice were given prepared food for 3 to 5 days with no alternative food sources.

## 2.10 Tissue Harvest and Preparations for Paraffin Sections

*RNA Competancy/In Situ Hybridization*

Embryos at 15.5dpc or younger were extracted in DEPC PBS (1mLDEPC/L, autoclaved). A piece of tail was removed for genotyping. The remainder of the embryo was fixed in at least 10 fold volumes of 4% paraformaldehyde at 4$^{o}$C on a nutator overnight. They were then washed 3 times for at least 5 minutes in DEPC PBS and stored at 4$^{o}$C until shipped on wet ice to UT-Southwestern or UT-Science Park for Paraffin embedding and tissue section as previously described [43].

*Protein Competancy/Immunohistochemistry*

Embryos and adult tissues (heart) were extracted in PBS and fixed in at least 10 fold volumes of 10% formalin or 4% paraformaldehyde at room temperature on a nutator for 24-72 hours. Embryos older than 15.5dpc were skinned before fixation. A piece of tail was removed before fixation and kept for genotyping. All samples were washed 3 times for at least 1 hour in PBS and stored at 4$^{o}$C until shipped on wet ice to UTSouthwestern or UT Science Park for Paraffin Embedding and tissue section as previously described [43].

## 2.11 Tissue Harvest for RNA Extraction

*Embryonic tissue*

Embryos were extracted as above. For skeletal muscle, limbs were cut off at the shoulder and hip joints. The wrists, ankles, and digits were removed with forceps. The skin was removed from embryos older than 15dpc. The tongue was removed by opening the jaw and pinching off the tongue with forceps. Heart tissue was removed through the chest cavity. Isolated tissues were immediately added to 100ul Trizol and dounced manually in an eppendorf tube. After douncing, the volume was adjusted to 800ul. A piece of tail was removed and kept for genotyping. Embryonic tails were cooked at 60$^{o}$C for 3 hours and genotyped as previously described. Trizol homogenates were stored at -80$^{o}$C indefinitely.

*Adult Tissue*

Whole tissue was extracted from the mouse at various time points and immediately added to 3mL trizol in a round bottomed polystyrene culture tube. A piece of tail was removed and used to confirm the genotype. Tissues were then homogenized with a polytron homogenizer at setting 22 for 1 minute at room temperature. When the left and right ventricles were isolated, the left ventricle was separated from the whole heart with scissors. The left ventricle and right ventricle plus septum were independently manually homogenized in 200ul of trizol. After homogenization 700ul of trizol was added. All trizol homogenates were stored at -80$^{o}$C indefinitely.

## 2.12 *skNAC* Knockout

A targeting construct was created to replace the muscle-specific exon 2 with an FRT-flanked PGK-Neo cassette. Genomic fragments surrounding exon 2 of *skNAC* were

isolated by SacI digestion (6.75 kb) for the 5′ arm and XbaI digestion (1.3 kb) for the 3′ arm and subcloned into a vector containing a PGK-Neo cassette. The targeting construct was linearized with AdhI and electroporated into 129S6 ES cells. Targeted ES cells were selected with G418, and correct targeting was confirmed by Southern hybridization after HindIII and HpaI digestion with the corresponding 5′ and 3′ probes, respectively. Two correctly targeted ES cell clones were injected into C57BL/6 blastocysts to generate chimeric mice, which were bred with C57BL/6 or 129S6 mice. The offspring were genotyped by PCR. RT-PCR was done to test deletion of the second exon and preservation of *αNAC* expression in *skNAC–/–* embryos. Primers for the *skNAC* transcript were 5′-gttactgacccagcaatggctcct-3′ and 5′-tggaggagtgggagaaagcttattaag-3′. Primers for *αNAC* were 5′-cacccggtgaagccacagaaaccgtc-3′ and 5′-aactttgaatttctcagcagctgc-3′.

### 2.13 Microarrays

E11.5 hearts of wild type and *skNAC–/–* embryos were collected from timed pregnancies and three embryos for each genotype were processed individually for RNA isolation with Trizol. RNA was amplified and labeled with MessageAmp II-Biotin Enhanced Kit (Ambion). Affymetrix GeneChip arrays (mouse genome 430 2.0) were used for hybridization in triplicate. Hybridization quality was assessed with the affyPLM package (Bioconductor). Preprocessing was performed with the RMA algorithm. Log-ratios were calculated with a linear model (Limma Package, Bioconductor).

### 2.14 RT-PCR

Trizol homogenates were thawed and RNA was isolated following Invitrogen's Trizol protocol, including the optional centrifugation to remove soluble material following thaw. For embryonic samples, the isopropanol precipitation was supplemented

with 5-10ug of RNAse free glycogen. After drying the RNA pellet, samples were resuspended in 25ul of deionized DEPC water and immediately DNase treated by adding 15ul mix D (2.7x DNaseI buffer (Invitrogen), 1U/ul RNaseOut (Invitrogen), 2U/ul DNaseI (Roche)) to a final volume of 40ul and incubating at 37$^{o}$C for 1 hour. The enzyme was heat inactivated by adding 4ul of 25mM EDTA and incubating at 70$^{o}$C for 10 minutes.

RNA quality and concentration was analyzed using a nanodrop spectrophotometer. The same mass of total RNA (1-10ug) was used in all comparable samples. 5ul mix 1 (40uM oligo dT, 5.5uM dNTP) and DEPC water were combined with the total RNA sample to an intermediate volume of 35ul, except for the negative control, which was adjusted to 37ul. Samples were boiled for 5 minutes and iced for 5 minutes before adding 13ul of mix 2 (3.8x 1st Strand Buffer, 20mM DTT, 1U/ul RNaseOut (Invitrogen)). After prewarming the reaction mixture for 2 minutes at 37$^{o}$C, 2ul of 200U/ul MMLV (Invitrogen) was added to each reaction except the negative control. Samples reverse transcribed for 1.5 hours at 37$^{o}$C. The enzyme was heat inactivated at 85$^{o}$C for 5 minutes and the samples were stored at 4$^{o}$C.

1-2ul of the cDNA reaction were used for PCR with various gene primer sets. Reduction of *Smyd1* transcripts in *Smyd1*$^{Flox/Flox}$ or *Smyd1*$^{Flox/KO}$ with *Nkx2.5-Cre*, *MyoG-cre*, *αMHC-CrePR1*, and *MLC-rtTA* and *Tet-ON-Cre* was confirmed by RT-PCR. RT-PCR Primers most widely used for *Smyd1* are 47660: ctgcgggcagtgcaagtt and Smyd1e6R: gctcacaggagcagtcaaagtag. RT-PCR Primers for *β-actin* are β-actF: tacgagggctatgctctc and β-actR: cgcagctcagtaacagtc. Other primers were designed to cross introns and are listed along with thermocycling conditions in Appendix 2.

## Chapter 3. Transcriptional Regulation of *Smyd1* Expression

Transcriptional control can be regulated in concert with several genes at a given locus by affecting the chromatin, e.g. histone modifications that induce chromatin to be more or less accessible to polymerases and transcription factors. Control can also be mastered at the level of individual genes by sequence dependent binding of transcription factors to DNA. Because there is a 70kb intron between the thymus specific *Smyd1* first exon, exon 1a, and the muscle specific *Smyd1* first exon, exon 1b, we predicted that this expression pattern was derived from dual promoters. The thymus specific first exon is only approximately 200 base pairs from *CD8β* coding sequence. Since both of these transcripts share thymocyte specific expression, it is likely that factors activating and repressing *Smyd1* transcription share the same binding sites and interacting partners with the *CD8β* locus. However, in order to hypothesize what specifically affects the transcription of the muscle specific *Smyd1* isoforms, the putative promoter sequence, a region immediately upstream of exon 1b, was analyzed for putative transcription factor binding sites using the genomatix software suite (www.genomatix.de), specifically, the MatInspector transcription factor database (see appendix 1) [46].

*Homology in non-coding sequences between Mouse and Human: Region 1*

The described region was then further analyzed by comparing the region in the mouse and human genomes (FIG 3.1). These two genomes display a surprising amount of conservation in the noncoding region 600 bp upstream of the first exon of *Smyd1a* and *b*. In this region interesting and conserved putative transcription factor binding sites include: MTBF, TATA, *Zbp-89/Maz/Sp1*, *Tcf11*, *Meis-1*, and *Nkx2.5*. Also, there is a

large cluster of conserved sequence and sites that include: *MyoD/Myf5*, *Pbx-1/Meis-1*, *Tgif*, *Comp1*, *Mef2*, and MTBF.

*MTBF: MyoTube Binding Factor*

MTBF is a sequence commonly found in promoters specific for myotube expression and was originally found between *MyoD* and *Mef2* binding sites of the *Desmin* promoter. This site was found to be necessary for maximal myotube expression of *Desmin*, a protein expressed specifically in skeletal muscle and heart [47]. It is unknown what protein(s) binds to this sequence.

*TATA Box:*

The TATA Box is a common element in a class II promoter and is usually found 25 base pairs upstream of the transcription initiation site.

*Zbp-89/Maz/Sp1*

*Zbp-89* is a ubiquitously expressed protein that has several functions, including DNA binding transcription factor. *Zbp-89* is thought to be capable of either transcriptional activation or repression. There are at least 2 examples in which *Zbp-89* acts as a repressor via inhibition of *Sp1*, either through binding site competition or by direct interactions [48]. There are 2 conserved *Zbp-89* sites in the putative muscle *Smyd1* promoters in mouse and human. One of these sites overlaps a putative *Maz* binding site and the other a putative *Sp1* binding site. *Maz* and *Sp1* are C2H2-type zinc finger containing proteins that are known to bind specifically to GC-rich DNA sequences. The consensus site for *Maz* and *Sp1* are very similar and, in some cases, are synonomous. Both of these proteins are

thought to have repression and activation activities [49]. Hypothetically, *Zbp-89* may compete with *Maz* as it does with *Sp1* in order to repress transcription.

*Tcf11*

*Tcf11* is a widely expressed transcriptional activator. The activation activity of *Tcf11* has a dose-dependent and inverse relationship with *MafG* [50].

*Meis-1*

*Meis-1* is a homeodomain protein that is a known cofactor for Hox transcriptional activation. This cooperation leads to increased affinity and specificity for the appropriate transcriptional targets [51].

*Nkx2-5*

*Nkx2-5* is a pivotal transcription factor in orchestrating cardiac development. It contains a homeodomain and binds directly to DNA.

*MyoD/Myf5*

*MyoD* and *Myf5* are Myogenic Regulatory Factors (MRFs) that bind to DNA at E-boxes to induce the myogenic program. They are considered to be "master regulators" of myogenesis and, while they are at least partially redundant of each other, they are essential for myoblast determination.

*Pbx-1/Meis-1*

*Pbx-1/Meis-1* is a binding site for a heterodimer of the two proteins. This heterodimer binds DNA cooperatively with all *E2a*-MRF (*MyoG*, *MyoD*, *Myf5*,

*Myf6*) heterodimers [52]. Not only does this heterodimer enhance binding of the myogenic regulatory factors to E-boxes, but it is essential for binding to non-cononical E-boxes, such as those within the *Myogenin* promoter [53].

*Tgif: 5'TG3' Interacting Factor*

*Tgif* is a homeobox protein in the TALE superclass. It has recently been shown to repress transcription via a specific interaction with RXRα and the retinoid response element (RRE) [54].

*Comp1: Cooperates with Myogenic Proteins 1*

*Comp1* is a protein known to bind both DNA and *Myogenin* and enhance the interaction between *Myogenin* and DNA containing E-box sequences [55].

*Mef2: Myocyte-specific Enhancer binding Factor 2*

*Mef2* was initially characterized in terms of the role in myogenesis. It has since been discovered to play an important role in a wide variety of tissues. Vertebrate genomes encode four *Mef2* genes, A through D. All four genes share a very conserved MADS domain at the N-terminus that mediates DNA binding and dimerization. *Mef2* proteins have a divergent region responsible for interactions with cofactors and a C-terminal transactivation domain [56].

## Fig 3.1 Human and Mouse *Smyd1* Regulatory Region Alignment

SMYD1 promoter region sequence comparison; mouse to human genome; Hits on chr2:88659962-88660712; Berkley Genome Pipeline; Bold conserved regions based on 2 sequence Blast; +1 based on: NM_009762.1 (Mm); NM_198274 (Hs)

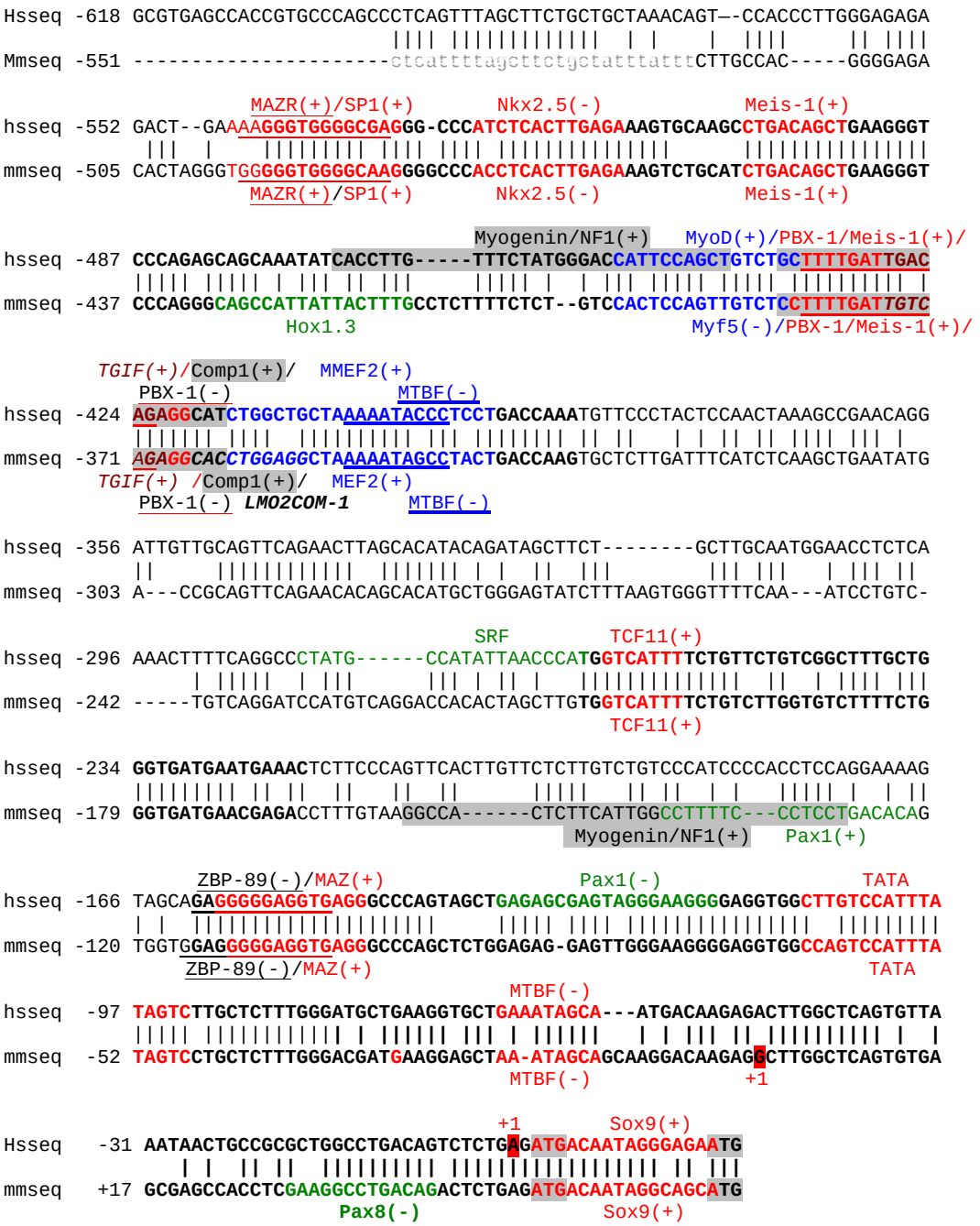


**Fig 3.1 Regulation of *Smyd1* must be highly conserved as the non-coding regulatory region has been conserved in mammalian evolution.** Matching base pairs are indicated by a |, mismatches are signified with a space, and putative binding sites are annotated in matching text. Regions in bold are the most highly conserved at the sequence level.

*Homology in non-coding sequences between mouse and human: region 1*

When extending the 600bp region to 1kb for analysis, there is added DNA homology after an approximately 300 bp insert into the human sequence. This region in both species encodes many of the same putative transcription factor binding sites. However, these are not conserved at the level of DNA sequence or proximity to the coding start site. These putative sites include *Srf*, *Myogenin*, *Oct1*, and *Bcl6*. There are also duplicates of the following *TATA*, *Pbx/Meis*, *Maz,* and *MTBF*.

*SRF: Serum Response Factor*

*Srf* is an ancient transcriptional activator with a MADS DNA interacting domain. *Srf* is thought to be a scaffolding platform that mediates expression of entire gene programs and heightened expression coincides with striation of skeletal and cardiac muscle [57].

*Myogenin*

*Myogenin*, or *MyoG,* is a MRF, which are thought to be "master regulators" of the myogenic program. *Myogenin* is known to play an essential role in myoblast to myotube differentiation.

*Oct1*

*Oct1* is an ubiquitously expressed transcription factor and has a general affect on transcription as well as effects on specific gene expression [58].

*Bcl6*

*Bcl6* is most highly studied in B-cells and T-cells, but is moderately expressed in most tissues. *Bcl6* is a known transcriptional repressor [59].

Extending sequence analysis to approximately 3kb upstream of *Smyd1* in mouse and human genomes leads to very little DNA homology that is not already mentioned.

*Homology with other skeletal and cardiac muscle promoters*

Because there were so many interesting putative sites, I attempted to narrow down candidates by extending the analysis to include other genes with similar expression patterns. *Desmin* is an intermediate filament that is expressed in both the heart and skeletal muscle [60]. *Hop*, on the other hand, is expressed only in the heart [61]. Comparing potential binding sites from mm*Smyd1*, hs*SMYD1* and mm*Desmin* putative regulatory sequences yields four conserved binding site clusters. While the same factors are present in the putative *Hop* regulatory region, they are not clustered in the same subgroups, suggesting that this pattern is important for skeletal muscle specific expression, but not for cardiac specific expression (Fig 3.2).

The second, and most interesting, cluster contains *Sp1*/GC box elements, *MazR*, *Meis1*, *Mef 2* and *MyoD* or *Myf5* binding sites. This cluster is very likely essential for skeletal muscle expression. Although the *Pbx* binding site does not register on the 4 sequence analysis, both putative *Smyd1* promoters contain both *Meis1* and *Pbx* adjacent to the MRF site, *MyoD* or *Myf5*. As mentioned above, *PBX/Meis1* heterodimers bind cooperatively with MRF heterodimers. Although the particular MRF is not identical in all four promoters, the varying binding proteins, *Myf5* and *MyoD*, are thought to have overlapping functions [62]. *Mef2* is also an important transcriptional activator for the myogenic program. While the *SP1* and *Maz* elements are not accompanied with a *Zbp-89*

site, they are represented in three promoters with similar expression, suggesting that they play a role independent of *Zbp-89*.

The first observed cluster includes a TATA box and an MTBF binding site. Although this TATA box is not located near the canonical position, -25, it is possible that the true or alternative transcription start site has not yet been identified. The third cluster contains *Stat* and *Maz* and the fourth cluster contains, *Evi1*, TATA, and *Stat*.

*Stat: Signal Transducers and Activators of Transcription*

Seven members of the Stat family have been identified in both the human and the mouse genome. Stat family members can receive a signal at the cell membrane and activate transcription without the use of intermediate players. They are often regulated by phosphorylation [63].

*Evi1: Ectopic Viral Integration Site-1*

*Evi1* was first described in Leukemia studies. Viral integrants into the host genome induced activation of *Evi1*, which caused leukemia. While *Evi1* is minimally or unexpressed in neonatal or adult skeletal muscle, it is expressed ubiquitously at embryonic day 9.5 and is present in the developing limb bud and heart. *Evi1* binds DNA in a sequence specific manner and can activate transcription through several zinc finger domains. *Evi1* associates with co-activators and co-repressors [64].

Based upon the above analysis, I created constructs that used 3kb, 1kb, and 0.5kb directly upstream of the transcription initiation and cloned these putative regulatory domains into a vector that contained a promoterless AUG-β-gal. The 1kb and 3kb constructs were

used to create transgenic mice. While 1kb is sufficient to drive expression of *Smyd1* in a semi-endogenous manner, the 3kb construct gave the same pattern, albeit stronger LacZ expression. Transgenic mice created from either construct failed to express the reporter in the left ventricle (Fig. 3.3). This aberrant expression is not unusual in transgenic reporters. Like *Smyd1*, many other genes are endogenously expressed in both the left and right ventricles but transgenic attempts to fuse large segments of the promoter to a LacZ reporter have failed to produce expression in the left ventricle. However, directed recombination of lacZ into the endogenous site of these genes does express the reporter equally in both chambers, suggesting that there is a yet unknown, but non-unique element in left ventricle expression of many cardiac genes, including *Hop* [61], *Mlc2v* [23] and *Nkx2.5* [65].

I also created mutants in the proximal and distal *Mef2* sites, the *Myf5* and *MyoD* sites (Fig. 3.4), although these constructs were never used to create transgenic mice or analyze the effects of the mutated sites on transcription. Meanwhile, collaborators in Eric Olsen's lab at UT Southwestern established similar transgenics using a vector that contained *hsp68-LacZ*, which provides a basal promoter fused to the reporter gene (Fig. 3.5). These transgenic mice expressed LacZ in the same pattern as those with simply the initiation codon, but at a much higher level (Fig. 3.6). This set of constructs also displayed higher expression when 3kb of regulatory region was fused to the basal promoter, but 0.6kb was sufficient for expression (Fig. 3.7).

Because *Mef2c* is known to play a definitive role in both cardiac and skeletal muscle myogenic program and conserved putative binding sites were present, we hypothesized that *Mef2c* binds directly to the *Smyd1* promoter in order to act as a transcription factor. In fact, when electromobility shift assays were done with *MEF2c* protein on the wildtype construct, we detected a supershift, which supports this

hypothesis (Fig. 3.8A). A mutation was made in the putative *Mef2* site on the hsp-90 basal promoter background and this construct was negative for *Mef2c* binding (Fig. 3.8B). Mutating the *Mef2* site also ablated expression in the cardiac tissue when transgenics were created, but skeletal muscle expression was intact (Fig 3.8C).

From sequence analysis, I predicted that this *Mef2* site is important in skeletal muscle expression due to motif conservation with mouse and human *Smyd1* muscle regulatory factors as well as that with *Desmin*. However, there is a second, more distal *Mef2* site in the mouse putative promoter that may compensate for the lack of the mutated *Mef2* site. While 0.6kb of the promoter is sufficient for driving expression, expression is heightened when 1kb, which includes the distal *Mef2* site, is available. I predict that the distal *Mef2* site is important for skeletal muscle expression in the endogenous context. In fact, a *Mef2c* conditional knockout was recently published from which it was shown that *Smyd1* is knocked down two-fold in the skeletal muscle of a *MyoG-Cre* induced *Mef2c* knockout [66].

We also made mutations of 3 E-boxes, DNA motifs known to bind to MRF hetero and homodimers, in the background of the 600bp-hsp90 construct. Only one of these E-boxes was recognized in the genomatix computational approach as a MRF binding site, *MyoD*. While all three core E-box sequences are present, only one of the E-boxes is recognizable as an MRF binding site by genomatix software. Because single mutations were not created, we cannot predict the role of each individual E-box sites. When all three E-box sequences are mutated, cardiac expression of the reporter is maintained, but skeletal muscle expression is lost in the context of the 600bp-hsp90 basal promoter construct (Fig. 3.9).

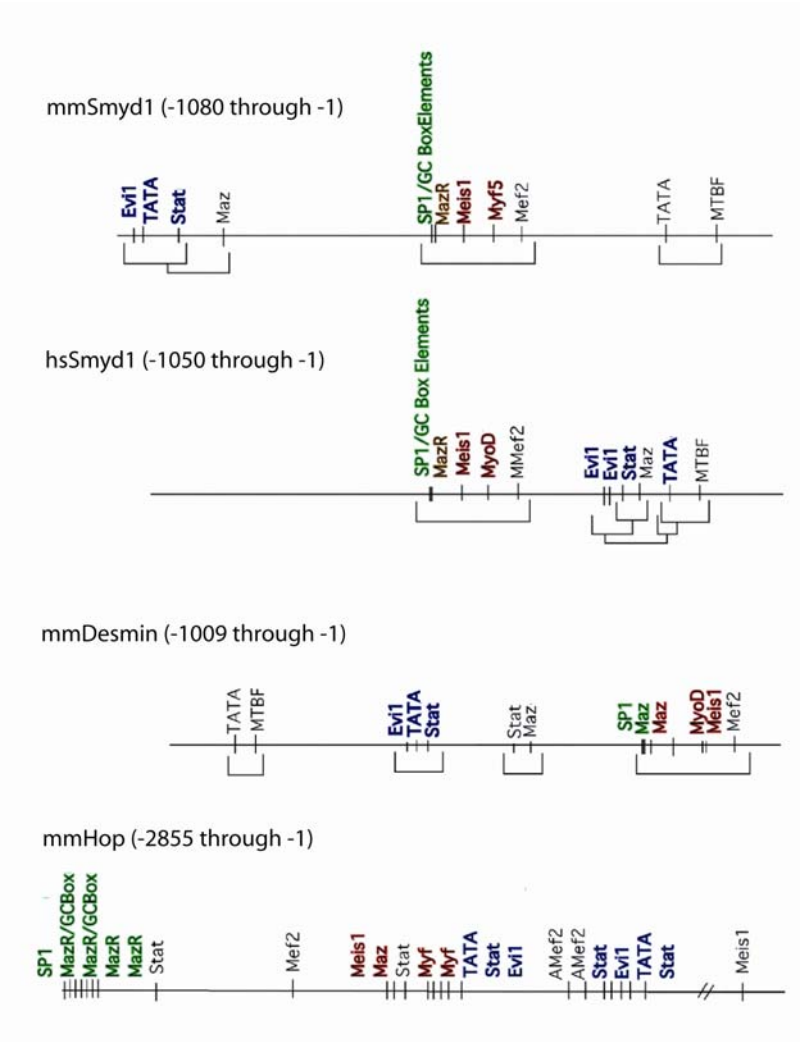


**Fig 3.2 Several Motifs are conserved beteween mm*Smyd1*, hs*SMYD1*, and mm*Desmin*, but not with mm*Hop*.** There are many putative transcription factor binding sites in common between all four promoters. However, the heart only promoter of mm*Hop* does not conserve the same organization as the cardiac and skeletal muscle expressing promoters of mm*Smyd*1, hs*SMYD1*, and mm*Desmin*. There are four clusters, indicated in brackets, in common between these 3 promoters. The most impressive of which encodes putative binding sites for: *Sp1*, *Maz*, *Meis1*, *MyoD* or *Myf5*, and *Mef2*.

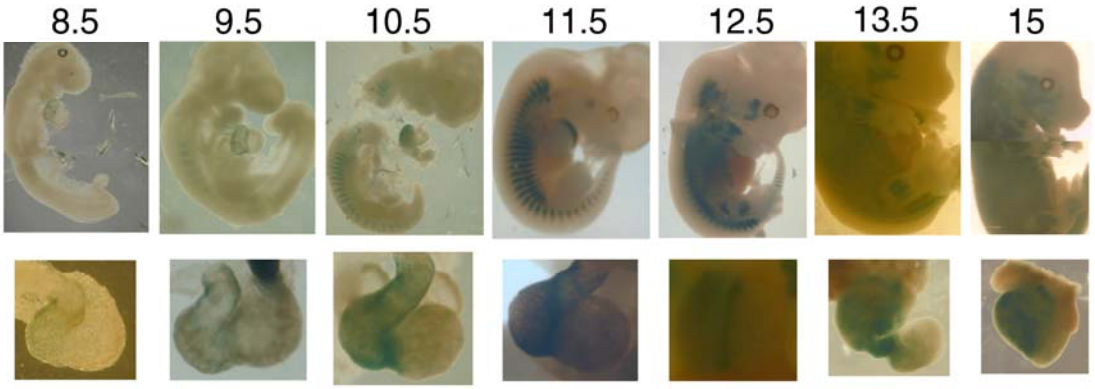


**Fig 3.3 Time progression of reporter expression of a 1kb regulatory region directly upstream of *Smyd1*.** A *Smyd1* reporter is expressed as early as E8.5 in the right heart and E9.5 in somites. Expression is heightened and maintained throughout development. The top panel represents whole embryo staining at the respective ages. The bottom panel represents cardiac staining at respective ages.

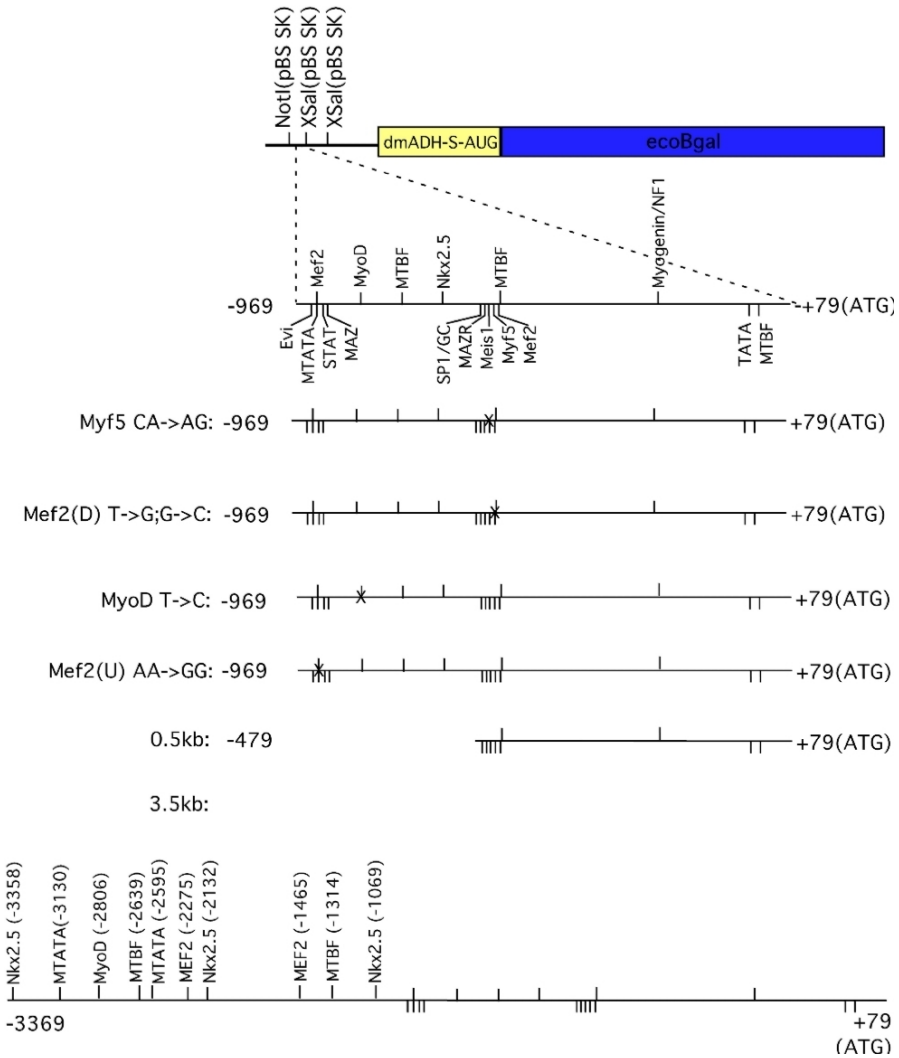


**Fig 3.4 Promoter Reporter Constructs in the Promoterless AUG Background.** 7 wildtype and mutagenic constructs were created on the promoterless background. Interesting putative transcription factor binding sites are annotated on the 1kb wild type construct and represented by hash marks on all others. Mutations are described on the left and marked with an X on the promoter. The 1kb wildtype and 3.5kb wildtype constructs were used to make stable transgenic lines.

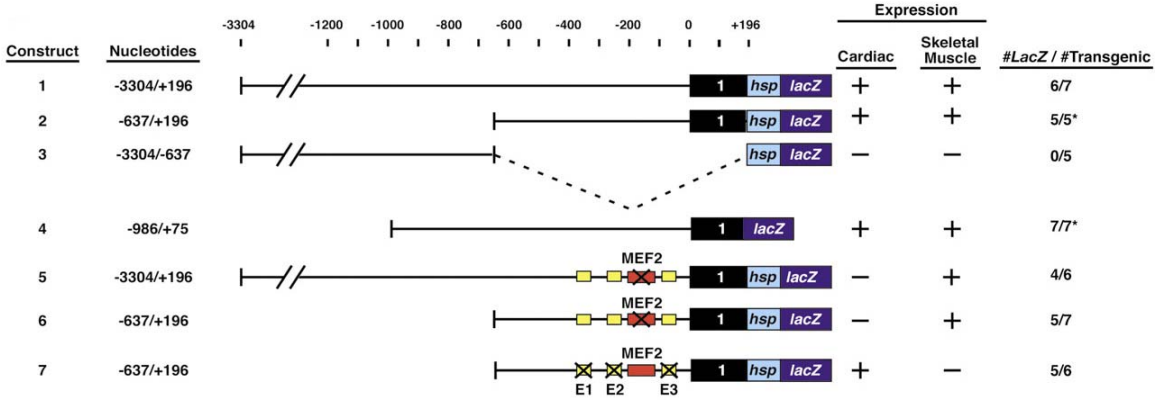


**Fig 3.5 Promoter Reporter Constructs Fused to a Basal Heat Shock Promoter.** Regions of 5' flanking DNA immediately upstream of *Smyd1* exon 1b used to create *lacZ* transgenes are shown. Nucleotides are numbered relative to the transcriptional start site, which is designated `0'. Construct number is indicated on the left, and the corresponding expression pattern is summarized on the right. All constructs, except construct 4, contain the indicated upstream genomic regions fused to the *hsp68* basal promoter and the *lacZ* gene. Construct 4 contains the region from –986 to +75 bp fused directly to promoterless *lacZ*. The *Mef2* site was mutated in constructs 5 and 6, and the three E-boxes were mutated in construct 7. Asterisk denotes constructs that were also used to generate stable transgenic lines. The numbers of F0 transgenic embryos compared with the total number of transgene-positive embryos are shown. Figure extracted from Phan and Rasmussen, et al [67].

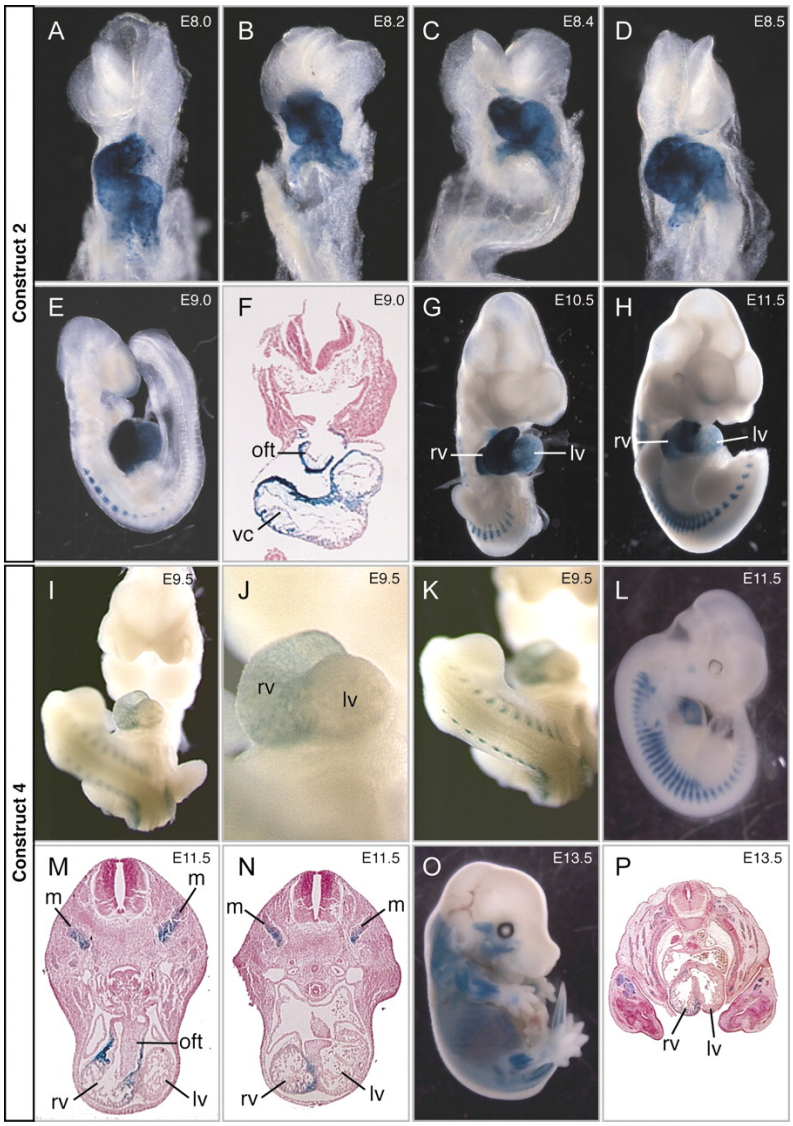


**Fig 3.6 Expression in the 1kb promoterless constructs is much weaker than expression in the 700 bp *hsp90* promoter constructs.** Expression patterns of transgenes directed by –637/+196 and –986/+75 regulatory regions of *Smyd1* at different developmental stages. (A-H) Transgenic embryos of the indicated ages from a stable transgenic line harboring construct 2 (–637/+196; see Fig 3.5). (A-D,G,H) Frontal views; (E) right-side view; (F) transverse section. Two independent stable transgenic lines displayed identical patterns of *lacZ* expression. (I-P) Transgenic embryos of the indicated ages from a stable transgenic line harboring construct 4 (–986/+75; see Fig 3.5). (J,K) Enlarged regions of the embryo shown in I. (M,N) Transverse sections at different levels of the embryo shown in L. (P) Transverse section of the embryo shown in O. rv, right ventricle; lv, left ventricle; oft, outflow tract; vc, common ventricular chamber; m, myotome. Figure extracted from Phan and Rasmussen, et al [67].

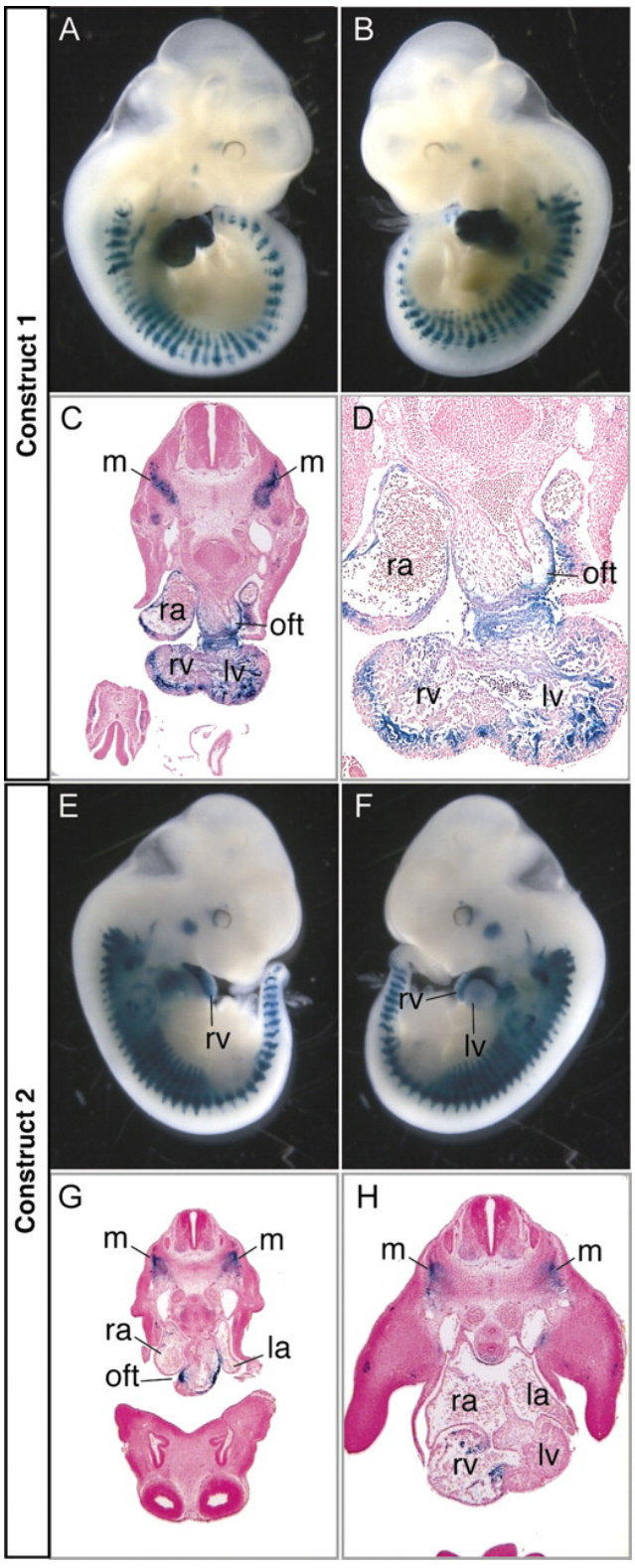


**Fig 3.7 Expression patterns of transgenes directed by distal (–3303/+196) and proximal (–637/+196) regulatory regions of *Smyd1* at E11.5.** Whole-mount views of F0 transgenic embryos at E11.5 with construct 1 (–3304/+196; A-D) and construct 2 (–637/+196; E-H) are shown (see constructs in Fig 3.5). (A,B,E,F) Left and right sides of embryos are shown. (C,D,G,H) Histological sections. (C,D) Robust expression of construct 1 is seen in the ventricular and atrial chambers of the heart and somite myotomes. Identical expression patterns were seen in 6 out of 7 embryos with this construct (data not shown). (G,H) Construct 2 is expressed specifically in the right ventricular chamber and somite myotomes. Identical expression patterns were seen in 5 out of 5 embryos with this construct (data not shown). m, myotomes; oft, outflow tract; ra, right atrium; rv, right ventricle; lv, left ventricle; la, left atrium. Figure extracted from Phan and Rasmussen, et al [67].

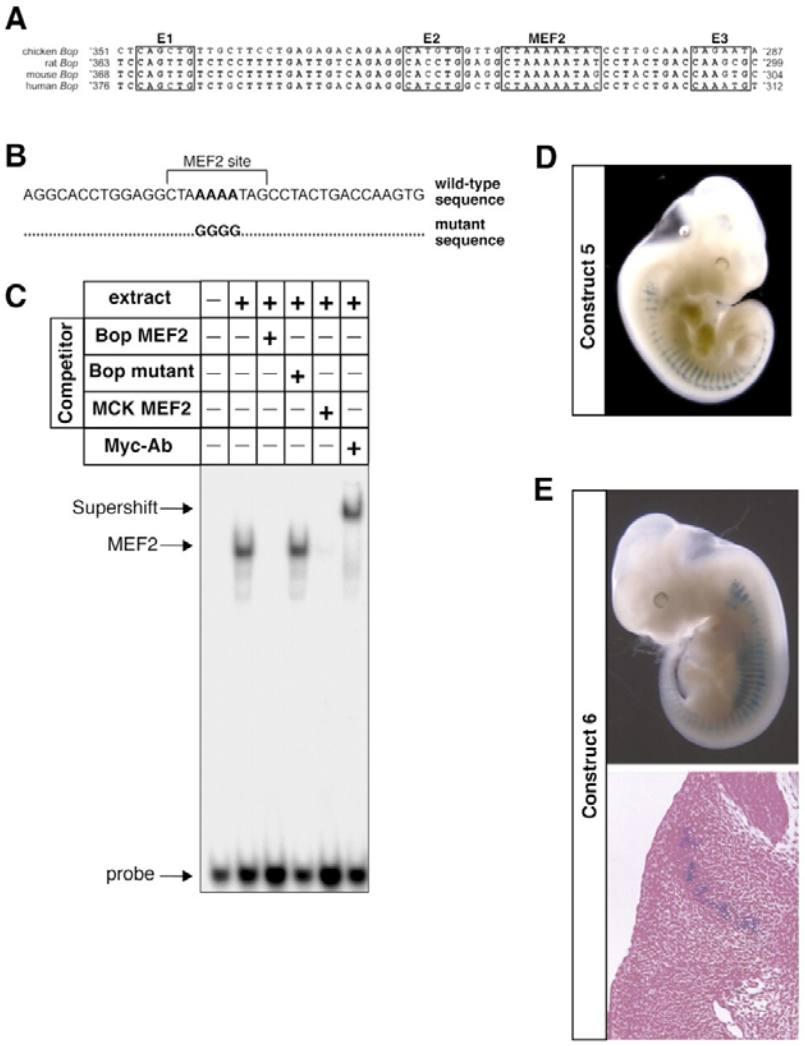


**Fig 3.8 A *Mef2*-binding site is required for cardiac expression of *Smyd1.*** (A) The sequence of the mouse *Smyd1* gene containing the essential *Mef2* site and the E-boxes is shown along with homology to the corresponding regions from other species. (B) 5' flanking sequences of mouse *Smyd1* genes showing conserved *Mef2* site at –329/-320. (C) A gel mobility shift assay was performed using nuclear extracts from COS-1 cells transfected with a *MYC-Mef2C* expression plasmid and the radiolabeled *Mef2* site shown in panel A as probe. Specific and nonspecific competitors were used at 50-fold molar excess. Antibody supershift used 1 µg of polyclonal anti-MYC antibody. (D,E) *Mef2* site is essential for *Smyd1* expression in the anterior heart field *in vivo*. F0 transgenic mouse embryos were generated using constructs 1, 2, 5 and 6 (see Fig 3.5), and stained for expression of ß-galactosidase at E11.5. Mutation of the *Mef2* site abolished expression in the anterior heart field without affecting expression in skeletal muscle. Four independent F0 transgenic embryos were analyzed with construct 5 (D) and five independent embryos were analyzed for construct 6 (E). All embryos with each transgene showed comparable expression patterns. Figure extracted from Phan and Rasmussen, et al [67].

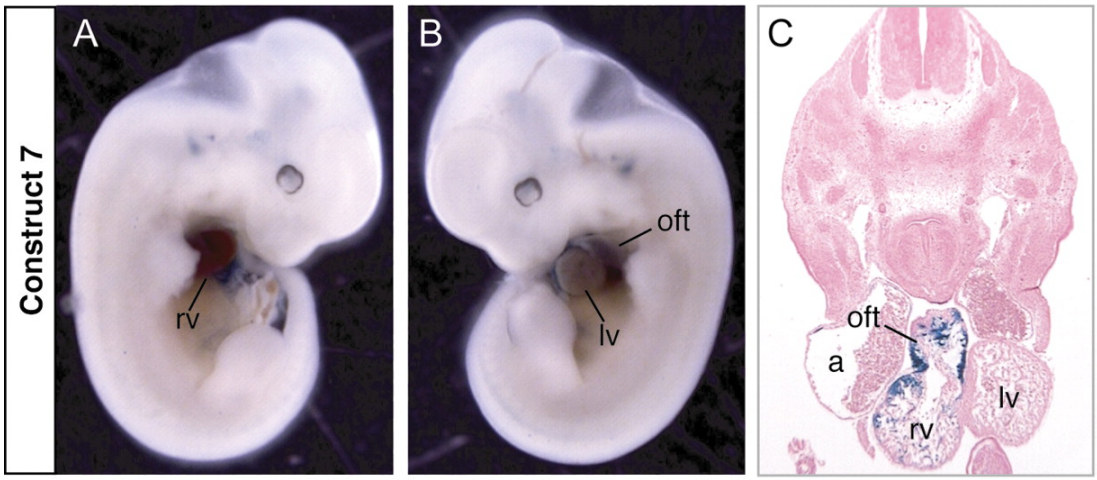


**Fig 3.9 E-boxes are required for *Smyd1* expression in developing skeletal muscle.** Mutation of the three E-boxes in the *Smyd1* upstream region (construct 7 in Fig 3.5) caused a loss of skeletal muscle expression without affecting expression in the heart, as seen in a whole-mount lateral view (A,B) and transverse sections (C) of five independent F0 transgenic embryos at E11.5. a, atrium; rv, right ventricle; lv, left ventricle; oft, outflow tract. Figure extracted from Phan and Rasmussen, et al [67].

*Discussion*

Conservation of DNA sequences between species is typically indicative of an essential function by those sequences. Typically this function is to code RNA or protein, however, when non-coding DNA is conserved it suggests that there are important DNA binding sites within the conserved regions. However, because there is not much variation at the DNA level, many putative binding sequences may also be conserved between species regardless of whether the site is a true *in vivo* target site. Therefore, common putative sites in non-conserved regions may be equally important—they are still conserved although the DNA sequence is not.

Predictions have been made based on computational analysis and further validated via literature reviews and some experimentation. I have found four clusters of putative binding sites that are likely important for normal expression in skeletal muscle based on the regulatory regions of multiple genes that produce skeletal muscle and heart specific expression. These predictions have not been tested experimentally, however, this could be done given the time and money to create more mutated promoter-reporter mice. The most impressive of the four clusters contains a *Pbx-1*, *Meis-1* and MRF binding site, *Mef2*, and *Sp1/Maz* sites. *Pbx* and *Meis* are known to bind as heterodimers cooperatively with all MRF heterodimers. *Mef2* is known for a large role in skeletal and cardiac muscle and we have shown this site to be essential for cardiac expression. *Sp1* and *Maz* are known to repress and activate in different contexts. This cluster of sites is not only conserved between mouse *Smyd1*, human *SMYD1*, and mouse *Desmin*, it is also in precisely the same location in mouse and human *SMYD1* regulatory regions. The second cluster is much smaller, but is also in the same location in both mouse and human regulatory regions and contains a TATA box adjacent to an unknown myotube specific enhancer. The third and fourth clusters are combined in mouse and human *SMYD1*

regulatory sequences and contain several *Evi1* sites, a TATA box, *Maz*, and *Stat*. While these are not spatially conserved, the mouse regulatory region does contain two *Lmo2* comples half sites that would allow the promoter to fold [68] and bring the mouse region approximately as close to the initiation site as the human region (Fig 3.10). While it is unlikely that the TATA box has any significance at this location, it does tend to cluster with the *Evi1* and *Stat* putative transcription factor binding sites.

Predictions specific to *Smyd1* regulatory regions have been made by comparing putative sites in mouse and human promoters. Some factors, like *Bcl6*, have been included despite seeming unlikely candidates in controlling the spatial expression of *Smyd1*. There was no evidence to exclude these factors and doing so would be arbitrary manipulation of data. However, by taking the factors we have shown to have importance and those that are likely to be important, I have created a schematic of the most likely players in the transcriptional regulation of *Smyd1* (Fig 3.10). This includes, a TATA box and Myotube specific enhancer, *Myogenin*, 2 *Mef2* sites, two *Lmo2* complex half-sites, *Myf5*, *PBx/Meis1*, *Meis1*, *Sp1*/*Maz*, *Comp1*, and *MyoD*.

It would make sense that *Myogenin*, *Myf5*, and *MyoD* are all involved in *Smyd1* transcription because these are three of the four MRFs and *Smyd1* expression is maintained in all known skeletal muscle in all known stages of development and adulthood. Therefore at different times and places, different MRFs may be necessary to activate *Smyd1* expression. By mutating three putative E-boxes, which coincide with the putative *Myf5* site, the putative *Lmo2* complex half site 1, and an unidentified region regulatory region, we have shown that skeletal muscle transcription is dependent on E-boxes and likely MRF heterodimer binding. Skeletal muscle transcription may also be dependent on the folding induced by the *Lmo2* complex which would position the distal *Mef2* site. We have shown that the proximal *Mef2* site is only essential for cardiac

expression, but we know that skeletal muscle *Mef2c* knockouts cause a two-fold reduction of *Smyd1*. Therefore it is likely that the distal *Mef2* site affects, but is not essential for *Smyd1* expression. The *Mef2* mutations analyzed were either on the 0.6kb regulatory context and the E-box mutations affected the integrity of the *Lmo2* complex half site, which contain and are dependent on E-boxes. *Comp-1* is also known to enhance the binding of MRFs to E-boxes.

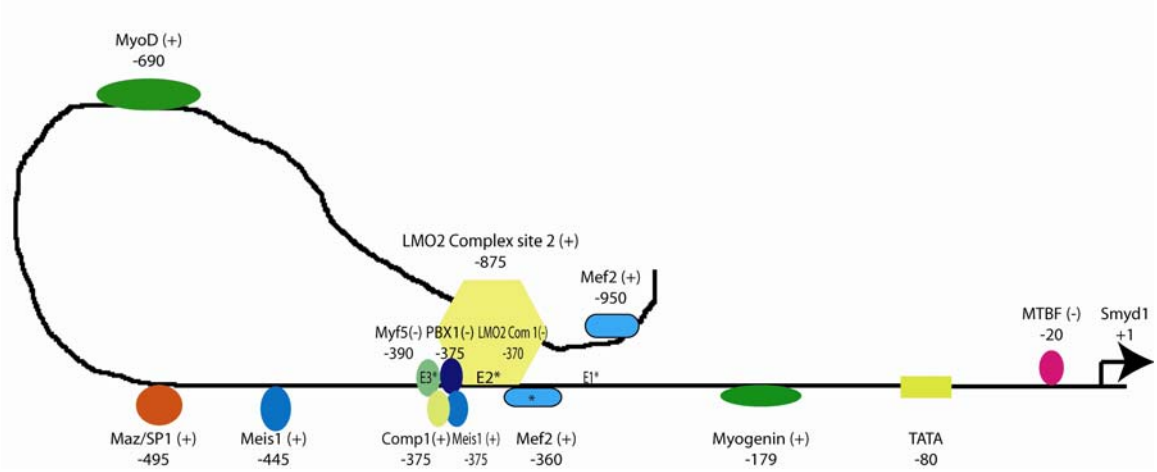


**Fig 3.10 A comprehensive prediction of important mm*Smyd1* promoter binding sites.** Based on experimentation we have shown that the *Mef2* site at -360 as well as the 3 E-boxes that overlap the putative *Myf5*, *LMO2* complex half-site 1, and an unidentified putative binding site are essential for proper transcription in several settings. The other putative binding sites are expected after extensive computational analysis of the sequence and literature reviews. Mutations analyzed are indicated with an *. The position is marked relative to the +1 transcription site. (+) and (–) indicate whether the putative site is on the coding or non-coding strand.

## Chapter 4. *Smyd1* is Essential in all Stages of Embryonic Cardiac Development.

Targeting vectors made for eukaryotic knock-ins or knock-outs typically include an antibiotic cassette, Neomycin, which allows for positive selection. However, this screening strategy integrates a Neomycin cassette into the gene locus. Some reports indicate that neomycin can provide splice acceptor and donor sites in order to incorporate itself into the mRNA, creating a hypomorphic gene product [69, 70]. Theoretically, the presence of the Neomycin cassette could also displace essential regulatory regions and affect regulation or expression of genes near the locus.

With the introduction of *Cre*/LoxP and *FLP*/FRT technology, it has become possible to remove the drug resistance cassettes after integration of the targeted vector into the mouse genome. *Cre* recombinase is a protein isolated from bacteriophage P1 that has been shown to specifically recognize a LoxP site, a 34 bp sequence, and induce recombination between any two LoxP sites [71]. If the LoxP sites are on the same chromosome or plasmid in opposite orientations, *Cre* recombinase causes an inversion. However, if the LoxP sites are in the same orientation on the same chromosome or plasmid, *Cre* recombinase excises all DNA between the two sites [72]. Similarly *FLP* recombinase, first identified in yeast, specifically recognizes and induces recombination of FRT sites, a distinct DNA sequence [73].

It has also become possible to use these same sites to knock out endogenous DNA sequences of interest conditionally, or only in specific tissues, cell lineages, or stages of life, by using transgenically expressed recombinases under the control of specific temporally and/or spatially regulated promoters. Segments of interesting DNA, usually encoding specific genes of interest, are flanked by either LoxP or FRT sites and introduced into the mouse genome via targeted knock-in technology. Mice with the

targeted knock-ins of LoxP or FRT sites are mated with those that carry a transgenic allele of the appropriate recombinase, which are fused to a known promoter in order to induce specific, predictable patterns of expression. Thus, the recombination event occurs at the genomic level only in the appropriate temporal and spatial manner. In many recent conditional knockouts both recombinases are employed. Often the FRT/FLP system is employed to specifically remove the Neomycin cassette and the LoxP/Cre system is used to induce the knockout of the gene of interest. Because *FLP*/FRT tools were not well developed when we initiated cloning for the mouse *Smyd1* conditional knockout, we created a targeting vector in which both a region of the *Smyd1* gene as well as the Neomycin cassette were flanked with LoxP sites.

The conventional *Smyd1* mouse knockout ($Smyd1^{KO/KO}$) yielded a very interesting early embryonic cardiac phenotype. In the *Smyd1* null embryos, there was an expansion of "cardiac jelly", excess hyaluronic acid within the chambers, a pericardial edema and lack of ventricular septation [10]. The single ventricular chamber had the molecular markers of a left ventricle, but lacked markers of right ventricles [10]. However, because $Smyd1^{KO/KO}$ was embryonically lethal, these mice did not allow us to examine the role of *Smyd1* in other temporal and spatial specific contexts and prevented us from collecting large amounts of *Smyd1* deficient tissue for molecular analysis. Also, we could not prove that the phenotype observed in the conventional knockout was due to an absence of *Smyd1* in any particular tissue. The phenotype could have been due to non-autonomous effects on the heart or effects caused by the introduction of the neomycin cassette to the locus instead of the absence of *Smyd1* gene products specifically in the heart. Therefore, we further investigated the role of *Smyd1 in vivo* by creating *Smyd1* conditional knockout mice.

As discussed in chapter 1.1, the *Smyd1* gene expresses multiple alternatively spliced isoforms and the initiation site varies for different isoforms. Therefore, similar to the conventional *Smyd1* knockout, we opted to target exons that were common in all known isoforms. We created a targeted knock-in line of mice that had LoxP sites flanking *Smyd1* exons 2 and 3 followed by a Neomycin cassette and a third LoxP site (Fig. 4.1A). This was electroporated into 129S6 ES cells and selected for with G418 and gancyclovir. Clones were screened for homologous recombination by Southern blotting (Fig. 4.1B). Correctly targeted ES clones were injected into C57BL/6 blastocysts to create 129.Bl6 chimeric mice, which were mated with C57BL/6 females to create germline knock-in mice. We then obtained existing *Cre* mice from various labs in order to further study *Smyd1* in cardiac and other tissues.

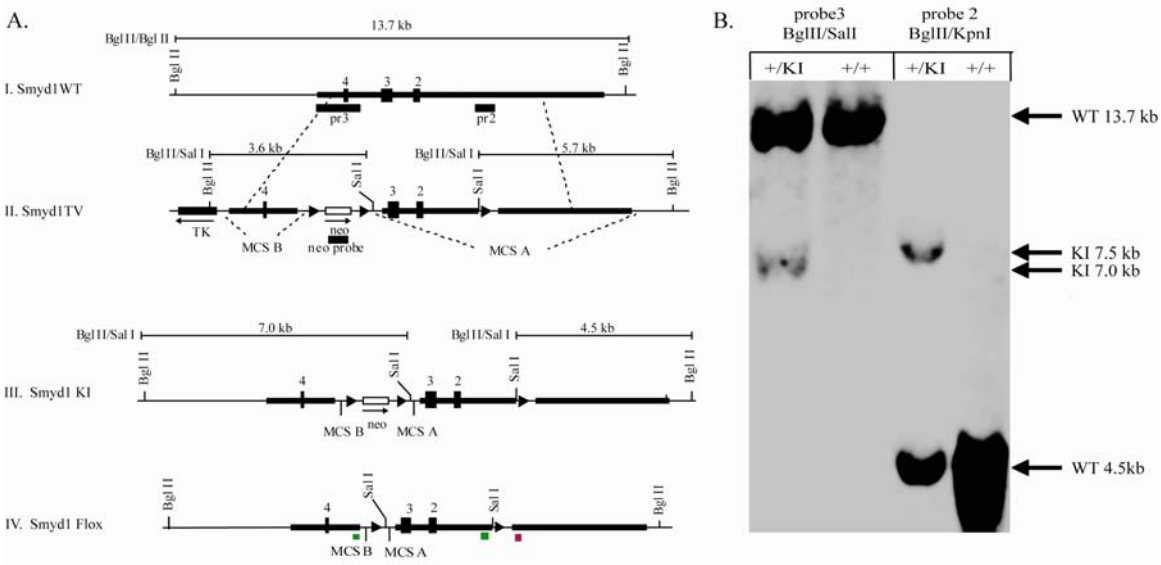


**Fig 4.1 Targeting strategy of *Smyd1* Flox allele**. **A.** The targeting vector (II) was electroporated into SM1 (129S6) ES cells. After selection with G418 and gancyclovir, surviving clones were screened for homologous recombination with the genomic locus (I). The resulting targeted knock-in (III) was identified by southern blot. ES cells showing correct integration in both arms were injected into C57Bl/6 blastocysts, which were then transferred into the uteri of 2.5 d.p.c. pseudo-pregnant females to create chimeric mice. The chimeric mice were mated to C57Bl/6 females to create a germline knock-in. The germline mice were crossed to a ubiquitous expressing *EIIA-cre* mouse and progeny were screened for the neo deletion (IV) by southern blotting. **B.** Southern blot of ES cells. Genomic DNA was digested by BglII and SalI and probed with pr3 to distinguish the 13.7kb wild type allele from the 7.0 kb targeted allele. Wildtype 129 DNA as well as the Targeting Vector were included as controls. Data not shown: Genomic DNA was digested by BglII and SalI and probed with pr2 to distinguish a 13.7kb wild type allele and a 4.5kb targeted allele. Controls were included.

## 4.1 The *EIIA-Cre* induced *Smyd1* knockout recapitulates the conventional *Smyd1* KO.

In order to determine that the phenotype derived from the conventional mouse knock-out, $Smyd1^{KO/KO}$, was due to the absence of *Smyd1* gene products and not to the introduction of the neomycin cassette, we recreated the Smyd1 knockout using *EIIA-cre* transgenic mice. This mouse line was a C57Bl6/FVB hybrid created in the Westphal lab. *EIIA-cre* is expressed in pre-implantation zygotes [74]. Therefore, *EIIA-cre* induces recombination during the single cell stage of embryogenesis and all cells in any particular embryo should contain a genome that reflects the same recombination event: a ubiquitous effect.

When mating *EIIA-Cre* (Bl6.FVB) mice to the original $Smyd1^{KI/+}$ mice (Bl6.129) several recombination events can occur on the $Smyd1^{KI}$ allele, including no deletion ($Smyd1^{KI}$), total deletion ($Smyd1^{\Delta KI}$), deletion of only the neomycin cassette ($Smyd1^{Flox}$), and deletion of the endogenous sequence, but not neomycin ($Smyd1^{\sim KO}$) (Fig 4.2). Resulting genotypes were screened via Southern blotting (Fig 4.1). $Smyd1^{Flox}$ mice, BL6.FVB.129 hybrids, were backcrossed one generation to C57BL/6J mice and screened for the absence of Cre and presence of $Smyd1^{Flox}$. These were kept for breeding to tissue specific *Cre* lineages. $Smyd1^{\Delta KI}$ were bred to homozygosity to compare to the original $Smyd1^{KO/KO}$ mice.

The $Smyd1^{\Delta KI}$ allele varies from the $Smyd1^{KO}$ allele in two ways. Firstly, $Smyd1^{\Delta KI}$ does not contain a neomycin cassette while $Smyd1^{KO}$ does. Secondly, $Smyd1^{KO}$ deletes an extra 290 base pairs in the intron between exons 3 and 4. The time of death and the gross phenotype of $Smyd1^{\Delta KI/\Delta KI}$ match that of $Smyd1^{KO/KO}$ (Fig 4.3), suggesting that the addition of the neomycin cassette and the varying intron segment plays no role in this phenotype. Therefore this eliminates the possibility that the *Smyd1* conventional

knockout phenotype is caused by a neomycin hypomorph and supports the hypothesis that the phenotype observed in both mouse lines is due to the absence of the *Smyd1* gene products.

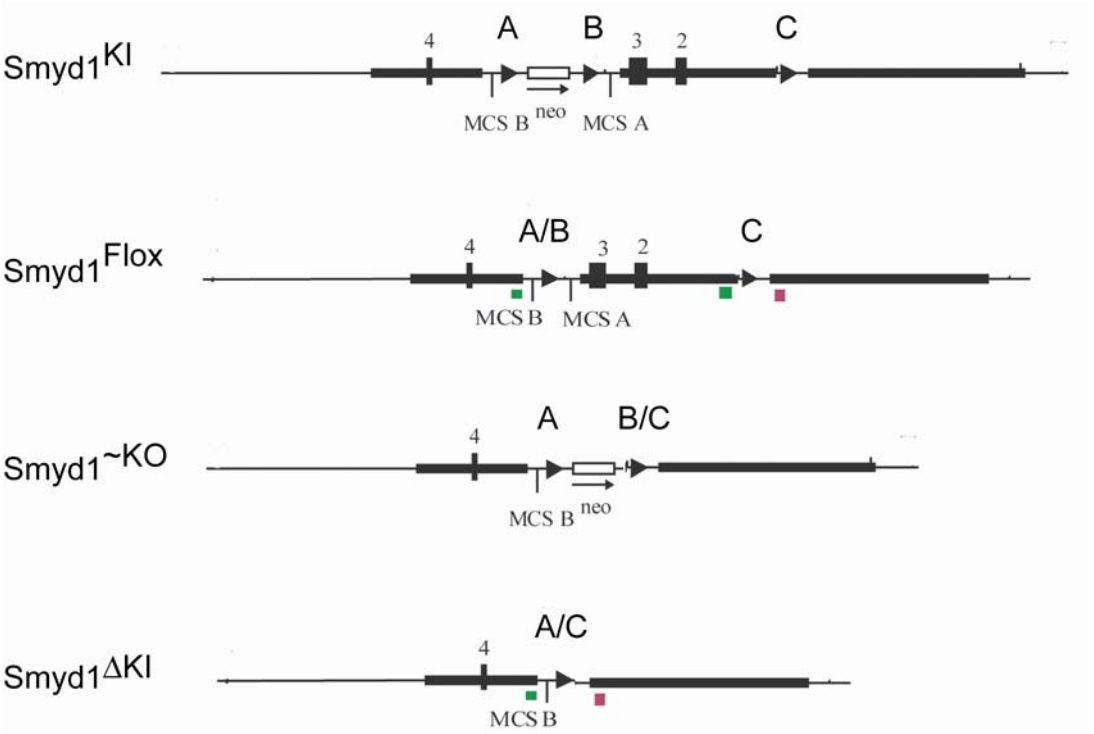


**Fig 4.2 Possible LoxP recombination events**. LoxP sites are designated as A, B and C. Recombination events can occur between A and B resulting in *Smyd1*$^{Flox}$, B and C resulting in *Smyd1*$^{\sim KO}$, A and C resulting in *Smyd1*$^{\Delta KI}$, or not at all maintaining *Smyd1*$^{KI}$. Results of these are depicted above.

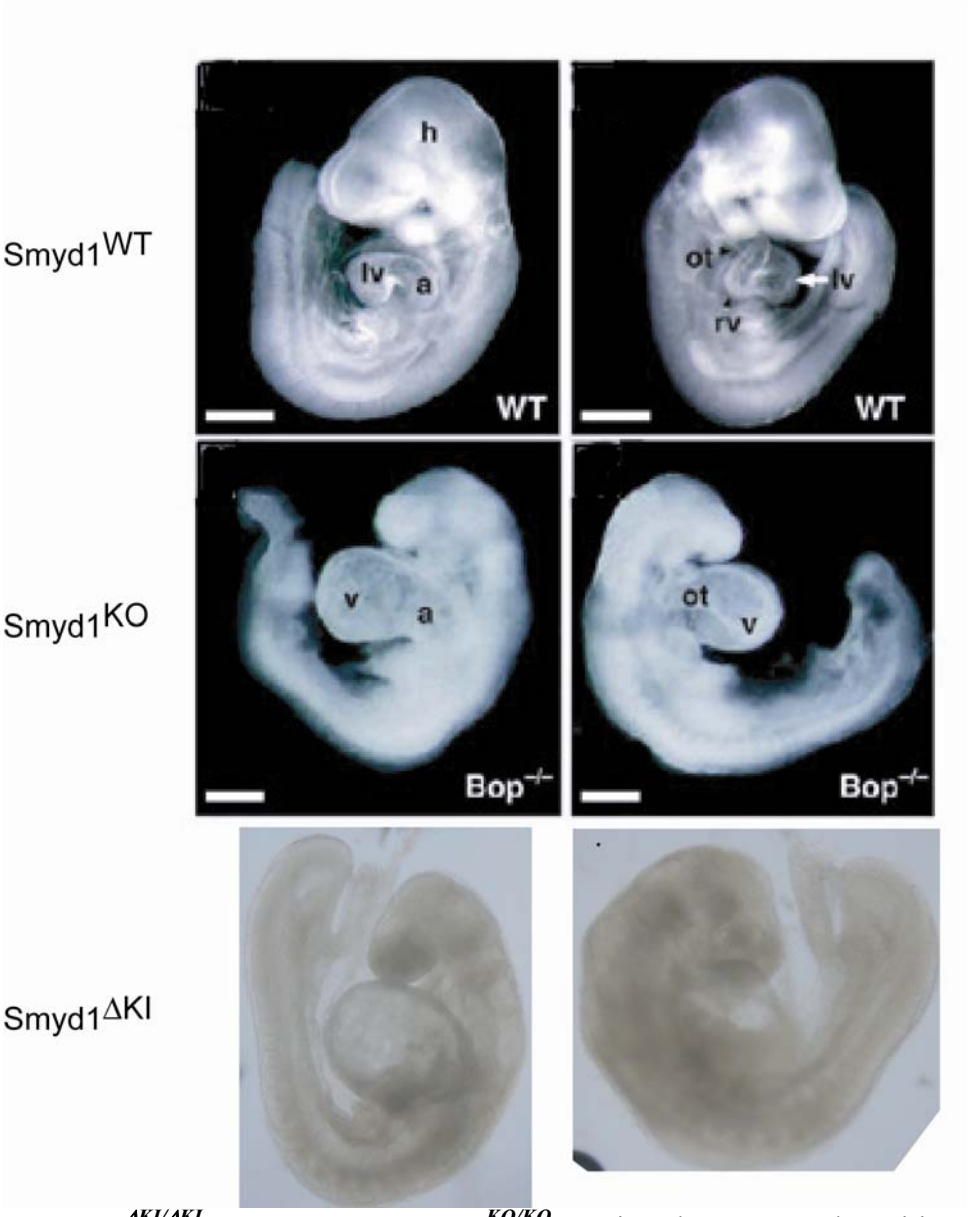


**Fig 4.3** ***Smyd1*$^{\Delta KI/\Delta KI}$ phenocopies *Smyd1*$^{KO/KO}$**. Mice that screened positive for the *Smyd1*$^{\Delta}$*KI* allele were timed mated to each other and their embryonic, homozygous offspring were analyzed for gross phenotype. The time of death and outward appearance match that of the conventional *Smyd1* knockout suggesting that the Neomycin cassette does not contribute to this phenotype.

## 4.2 The *Nkx2.5-Cre* mediated *Smyd1* knockout demonstrates that *Smyd1* is essential post ventricular septation.

*Nkx2.5-Cre* mice were obtained from the Olson laboratory on the B6C3F1 strain, which is an isogenic, first generation hybrid of C3H and C57Bl/6J. Recombination induced by *Nkx2.5-Cre* occurs as early as E8.0, during the linear heart-tube stage of cardiogenesis, and is confined to the cardiac ventricles at E8.5 with minimal recombination in the outflow tract and atria at E12.5 [75]. By crossing the $Smyd1^{Flox}$ allele to the *Nkx2.5-Cre* transgene, we analyzed whether the *Smyd1* deficient phenotype is due to an absence of Smyd1 in cardiac ventricles.

Because mice rarely show phenotypes due to heterozygosity or partial deletion, because we were not sure how effective *Cre* recombination would be, and because we had the $Smyd1^{KO}$ mice readily available, we induced our knockout on a mixed background: $Smyd1^{Flox/KO}$. Mice were mated in order to ultimately create $Smyd1^{Flox/KO}$;$Nkx2.5\text{-}Cre^{+/?}$ mice (BL6.FVB.129.C3H) as well as the following breeders of the same background: $Smyd1^{Flox/WT}$; $Nkx2.5\text{-}Cre^{+/?}$ and $Smyd1^{KO/WT}$; $Nkx2.5\text{-}Cre^{+/?}$. Once this combination of alleles was obtained, these mice were only crossed to each other, thereby maintaining a population that is similar and inbred, but non-isogenic.

*Tissue Surveys on the Smyd1 Locus*

In order to confirm that the $Nkx2.5\text{-}Cre^{+/?}$ transgene is effective on the $Smyd1^{Flox}$ locus as previously described on other loci, DNA recombination tissue surveys were conducted on heterozygous $Smyd1^{Flox/WT}$; $Nkx2.5\text{-}Cre^{+/?}$ mice. At the time of sacrifice, I collected several tissues and two PCR based recombination assays were used. The first assay detects $Smyd1^{WT}$ and $Smyd1^{Flox}$ alleles that have been unaltered by recombination. $Smyd1^{WT}$ does not contain LoxP sites; therefore, regardless of *Cre* expression, levels of

genomic *Smyd1*$^{WT}$ cannot decrease, providing an internal positive control. Therefore, despite variance in DNA preparations from sample to sample, the *Smyd1*$^{WT}$ band represents 100% of one allele. *Smyd1*$^{Flox}$, however, is the target of recombination. In a perfect system, tissues in which recombination occurs should not yield a *Smyd1*$^{Flox}$ band. However, most *Cre* recombination is not 100% effective and, therefore, we expected to see the band diminish or disappear in tissues displaying strong to complete recombination. These PCR assays were not confirmed to be in the linear range; and therefore we can not quantitate the products via densitometry, but by comparing the *Smyd1*$^{WT}$ and *Smyd1*$^{Flox}$ bands, we can qualitate the efficiency of recombination.

In order to test this, I performed controls on both plasmid and genomic tail DNA. A plasmid containing an XbaI fragment from a BAC-clone served as a wildtype plasmid template, while the targeting vector, pKI, used to make the conditional knockout served as Flox plasmid template. By combining these two plasmids in known molar ratios, I assessed how well the ratio of PCR products represents the ratio of the two alleles. The Flox and WT bands appear to be equal when the WT allele is 13 fold more concentrated than the Flox allele (Fig 4.4A). This data suggests that the PCR reaction favors the Flox product. It is not unusual for PCR reactions to prime products at different rates, although it is unusual for the reaction to favor the larger product. The 30:1 WT/Flox molar ratio detected an obvious decrease in the Flox product.

To determine if this favoritism also occurs in genomic DNA, I obtained genomic tail DNA that was determined to be homozygous for either the WT or the Flox allele. By combining these two genomes in known mass ratios, I assessed whether the findings for plasmid DNA are true for genomic DNA. In the genomic setting, it appears that the 5:3 WT:Flox ratio gives equal products (Fig 4.4B). Therefore the favoritism towards the Flox allele is still present, although it does not appear as severe with genomic templates.

A decrease in the Flox band appears in the sample with 5:2 WT/Flox ratio, which simulates 60% recombination. The Flox band disappears at the 100:1 WT/Flox ratio, which simulates a 99% recombination.

Tissue surveys were done on entire tissues. However, *Nkx2.5-Cre* induces recombination primarily in the ventricles. Although atria and valves are much smaller than ventricles, DNA from this unaffected tissue can still mask recombination in affected tissue. There is also a possibility of contaminating tissues with red blood cells, fur, or other surrounding tissues. Due to these factors, recombination is difficult to detect with this assay and a decrease in the $Smyd1^{Flox}$ band compared to the $Smyd1^{WT}$ band will only be detectable if recombination is strong. Therefore, recombination was also assayed via the Floxdel assay, which detects only a recombination product. By coupling data sets from both reactions I can semi-quantitatively analyze whether recombination is occurring in predicted patterns and whether this recombination is strong.

PCR products from the Floxdel assay on tissues harvested from mice carrying the *Nkx2.5-Cre* transgene only occurred in heart tissue and in mice that also had the targeted $Smyd1^{Flox}$ allele (Fig 4.4C). However, the Flox/WT PCR results did not consistently display a decrease in the $Smyd1^{Flox}$ allele compared to the $Smyd1^{WT}$ allele. In Mouse #7, there is an obvious reduction in the $Smyd1^{Flox}$ allele compared to the $Smyd1^{WT}$ allele. This reduction is not obvious in surveys from mouse #4. Therefore, while recombination is occurring in the appropriate tissue, it is not complete and it is possible that the rate of recombination varies from individual to individual. RT-PCR confirms that expression of *Smyd1* transcripts is reduced but not complete (Fig 4.5A) and real time PCR shows a 90% knockdown (Fig. 4.6B).

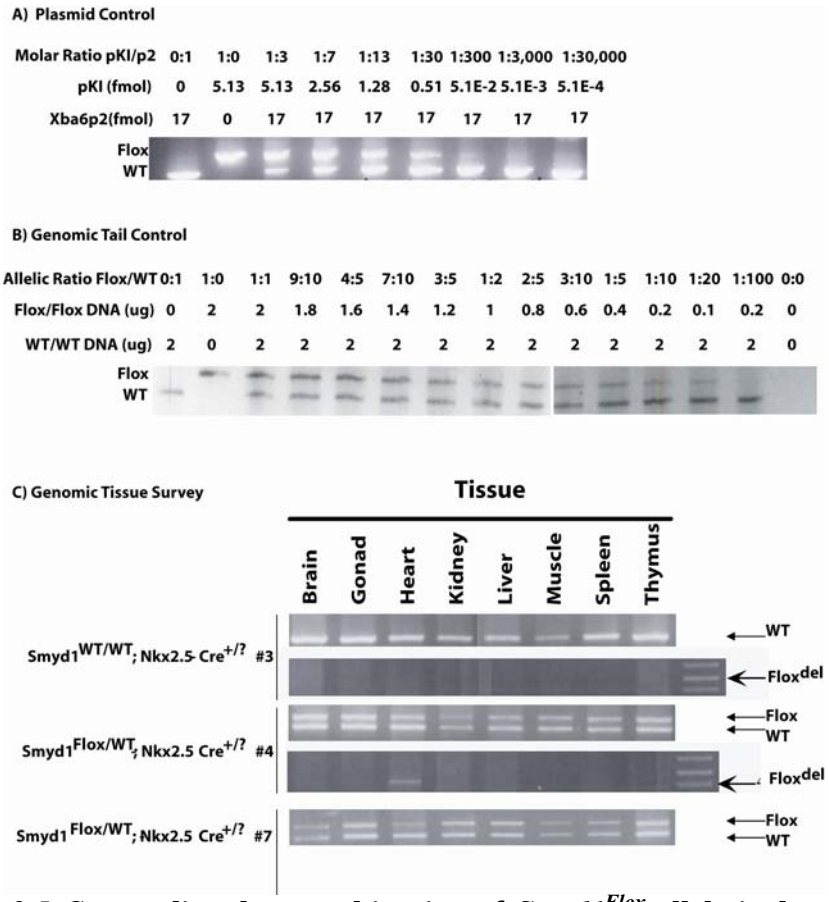


**Fig 4.4** ***Nkx2.5-Cre*** **Mediated Recombination of** ***Smyd1$^{Flox}$*** **allele is detected only in the heart.** **A)** Combining a WT and Flox plasmid in known ratios displays that PCR favors the Flox plasmid. However, a decrease in the Flox band is detectable when the WT band is significantly more prevalent than the Flox band. **B)** PCR reactions also favor the flox product when genomic DNA is combined in known ratios. A flox band that is weaker than the WT band signifies that the WT band is significantly more prevelant than the Flox band. **C)** Analysis of Mouse tissues does not commonly detect a decrease in the Flox band due to recombination. The recombination assay detects cardiac specific recombination, as predicted. While the Flox/WT bands do not consistently represent an expected effect on the Flox allele, they do provide a positive control for all samples are confirm that PCR grade DNA is present in relatively equal amounts in samples from all tissues.

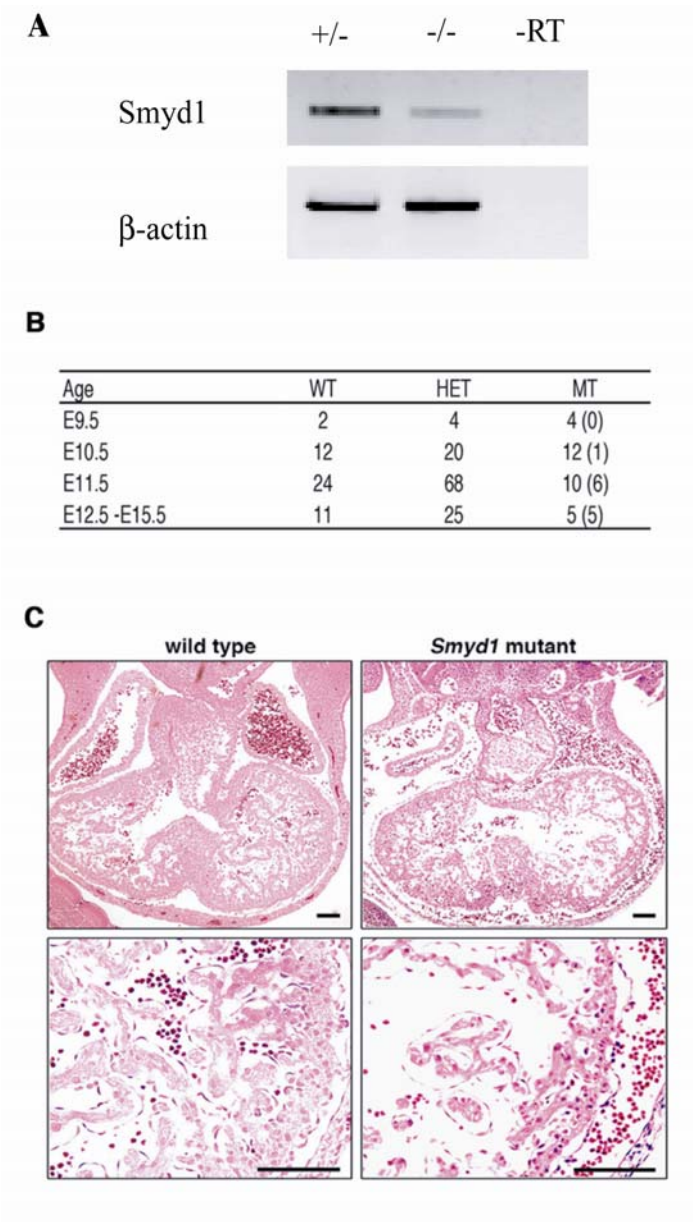


**Fig 4.5 When *Smyd1*$^{Flox}$ is combined with *Nkx2.5-Cre, Smyd1* is effectively knocked down and mice display an embryonically lethal cardiac phenotype. A)** A Smyd1 knock-down is detectable, although incomplete at E10.5 as detected by RT-PCR. **B)** To investigate the lethality, *Smyd1*$^{WT/Flox}$*; Nkx2.5-Cre* were crossed with *Smyd1*$^{WT/KO}$*; Nkx2.5-Cre*. *Smyd1*$^{WT/WT}$; *Nkx2.5-Cre* is referred to as WT and *Smyd1*$^{Flox/KO}$; *Nkx2.5-Cre* is referred to as MT. Both *Smyd1*$^{KO/WT}$ and *Smyd1*$^{Flox/WT}$ with the *Cre* are referred to as HET. Analysis of timed matings determines that CKO mice are present and normal in Mendelian ratios at E9.5. Early abnormalities are detectable at E10.5 and all CKO mice are dead by E12.5. Numbers shown are observed animals and numbers in parenthesis indicate phenotypic or dead embryos. **C)** H&E stained transverse sections show pericardial edema, cardiac hemorrhaging, a thinned pericardium, decreased trabeculation, and poor septation.

*Phenotype of Nkx2.5-Cre mediated Smyd1 CKOs*

Despite the incomplete effect of the *Nkx2.5-Cre* on the Smyd1 locus, this conditional knockout, or knockdown, did prove to be embryonic lethal. The first 50 pups born from matings described above yielded 20% *Smyd1* wildtype homozygotes, 80% *Smyd1* heterozygotes, and 0% genotypic *Smyd1* nulls (table 4.1). Expected Mendelian ratios are 25%, 50%, and 25%, respectively. The large number of heterozygotes may indicate that there is selective segregation of the mutant alleles. However, this was not analyzed as the more interesting feature is that homozygous mutants did not exist at weaning. There was no consistent or significant reduction in litter sizes between birth and weaning, the age at which we performed genotyping.

To determine the time of death of homozygous *Smyd1* mutants, we analyzed embryos from timed matings. At E9.0, when the conventional *Smyd1* nulls are visibly affected, *Smyd1* CKO embryos appeared to be normal. We determined that the time of death varied, but was typically between E10.5 and E12.5 (Fig. 4.5B). At E10.5 Mendelian ratios were not abnormal and only 1 of 12 conditional knockout embryos displayed a phenotype. By E11.5 Mendelian ratios were significantly lower than expected ($p<0.01$) and 60% of observed CKO embryos displayed a marked phenotype. 100% of CKO embryos analyzed between E12.5 and E15.5 lacked a functional heart (no heart beat). Although the Mendelian ratios did not continue to decrease between E11.5 and E15.5, all observed embryos were developmentally stunted at E12.5 and no survivors were found at birth.

Embryos were analyzed for phenotype at E11.5. At this age, embryos display pericardial edema, cardiac hemorrhaging, poor trabeculation, disorganized ventricular septum, and a thinned ventricular wall (Fig. 4.5C). Despite the fact that *Nkx2.5-Cre* induced recombination is reported to occur by E8.0, *Nkx2.5-Cre* induced *Smyd1* CKOs

survive 2 days longer than the *Smyd1* conventional knockouts. This is not surprising because 10% of the normal levels of *Smyd1* transcripts exist in the *Smyd1* CKO (Fig. 4.6B) as opposed to 0% in the *Smyd1* KO. However, it is still possible that *Smyd1* is involved in molecular events prior to *Nkx2.5-Cre* mediated deletion and that the full expression and function of *Smyd1* prior to E8.0 partially rescues the phenotype between E9.0 and E11.5. In the *Smyd1* CKO, septation does occur, however the septum is not normal. Therefore, even if embryos are rescued through the septation process, *Smyd1* gene products are still required.

| Smyd1 Allele | Number of Animals Observed | Observed Ratio | Expected Ratio |
|---|---|---|---|
| WT | 10 | 0.20 | 0.25 |
| Het | 39 | 0.80 | 0.50 |
| MT | 0 | 0.00 | 0.25 |

**Table 4.1 *Nkx2.5-Cre* induced *Smyd1* CKOs are not viable.** *Smyd1*$^{WT/Flox}$*; Nkx2.5-Cre* were crossed with *Smyd1*$^{WT/KO}$*; Nkx2.5-Cre*. *Smyd1*$^{WT/WT}$ ; *Nkx2.5-Cre* is referred to as WT and *Smyd1*$^{Flox/KO}$; *Nkx2.5-Cre* is referred to as MT. Both *Smyd1*$^{KO/WT}$ and *Smyd1*$^{Flox/WT}$ with the *Cre* are referred to as HET. Litters were genotyped at weaning and observed numbers were recorded. There were no significant reduction in litter size between birth and weaning.

*Nkx2.5-Cre Mediated Smyd1 CKO Phenocopies skNAC*

Because *Smyd1* has primarily been found to interact with a protein called *skNAC* in a muscle library yeast 2-hybrid [6] and a cardiac library yeast 2-hybrid [7], we analyzed the possibility that these genes have cooperative roles. *Smyd1* and *skNAC* are expressed in very similar patterns; both have restricted expression patterns that include the myocardium, somitic myotome, and muscle cells. *Smyd1* is expressed as early as E7.75 and through adulthood in the myocardium of the mouse heart and is found in the myotome of somites as early as E12.5 [10]. Reporter assays suggest that *Smyd1* is expressed in somites as early as E9.0 [67]. *Smyd1* expression continues in developing and adult muscle tissue. *skNAC* expression is not detectable as early as *Smyd1* expression, but weak expression in the mouse looping heart is detectable by E8.5 (Personal Communication, C. Park and D. Srivistava). Cardiac expression of skNAC continues in the myocardium through adulthood. Somitic myotome expression is detectable at E12.5 and muscle expression continues through adulthood.

*skNAC* KO mice survive to adulthood, however, in non-Mendelian ratios. Approximately half of the *skNAC* knock out embryos displayed cardiac edema by E10.5 and like the *Smyd1* CKO, at E11.5, many *skNAC* KO embryos display pericardial edema, cardiac hemorrhaging, poor trabeculation, disorganized ventricular septum, and a thinned ventricular wall (Personal Communication, C. Park and D. Srivistava). In collaboration with the Srivistava lab at UCSF, microarrays on *skNAC* deficient E11.5 mouse hearts that did not display cardiac dysfunction were performed and candidate genes were analyzed by qRT-PCR in both *Smyd1* CKO and *skNAC* KO hearts. Surprisingly, only 23 genes were significantly deregulated in *skNAC* null hearts. Of these, 17 genes are characterized (Fig 4.6A). Most targets were also misregulated in the *Nkx2.5-Cre* induced *Smyd1* CKO (Fig 4.6B).

*Myoglobin*, which is a cytoplasmic hemoprotein expressed in cardiac and skeletal muscle is the most significantly down regulated gene with a 4 fold reduction in *Smyd1* CKOs and 2 fold reduction in *skNAC* KOs. *Myoglobin* is a known transcriptional target of *skNAC* [9]. It has since been shown that the methylation of *skNAC,* hypothetically by *Smyd1,* is essential for *skNAC* to bind to the promoter of *Myoglobin* (personal Communication L. Zhu, P. Tucker). We also show that hydroxymethylbilane synthase (*Hmbs*) and erythroid differentiation regulator 1 (*Erdr1*), enzymes known for involvement in heme synthesis, were upregulated. The fact that the heme transport protein, *Myoglobin*, was down regulated and heme synthesis enzymes, *Hmbs* and *Erdr1*, were upregulated, suggests that embryos are undergoing oxidative stress, which is also marked by the upregulation of anti-oxidants, *Metallotheinein1*, *Metallotheinein 2*, and *Vanin1*. However, *Vanin1* is oppositely misregulated in the two knockout systems. In the *skNAC* null, *Vanin1* is overexpressed but in the *Smyd1* CKO, *Vanin1* is down regulated.

Transcriptional factors such as *Shox2* and elongation factor *RNA polymerase II*, as well as a voltage gated potassium channel, *Kcne1*, are also interesting targets discovered in this microarray. Interestingly, *skNAC* levels were down regulated in the *Smyd1* CKO. This indicates that *skNAC* works downstream of *Smyd1* as well as parallel to *Smyd1*. Other known candidate genes were also analyzed in these mice. Because *Irx4* and *Hand2* were found to be misregulated in the conventional *Smyd1* KO, these targets were confirmed in the less severe *Smyd1* CKO as well as in the *skNAC* KO. *Irx4* was similarly downregulated in all 3 knockout systems. However, *Hand2* was only affected by the *Smyd1* knockouts. Whether or not this effect is direct, it demonstrates that *Smyd1* has transcriptional effects independent of *skNAC*.

**A**

| Gene Name | Gene Symbol | Relative Change |
|---|---|---|
| oxygen transport | | |
| myoglobin | Mb | 2.0 DOWN |
| Membrane protein | | |
| potassium voltage-gated channel | Kcne1 | 1.6 DOWN |
| Secreted Factors | | |
| chemokine ligand 12 | Cxcl12 | 1.5 DOWN |
| protease, serine,11 | Prss11 | 1.6 DOWN |
| periostin | Postn | 1.7 DOWN |
| alpha fetoprotein | Afp | 4.7 UP |
| wingless-related MMTV integration site 2 | Wnt2 | 1.5 UP |
| Transcription Factor | | |
| short stature homeobox 2 | Shox2 | 2.0 UP |
| elongation factor RNA polymerase II | Ell2 | 1.8 UP |
| Anti-oxidant | | |
| metallothionein 1 | Mt1 | 1.7 UP |
| metallothionein 2 | Mt2 | 2.0 UP |
| vanin 1 | Vnn1 | 1.6 UP |
| Metabolism | | |
| hydroxymethylbilane synthase | Hmbs | 1.6 UP |
| erythroid differentiation regulator 1 | Erdr1 | 1.5 UP |
| aldehyde dehydrogenase family 1, A1 | Aldh1a1 | 1.5 UP |
| Calcium signaling | | |
| visinin-like 1 | Vsnl-1 | 1.7 UP |
| Cell Structure | | |
| keratin complex 2, basic, gene8 | Krt2-8 | 1.9 UP |

**B**

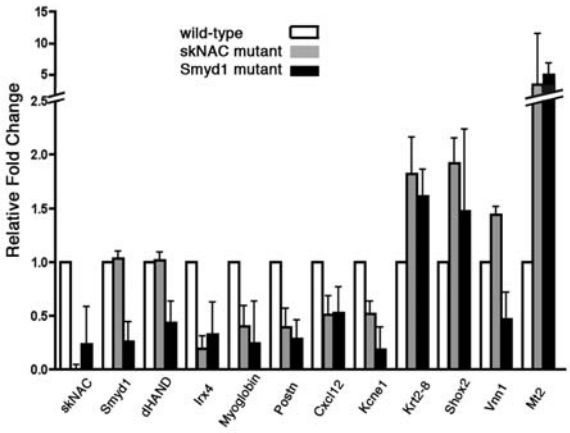


**Fig 4.6 The 17 characterized gene targets discovered in microarrays of the *skNAC* knockout were confirmed to be differentially expressed in both the *skNAC* knockout and the *Smyd1* knockdown. (A)** Differentially expressed genes in the *skNAC*–/– hearts. Microarray analysis was performed on three individual E11.5 hearts per genotype. Genes showing more than 1.5 fold change in expression are shown. **(B)** Quantitative real-time RT-PCR analysis for the hearts of *skNAC* and *Smyd1* mutants. Data are presented as mean of each genotype (for *skNAC*–/– analysis, n=3 and for *Smyd1* mutants, n=4). The error bars indicate the combined standard error.

## 4.3 *αMHC-cre* induced recombination of the *Smyd1*$^{Flox}$ locus occurs in the heart and gonad tissues and is embryonically lethal.

*αMHC-cre* transgenic mice (FVB) were obtained from the laboratory of Michael Schneider at Baylor College of Medicine. Recombination induced by *αMHC-cre* is specific to cardiac tissue in the mouse and to ventricular myocytes, but not in ventricular fibroblasts in primary cultures [76]. This transgene was later shown to be expressed in the ventricular and atrial myocardium as early as E8.5 [75]. The same breeding strategy described in the previous section was employed to create *Smyd1*$^{Flox/KO}$*;αMHC-Cre*$^{+/?}$ mice (BL6.FVB.129 hybrids) as well as the following breeders of the same background: *Smyd1*$^{Flox/+}$*; αMHC-Cre*$^{+/?}$ and *Smyd1*$^{KO/WT}$; *αMHC-Cre*$^{+/?}$. Once this combination of alleles was obtained, these mice were only crossed to each other, thereby maintaining a similar and inbred, but non-isogenic, hybrid population that is distinct from the *Nkx2.5-Cre* induced CKO line. Interestingly, several of the *Smyd1*$^{Flox/WT}$; *αMHC-cre*$^{+/?}$ breeders died prematurely between 3 and 7 months of age.

*Tissue Surveys on the Smyd1 Locus*

Like with the *Nkx2.5-Cre* line, we initially analyzed *αMHC-Cre* effeciency on DNA recombination of the *Smyd1* locus in heterozygous mice. With this line of *Cre* we saw no evidence of a reduced Flox band in any tissue (Fig. 4.7). However there was a strong Floxdel band in the heart as well as a band in the gonads (Fig. 4.7). Therefore, it appears that recombination is weakly occurring in the heart and also in gonads. Gonads often display basal expression of a slew of otherwise restricted genes. This particular *Cre* transgene may have integrated into a locus in which some basal expression is allowed in the gonads and therefore *Cre* can be expressed to induce some amount of recombination. Although *Cre* expression alone in the gonads should not affect *Smyd1* biology due to the

lack of endogenous *Smyd1* expression in the gonads, this could create error in genotyping assays. If recombination occurs in germ cells the Floxdel allele may pass on to offspring and would escape detection by our standard genotyping assay. If two of the other three potential alleles, $Smyd1^{KO}$, $Smyd1^{Flox}$, or $Smyd1^{WT}$, are detected, this possibility is ruled out. However, if only one allele is detected, it is assumed to be homozygous for that allele and, although $Smyd1^{KO/KO}$ and $Smyd1^{KO/Floxdel}$ are comparable, animals that are genotyped as $Smyd1^{WT/WT}$ or $Smyd1^{Flox/Flox}$ may actually be heterozygous for the $Smyd1^{Floxdel}$ allele. These animals would then be heterozygous for a *Smyd1* deficient allele. Neither of these genotypes would be selected for breeding, but may be selected for control animals in experiments.

*Embryonic Lethality of αMHC-Cre mediated Smyd1 CKOs*

Despite weak recombination, in the first 96 pups (9 litters), zero conditional knockouts survived to weaning age, while five mice that contained the same *Smyd1* alleles, but were *Cre* negative survived to weaning age. With the above mating scheme we would expect 3/16, or 18 total, to be $\alpha MHC\text{-}cre^{+/-}$; $Smyd1^{Flox/KO}$ and 1/16, or 6 of 96, to be $Smyd1^{Flox/KO}$ (Table 4.2). Therefore, the *Cre* negative genotype is surviving at normal ratios, but the *Cre* induced conditional knockout is lethal. Because we saw no reductions of litter size between birth and genotyping, it can be said that this conditional knockout is embryonically lethal.

Several embryonic ages were briefly analyzed. At E11.5, 2 CKO embryos out of 11 were observed and 1 of these was viable. With this mating we expected 25% CKOs, therefore this is not significantly different from the expected Mendelian ratio. At E13.5, 4 of 9 embryos were genotyped as CKOs and 2 of these were viable. With this mating, we expected 12.5% (1 of 9) CKOs, therefore we observed well above Mendelian ratios of

this combination of alleles. However, we saw expected ratios of each allele. The sample number is likely too low to be significant. At E16.5, Mendelian ratios were within the expected range, however, only one of four observed CKO embryos was viable. The *αMHC-cre* induced *Smyd1* knockout is embryonically lethal and displays a phenotype in some embryos by E13.5 and death is common by E16.5.

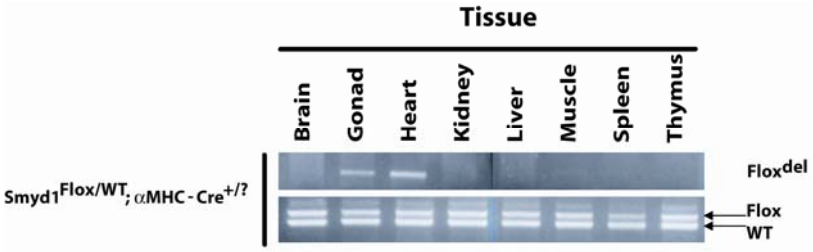


**Fig 4.7 *αMHC-cre* induced recombination of *Smyd1* occurs in the heart and gonads.** DNA was prepared from the tissues of a heterozygous *Smyd1* deleting mouse. Recombination was assayed in the upper panel and the positive control Flox/WT assay was assayed in the lower panel.

| Genotype | Expected Ratio | Observed Ratio |
|---|---|---|
| Smyd1 WT | 0.3750 | 0.52 |
| Smyd1 HET | 0.4375 | 0.48 |
| Smyd1 MT | 0.1875 | 0.00 |

**Table 4.2 *αMHC-cre* induced recombination of *Smyd1* is not viable**. *Smyd1*$^{WT/Flox}$*; αMHC-Cre* were crossed with *Smyd1*$^{WT/KO}$*; αMHC-Cre*. From the first 9 crosses we obtained an N of 96. Genotypes with two cardiac deficient *Smyd1* alleles are referred to as MT, genotypes with one cardiac functional allele are referred to as HET and genotypes with two cardiac functional alleles are referred to as WT. Litters were genotyped at weaning and observed numbers were recorded. There were no significant reduction in litter size between birth and weaning.

## Discussion

The cardiac phenotype of the *Smyd1* conventional knockout is not due to a hypomorphic allele of a *Smyd1*-neomycin fusion protein because the same result is obtained when the neomycin cassette is removed. The cardiac phenotype observed in the conventional knockout is due to the lack of Smyd1 in the cardiac chambers and likely the lack of Smyd1 in the ventricular chambers. The *αMHC-cre* induced knockout occurs weakly in the myotome of the atria and ventricles. While this knockdown is not efficient, it does cause heart defects which lead to lethality. The *Nkx2.5-cre* induced knockout is a little stronger and occurs a half-day earlier, but is restricted to the cardiac ventricles during embryogenesis. Although this knockdown does not exactly phenocopy the conventional *Smyd1* knockout, many of the same defects are observed. Unlike the conventional knockout, the *Nkx2.5-cre* induced conditional knockout does undergo ventricular septation, but this is process is deficient. There is also poor trabeculation, enlarged chambers, and cardiac edema. Therefore, the same processes are affected; it is likely that they are partially rescued due to partial expression of *Smyd1*. It is also possible that deletion at E8.0 is not early enough to inhibit the effects of *Smyd1*. A third possibility is that Smyd1 expressed in the cardiac atria or valves can act non-autonomously on the cardiac ventricles and this effect in the *Nkx2.5-cre* knockdown is responsible for the partial rescue of the phenotype.

While *Smyd1* is expressed in cardiac atria and valves, these conditional knockdowns do not address the role of *Smyd1* in these tissues. Co-analysis of the *skNAC* conventional knockout with the *Nkx2.5-cre* induced *Smyd1* knockout revealed that the role of *Smyd1* in the cardiac ventricles between E8.0 and E11.5 is closely related to a known binding protein, *skNAC*. The phenotype of this *Smyd1* knockdown was more

penetrant at this age, however the phenotypic *skNAC* knockouts phenocopied the *Smyd1* knockdown.

While it has been shown that *Smyd1* and *skNAC* interact at the protein-protein level (Personal Communication, L. Zhu and P. Tucker), it was surprising to find that there was a reduction of *skNAC* transcript in the *Smyd1* knockdown. This suggests that there is also a transcriptional effect of *Smyd1* on *skNAC*. Not surprisingly, transcript levels of *Myoglobin* were significantly decreased in both knockouts. Oxidative stress markers were also shown to be deregulated as would be expected with decreased levels of *Myoglobin*. The levels of markers were not quite as high in *Smyd1* knockdowns as in *skNAC* knockouts, however, this is expected as neither *skNAC* nor *Smyd1* are completely knocked out in the *Smyd1* knockdown. This suggests that the severity of the *Smyd1* phenotype is not simply because *Smyd1* depletion has a more complete effect, being upstream of *skNAC*, but rather that *Smyd1* affects *skNAC* independent pathways.

While previous studies have shown that *Smyd1* affects the transcript levels of *Hand2* and *Irx4*, *Hand2* was only affected in the *Smyd1* knockdown while *Irx4* was affected in both systems. This suggests that although *Smyd1* may affect *Hand2* via a *skNAC* pathway and *Irx4* is downstream of *Hand2*, *Smyd1* affects *Irx4* independently of *Hand2* and *skNAC*. This further emphasizes the role of *Smyd1* in *skNAC* independent pathways.

## Chapter 5. *MyoG-Cre* Induced Embryonic Skeletal Muscle *Smyd1* Knockouts are Embryonic Lethal with Few Observed Changes in the Transcriptome

*MyoG-Cre* mice were obtained from the Olson laboratory on the B6C3F1 strain, which is a first generation hybrid of C3H and C57Bl/6J. Recombination induced by *MyoG-Cre* is initiated in anterior somites at E9.0. By E9.5 expression is detected in all somites and is confined and maintained in muscle tissue throughout embryogenesis [77]. The same breeding scheme described above was employed to create $Smyd1^{Flox/KO}$; $MyoG\text{-}Cre^{+/?}$ mice (BL6.FVB.129.C3H hybrids) as well as the following breeders of the same background: $Smyd1^{Flox/WT}$; $MyoG\text{-}Cre^{+/?}$ and $Smyd1^{KO/WT}$; $MyoG\text{-}Cre^{+/?}$. However, in order to avoid breeding *Cre* to homozygosity we began crossing the above genotypes with *Cre* negative mice harboring the $Smyd1^{Flox}$ and $Smyd1^{KO}$ alleles. It later became apparent that $Smyd1^{Flox/Flox}$; $MyoG\text{-}Cre^{+/-}$ and $Smyd1^{Flox/KO}$; $MyoG\text{-}Cre^{+/-}$ displayed the same phenotype. In order to simplify the number of genotypes created and to create higher numbers of interesting genotypes, we mated $Smyd1^{Flox/Flox}$ to $Smyd1^{Flox/WT}$; $MyoG\text{-}Cre^{+/-}$. These mice were only crossed to each other, thereby maintaining an inbred, but non-isogenic, hybrid population that is similar, but isolated from the other cre induced CKO lines.

*Tissue Surveys on the Smyd1 Locus*

As described above, we analyzed *Cre* effectivity on genomic DNA recombination of the *Smyd1* locus in heterozygous mice. As with other tissue surveys, there was no detectable deletion of the Flox allele compared to the WT allele. However, there was a strong recombination band specifically in the skeletal muscle tissue (Fig. 5.1).

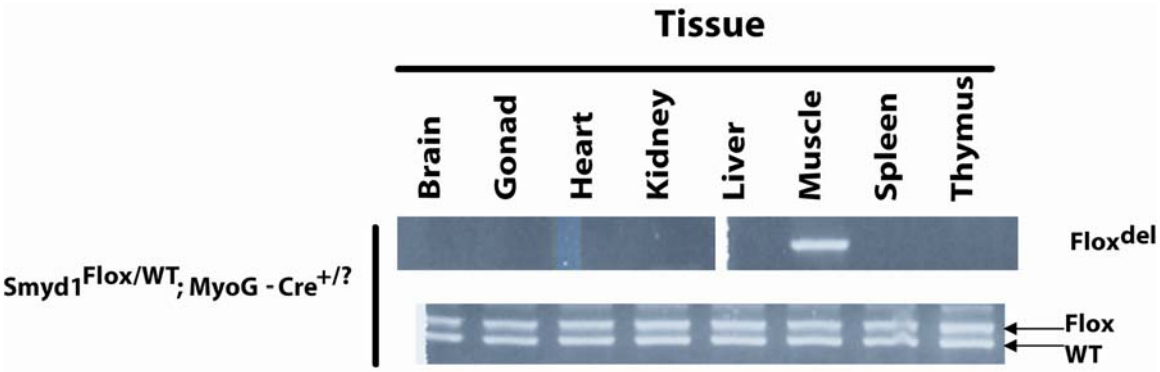


**Fig. 5.1** ***MyoG-Cre*** **induces recombination of** ***Smyd1*** **specifically in muscle tissue.** DNA was prepared from the tissues of a heterozygous *Smyd1* deleting mouse. Recombination was assayed in the upper panel and the positive control Flox/WT assay was assayed in the lower panel.

*Phenotype of MyoG-Cre mediated Smyd1 CKOs*

Surprisingly, of the first 350 pups weaned and genotyped, there were zero *Smyd1* conditional knockouts (Table 5.1). Many dead pups were found during and shortly after birth, but significant numbers of pups did not disappear between postnatal day 1 and weaning. The birth of a few litters was observed; conditional knockout pups were difficult births, likely because of an increased size due to extreme edema. These pups were often crushed during the birthing process. Those that were born intact did not show signs of movement or breathing. Genotypic ratios for these litters were analyzed between E9.5 and E17.5 and no significant variance from the expected Mendelian ratios were observed (Table 5.2). Fetuses were extracted at E18.5, an age at which they could be fostered and survive if viable. However, unlike their littermates, *MyoG-Cre* induced CKOs at this age made no attempts at locomotion or breathing. Therefore, these conditional knockouts are not dying simply due to birthing problems, but never show signs of life during late stage embryogenesis. Whether or not the heart was pumping at this age was not determined, but without locomotion these embryos could never survive. Therefore, this conditional knockout is perinatal lethal.

Embryos were analyzed from E11.5 through E18.5. At E12.5 embryos appear normal after sectioning (data not shown). Some embryos are outwardly abnormal as early as E13.5 and by E15.5 CKOs are easily discerned from littermates by eye. These pups displayed a very striking edema most visible in the back and shoulder area. In order to observe the skeletal muscle in these mice, pups at various ages were sectioned and compared to their littermates. At E15.5, there were obviously fewer myogenic cells, although different muscle groups were affected differently. The limbs of conditional *Smyd1* knockouts contained fewer cells than controls and correspondingly showed weaker in situ hybridization for a *Myogenin* probe, a marker of cells of myogenic

lineages (Fig. 5.2A-B). However, the pectoralis and the sternohyoid muscle were no longer represented in sections (Fig 5.2A-B). Haematoxylin and eosin (H&E) stained cross-sections of the scapula region shows that the surrounding muscle cells are less dense and less organized (Fig. 5.2C-D). Also, there is an appearance of very large multinucleated cells, which are likely myoblasts that have fused to become multinucleated, but have failed to elongate into myotubes (Fig. 5.2E-F). The tongue displays the same phenotype as the shoulder muscle (Fig. 5.2G-J). However, the masseter muscle appears totally normal (Fig. 5.2K-N).

To determine whether the apparent lack of cells and cell density is due to a decrease in proliferation, sections were analyzed by immunohistochemistry for *Ki67*, a protein that is present in all active stages of the cell cycle, but absent in the quiescent stages [78]. There does not appear to be significant differences in *Ki67* staining in muscle tissue (Fig. 5.3). However, it does appear that the leading edge of proliferation in the tongue epithelium is lagging (Fig 5.3C-D) and the amount of brown adipose tissue (BAT) is increased as well as more densely stained (Fig. 5.3A-B). While differences in tongue sections may be due to misshaped tongues, the excess brown adipose tissue can be confirmed in other sections (Fig. 5.2A-B). We have also shown that C2C12 siRNA knockdowns of *Smyd1* are richer in triglycerides, a product of white adipose tissue, than control cultures (personal communication, L. Zhu).

| Genotype | Expected Ratio | Observed Ratio |
|---|---|---|
| Smyd1 WT | 0.30 | 0.38 |
| Smyd1 HET | 0.50 | 0.62 |
| Smyd1 MT | 0.20 | 0.00 |

**Table 5.1 Skeletal Muscle specific knock-down of *Smyd1* is lethal prior to weaning.** Various *Smyd1*$^{Flox/KO/WT}$*; MyoG-Cre* crosses were made. From these crosses we obtained an N of 350. We expect 20% of pups to be homozygous for *Smyd1* deficient alleles, 50% of pups should be heterozygous and 30% should be wild type. Litters were genotyped at weaning and observed numbers were recorded. There were no significant reduction in litter size between birth and weaning.

| Age (DPC) | Smyd1 MT | Smyd1 HET | Smyd1 WT | Total |
|---|---|---|---|---|
| 9.5 | 5 | 11 | 6 | 22 |
| 12.5 | 4 | 9 | 10 | 23 |
| 13.5 | 2 | 3 | 3 | 8 |
| 14.5 | 0 | 4 | 0 | 4 |
| 15.5 | 6 | 11 | 8 | 25 |
| 16.5 | 6 | 5 | 2 | 13 |
| 17.5 | 2 | 4 | 3 | 9 |
| Totals | 25 | 47 | 32 | 104 |
| | | | | |
| Expected Ratio | 0.25 | 0.5 | 0.25 | |
| Observed Ratio | 0.24 | 0.45 | 0.31 | |

**Table 5.2 Skeletal Muscle specific knock-down of *Smyd1* does not induce resorption.** *Smyd1*$^{WT/Flox}$*; αMHC-Cre* were crossed with *Smyd1*$^{WT/KO}$*; αMHC-Cre*. *Smyd1*$^{WT/WT}$; *αMHC-Cre* is referred to as WT and *Smyd1*$^{Flox/KO}$; *αMHC-Cre* is referred to as MT. Both *Smyd1*$^{KO/WT}$ and *Smyd1*$^{Flox/WT}$ with the *Cre* are referred to as HET.

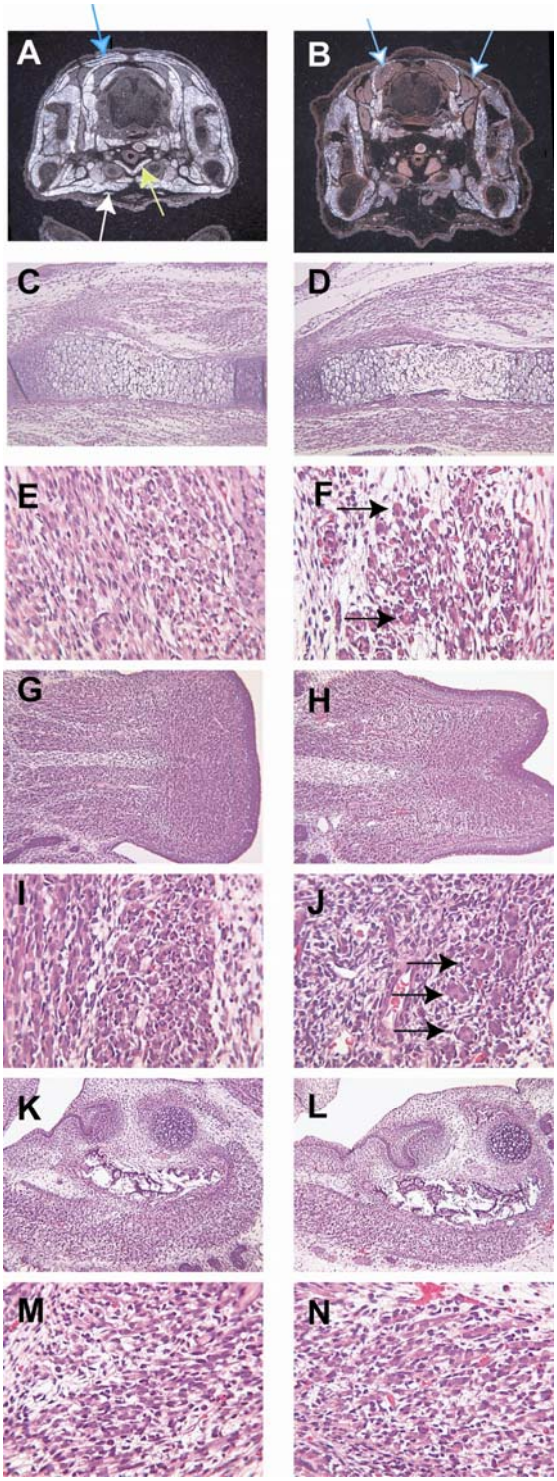


**Fig. 5.2 Skeletal muscle development in the *MyoG-Cre* induced *Smyd1* CKO is significantly derailed by E15.5.** A-N) H&E stained transverse sections. *Myogenin* in situ hybridization (A, B) show that there are much fewer muscle cells in the KO (B) compared to the WT (A). It also demonstrates that not all muscles are affected equally, e.g. the pectoral (white arrow) and sternohyoid (yellow arrow) muscles are unrepresented in the KO, the dorsal muscles (solid blue arrow) are significantly diminished, but other muscles, like limb muscles are clearly less affected. There also appears to be an excess of brown adipose tissue (blue and white arrow) in the KO. Failure of the muscles to develop can also be seen with pure H&E stains (C-N). C) WT Scapula, 10x, D) KO Scapula, 10x. E)WT scapula, 40x F) KO scapula 40x. Here the cells do not appear to be elongating along the same plane as cells are doing in the WT. Also many cells are forming "myo-balls" in which they appear multinucleated, but are failing to elongate in the KO (black arrow) (G-J). G) WT Tongue, 10x. H) KO tongue, 10x. I) WT tongue 40x J) KO tongue 40x. "Myo-balls" are also evident in the tongue. However, there do not appear to be problems in masseter muscle development (K-N). K) WT masseter 10x L) KO masseter 10x M) WT masseter 40x N) KO masseter 40x.

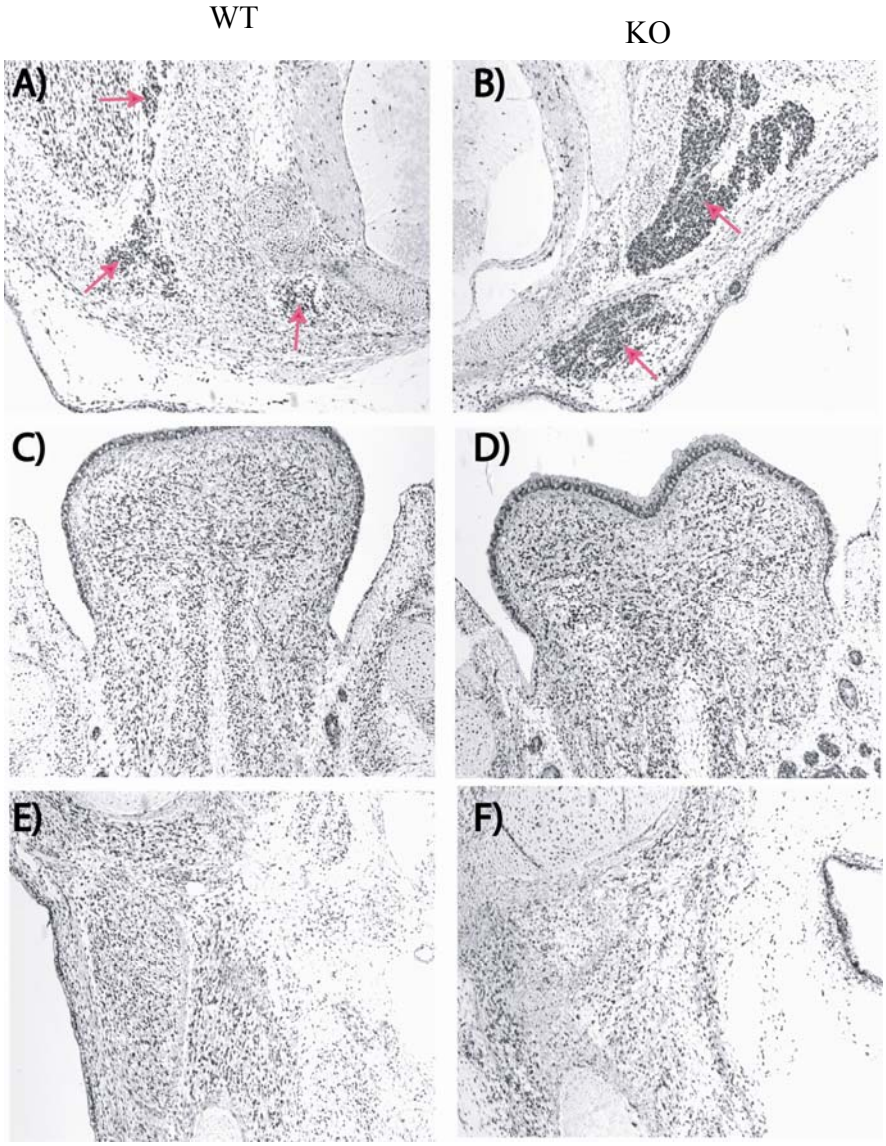


**Fig 5.3 Proliferation of Skeletal Muscle in *MyoG-Cre* induced *Smyd1* knockouts is unaffected, but there are non-autonomous effects on BAT.** Brown adipose tissue (pink arrows) in the dorsal regions (A, B), shows less *Ki67* staining in the WT (A) than the KO (B). Also, the tongue (C, D) shows proliferation differences in the epidermal region at the tip. This line of proliferation appears to lag in the KO (D) compared to the WT (C). There do not appear to be large differences along the humerus (E, WT; F, KO).

*Molecular Changes*

In an attempt to identify pathways in which *Smyd1* functions, I performed RT-PCR to analyze the transcriptome. *Smyd1* is a putative histone methyl transferase and has displayed HDAC dependent effects on transcription in *luciferase* assays [10]. Therefore, it is likely an important transcriptional factor. We decided to analyze the transcript pool just prior to observing the phenotype in an attempt to preclude secondary effects of *Smyd1* on other transcription factors from our data sets. By RT-PCR at E13.5, there is an obvious inverse correlation between *Smyd1* transcripts and *Cre* transcripts (Fig. 5.4). Because the myogenic program was significantly derailed, the first analysis included the myogenic regulatory factors, including *Myogenin*, *MyoD1*, *Myf5*, and *Myf6*. Surprisingly none of these factors appear to be affected at the transcriptional level (Fig. 5.4). I have also analyzed *UCP-1* and *Cidea*, markers of brown adipose tissue formation and found no differences (data not shown). Most transcripts that were analyzed, including *skNAC*, *Myoglobin*, *TpnIs*, and *MyHCIIa*, are not affected.

*TNFα* is known to have a role incorporating fatty acids into muscle tissue while there is a direct correlation between *IL6* and fatty acid oxidation by skeletal muscle [79]. Both of these are inflammatory cytokines that are involved in maintenance and differentiation of skeletal muscle. *TNFα* plays a bipolar role in muscle regulation; in the absence of *TNFα*, myogenic differentiation is blocked by failure to activate *p38* [80], however *TNFα* is also known to play a role during cachexia, or wasting atrophy [81]. Over expression and knockdown experiments of *IL6* show that myogenic markers, *Myogenin*, *α-actin*, and *p21*, are directly correlated to *IL6* levels through both the *p38*/*MAPK* and *NFκB* pathways [82]. Although *TNFα* levels are highly variant in these mice, it does not appear to be dependent on levels of *Smyd1* transcript (Fig. 5.5). *IL6*

levels are tightly correlated to *Smyd1* levels, however, opposite effects are observed in male versus female mice (Fig. 5.5).

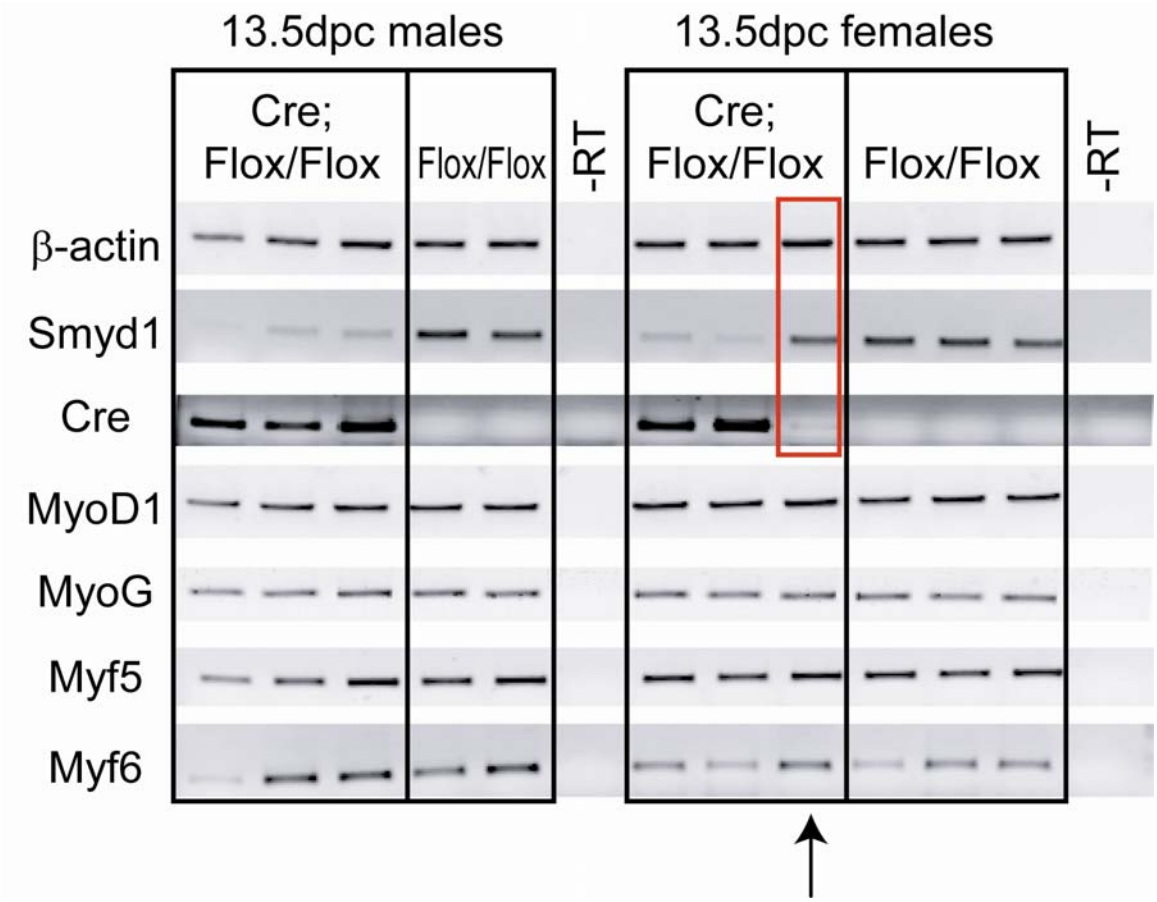


**Fig 5.4 Skeletal Muscle *Smyd1* deficiency does not affect transcription levels of MRFs (Myogenic Regulatory Factors).** Limbs obtained from E13.5 embryos were homogenized in "Trizol" and genotyped for *Cre*, *Smyd1*, and gender. Appropriate samples were used for reverse transcription and then RT-PCR. Some samples that are genotypically *Cre* positive do not express well and therefore *Smyd1* levels are not highly reduced (arrow and red box).

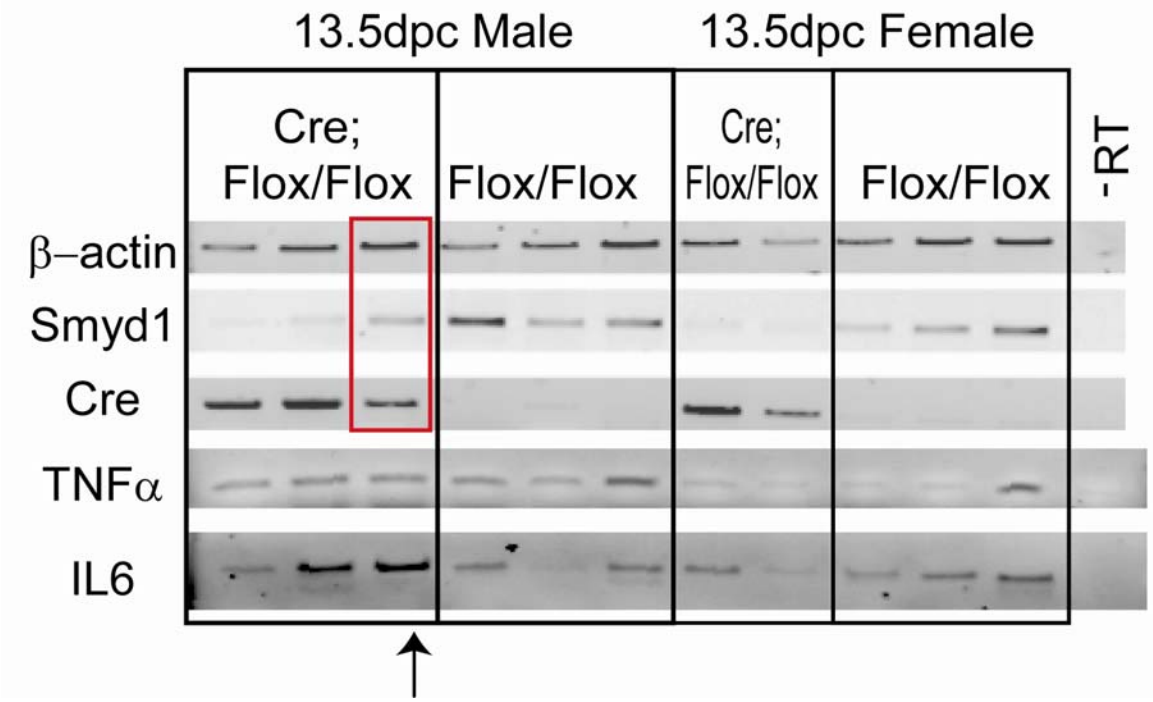


**Fig 5.5 Transcript levels of *TNFα* and *IL6* vary in *Smyd1* deficient skeletal muscle.** Limbs obtained from E13.5 embryos were homogenized in "Trizol" and genotyped for *Cre*, *Smyd1*, and gender. Appropriate samples were used for reverse transcription and then RT-PCR. Some samples that are genotypically *Cre* positive do not express well and therefore Smyd1 levels are not as highly reduced (arrow and red box). *TNFα* transcripts do vary, but this is more dependent on gender than on *Smyd1* levels. *IL6* transcripts vary dependent on Smyd1 levels, although different effects are observed in different genders. *Smyd1* deficiency in males leads to increased levels of *IL6* whereas *Smyd1* deficiency in females leads to slightly decreased levels of *IL6*.

*MyoG-Cre induced CKO Adult Survivors*

After several years of mating these mice and genotyping over 190 litters, I observed absolutely zero living post-natal *MyoG-cre* induced *Smyd1* CKOs on this non-isogenic background. However, two of four litters produced by one particular mother-son cross have yielded pups in which *MyoG-cre* induced *Smyd1* CKOs survive and appear outwardly normal into adulthood. With this particular cross *Smyd1* CKOs do not survive at normal Mendelian ratios, however crosses from their progeny exceed expectations of *Smyd1* CKOs (Fig. 5.6).

Initially these mice were sacrificed and tissues were harvested for RNA and DNA. DNA analysis revealed that a recombination event was indeed occurring at the *Smyd1* locus and specific to muscle; however, this recombination was not complete (Fig. 5.7). It is obvious that this recombination event did affect *Smyd1* expression at the RNA level (Fig. 5.8). On occasion I have observed embryonic individuals that were genotypically *Smyd1* CKOs, but that have had normal *Smyd1* levels. When *Cre* expression is analyzed, these embryos have expressed *Cre*, but not as highly as other age matched embryos (Fig. 5.4 and Fig 5.5). It is possible that the random locus into which this *Cre* transgene integrated is affected by repressors or enhancers that vary in different strains of mice. Because the mouse lines I am working with are not isogenic, over time this particular branch of this line of mice could have either lost or gained a genetic modification or homozygosity of a modification that delays *Cre* expression. If this explanation is accurate, it maintains that *Smyd1* is necessary for skeletal muscle development, but indicates that *Smyd1* is not essential for maintaining normal function in uninjured, developed skeletal muscle. Delayed deletion of *Smyd1* will rescue the skeletal muscle *Smyd1* conditional knockout mice.

Data obtained from siRNA knockdowns of *Smyd1* in C2C12 has suggested that *Smyd1* has an effect on *Calcineurin* regulation of slow fiber formation (Personal Communication, L. Zhu). Adult versions of the *Smyd1* muscle CKO allowed us to not just dissect and analyze limb muscle, but to further dissect and analyze specific muscle groups of known fiber types. The soleus muscle is composed primarily of slow muscle fibers, the tibialis anterior is primarily fast muscle fibers, and the gastrocnemius is a mixture of the two fiber types. Isolates of these three muscle groups have been analyzed by RT-PCR to determine if there is a fiber-type switch. It appears that the *Smyd1* knockdown is not as effective in the soleus muscle as in the gastrocnemius and the tibialis anterior. This corresponds to levels of *Cre* expression (Fig. 5.8). Although the soleus represents a weak knockdown of *Smyd1*, there is a surprisingly strong effect on some markers. As we would predict, the slow muscle markers, *MyHC-1* and *Tpn1s*, are not detectable in the fast tibialis anterior muscle. Although the hypothesis was that *Smyd1* had a positive regulatory role on slow muscle formation, the slow muscle markers, *Tpn1s*, *MyHC-1*, and *MyHCIIa* do not appear to be affected in either the soleus or the tibialis anterior. However, the glycolytic fast muscle myosin, *MyHC-IIb*, is up regulated in the slow soleus. Surprisingly, the intermediate oxidative-glycolytic Myosin, *MyHC-IIx* is decreased in the slow soleus of the *Smyd1* knockdown mouse, but increased in the mixed gastrocnemius and fast tibialis anterior (Fig. 5.8).

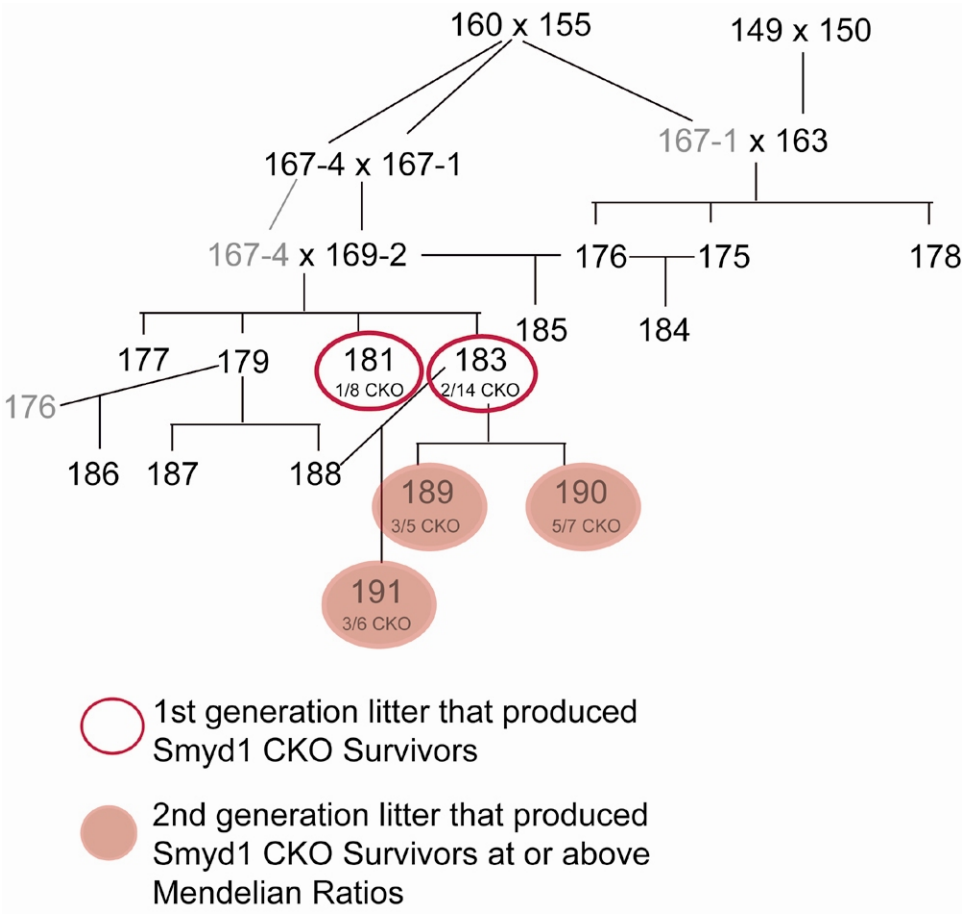


**Fig 5.6 A heritable change has occurred in a portion of the skeletal muscle specific *Smyd1* CKO mouse line allowing CKOs to survive.** Mice were bred in a non-gender specific manner by crossing *Cre*; *Smyd1*$^{Flox/Flox}$ mice to *Smyd1*$^{Flox/Flox}$ mice. Two of four litters from a mother-son cross yielded a few conditional knockout pups. Crosses from these lines produce CKO pups at or above expected Mendelian ratios. Numbers represented are litter numbers or litter number- individual number. Gray numbers are duplications necessary to show breeding with multiple partners.

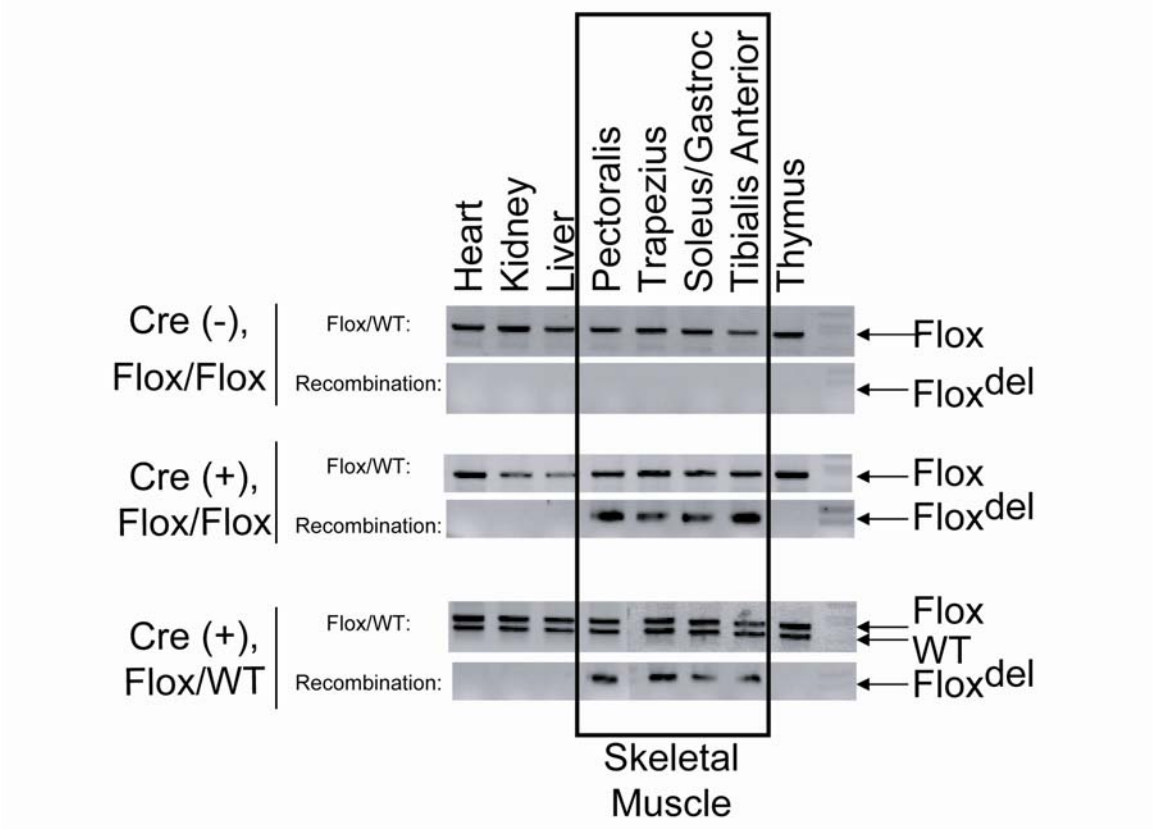


**Fig 5.7 Recombination is occurring specifically in the skeletal muscle of the *MyoG-Cre* positive *Smyd1*$^{Flox/Flox}$ survivors.** Mice were sacrificed and DNA was extracted from tissues and analyzed via PCR. A recombination product is occurring specifically in skeletal muscle when both the *Cre* transgene and *Smyd1*$^{Flox}$ allele are present.

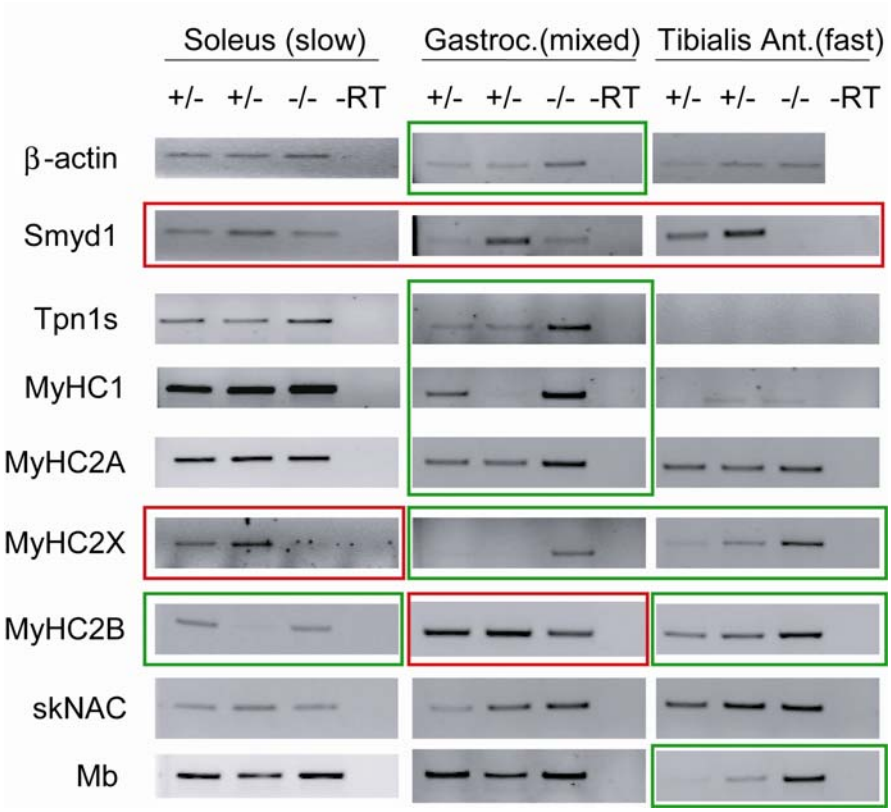


**Fig 5.8 Slow/Fast fiber type composition is altered in adult *Smyd1* deficient skeletal muscle.** Mice were sacrificed and tissues were homogenized in "Trizol" for RNA extraction, reverse transcription, and finally, RT-PCR analysis. Proper isolation of slow vs. fast muscle is indicated by the lack of troponin1 slow and *MyHC1* in the fast muscle. The heterozygous mouse represented on the left in all three columns does not express *Smyd1* as highly as the heterozygous mouse represented in the middle. The *Smyd1* knockdown is not as efficient in the soleus muscle compared to the others. However, expression levels of *MyHC2X* and *MyHC2B* are altered in all three muscles. In slow muscle *MyHC2x* is directly correlated to Smyd1 expression whereas in mixed and fast, *MyHC2x* is oppositely correlated to *Smyd1* levels. *MyHC2B* is oppositely correlated to *Smyd1* in slow and fast tissues whereas it appears to be directly correlated to *Smyd1* levels in mixed tissue.

*Discussion*

The phenotype of the *Smyd1* conditional knockout is quite striking because only one previous single gene mouse knockout has caused such a severe phenotype in skeletal muscle tissue. This phenotype was due to *Myogenin* deficiency, with which myoblasts appear normal, but differentiation from myoblasts to myotubes is blocked. The *Smyd1 MyoG-cre* induced knockout appears similar to the *Myogenin* knockout in many ways: while there are many myogenic cells, the appearance of myo-balls suggests that these cells are not properly elongating or differentiating into myotubes. Because myogenic regulatory factors, or MRFs, have long been considered responsible for proper muscle development, these genes, including *Myogenin*, *MyoD*, *Myf5*, and *Myf6*, we originally hypothesized that the lack of *Smyd1* disrupts myogenesis by affecting expression of one or several of these transcription factors. The most likely explanation would be that *Smyd1* serves as a muscle specific histone methyltransferase and globally affects the transcriptome. However, there was no obvious change at the transcript level of MRFs or of many other genes, which is further evidenced by the lack of changes in microarray data set (data not shown). The lack of changes in the transcriptome, suggests that perhaps the role of *Smyd1* involves posttranslational effects. While we have shown that *Smyd1* can methylate at least one non-histone protein, there is no evidence that any of the MRFs are methylated or otherwise posttranslationally modified in a manner which could be affected by *Smyd1*.

It is possible that *Smyd1* is involved in determining whether a cell will become muscle or adipose, involved in cross-talk with adipogenic energy stores, or involved in muscle absorption of nutrients. For example, adipose tissue and muscle tissue share similar precursors, it has been well documented that both primary muscle cell cultures and muscle cell lines can be induced to differentiate into adipose cells if given the

appropriate signals [83]. Therefore, it is likely that there are molecular pathways controlling this fating process *in vivo* and possible that *Smyd1* is involved in this pathway. Likewise, if *Smyd1* deficient muscle is incapable of signaling the break-down of BAT to release the metabolites, BAT would accumulate faster than in control embryos. Neither of these hypotheses would explain why the proliferation marker, *Ki67*, stained BAT more densely; however an alternate explanation is that *Smyd1* deficient muscle is unable to absorb nutrients. In this case, the excess nutrients in the system could be adsorbed by the brown adipose, causing an abnormal proliferation and expansion of BAT.

We have also observed that not all tissues are equally affected. Trunk muscles are nearly ablated, forelimb muscles are disorganized and improperly differentiated, but muscles like the masseter are normal. We have observed that muscles appear normal as late as E12.5, post migration from the somatic myotome. However, it is possible that the differential effect occurs earlier than the visible effect and that there is an essential molecular event involving *Smyd1* prior to this migration.

The genetic variation created by crossing four inbred strains as well as the, until recently, 100% penetrance suggested that *Smyd1* was required for muscle development in all mice and not in one or two specific backgrounds. This idea was recently brought into question when adult skeletal muscle conditional knockouts were obtained. Although the first generation of survivors did not present at expected Mendelian ratios, subsequent generations have. Therefore, it is likely that a genetic change occurred allowing mice to survive. While it is possible that this line of mice gained a modifier that no longer requires *Smyd1* expression, I hypothesize that this is not the case and that *Smyd1* is indeed essential in all backgrounds. This hypothesis is supported by previous data in which embryos are genotypically *Cre* positive, but expression of *Cre* is very low and

likewise expression of *Smyd1* is unaffected. In other words, the random genetic change has delayed *Cre* expression allowing *Smyd1* to be expressed through the critical stages.

The adult conditional knockouts will allow many interesting experiments that the perinatal lethal line would not. The excess fat observed in conditional knockouts suggests that there may be a defect in energy usage in *Smyd1* deficient muscle. In skeletal there are four types of myosin heavy chains, a slow oxidative *MyHC1*, a fast oxidative *MyHC2A*, a fast intermediate *MyHC2X*, and a fast glycolytic *MYHC2B*. The ratios of these in particular muscle groups are influence by aerobic exercise, the more a specimen exercises, the more oxidative *MyHCs* will be present. The type of *MyHCs* present dictates the energy pathways a muscle can use in order to utilize energy to contract. Early analysis of adult muscles has attempted to determine whether *Smyd1* expression affects fiber type and therefore energy usage. While this data does not completely coincide with data obtained from C2C12 siRNA knockdowns of *Smyd1*, it does support the idea that the lack of Smyd1 does affect the ratios of oxidative and glycolytic myosins. For the first time, this suggests that *Smyd1* does have an effect in adult skeletal muscle.

# Chapter 6. Adult Inducible *Smyd1* Knockouts Suggest that *Smyd1* Functions in Muscle Maintenance Pathways as well as Developmental Pathways.

## 6.1 *αMHC-cre-PR1* mediated *Smyd1* adult heart conditional knockouts suffer premature lethality, likely from chronic heart failure.

Several inducible versions of transgenic recombinases have been created that are controlled via drug administration. For example, *Cre* fusion proteins have been used which require interaction with drugs in order to enter the nucleus. Expression of these fusion proteins are easily directed by known tissue specific promoters, but the enzymatic action of *Cre* cannot occur until it is co-localized with DNA in the nucleus. We obtained *αMHC-Cre-PR1* transgenic mice (FVB) from the laboratory of Michael Schneider at Baylor College of Medicine. This transgenic fusion protein is expressed in a cardiac specific manner and is fused to a mutated progesterone ligand, PR1, which should only enter the nucleus by binding a synthetic progesterone receptor called RU486. The Schneider lab showed that recombination is specific to the heart and is induced by intraperitoneal injections of RU486 with minimal background until mice are 6 weeks of age [84].

Initial studies used the breeding scheme described above that avoids breeding the *Cre* transgene to homozygosity and creates $Smyd1^{Flox/KO}$; $\alpha MHC\text{-}Cre\text{-}PR1^{+/-}$ mice (BL6.FVB.129 hybrids), and $Smyd1^{Flox/KO}$ and $Smyd1^{Flox/Flox}$ mice as controls. It later became apparent that both CKO genotypes displayed the same phenotype. In order to simplify the breeding scheme, we mated $Smyd1^{Flox/Flox}$ mice to $Smyd1^{Flox/Flox}$; $MyoG\text{-}Cre^{+/-}$ mice to yield 50% CKOs and 50% Flox/Flox controls. These mice were only crossed to each other, thereby maintaining an inbred, but non-isogenic, hybrid population that is similar, but isolated from other CKO lines.

*Tissue Surveys on the Smyd1 Locus*

Recombination assays were performed at 4 weeks of age and recombination was detected only in hearts of mice that had received RU486 injections for 5 days, but not in mice that had received corn oil injections. However, as before, there was no decrease in the Flox allele compared to the WT allele (data not shown). Therefore recombination was not complete.

*Phenotype of αMHC-Cre-PR1 mediated Smyd1 CKOs*

Despite the fact that surveys suggest recombination was specific to the heart in the presence of RU486, mice that received either RU486 in corn oil or corn oil alone were symptomatic and frequently died. Mice carrying the CKO genotype, but that had not received drugs began dying at three to four months of age. DNA was harvested from individual tissues during autopsy of these mice and used to show that recombination was occurring and that it was heart specific (Fig. 6.1). These mice, however, were not heterozygous for $Smyd1^{WT}$. Therefore, the Flox/WT assay was not an acceptable positive control. Surprisingly, these mice displayed cardiac specific recombination despite the absence of RU486. The great variation in levels of the PCR recombination product could signify a great variation in the ability of *Cre* to induce recombination or could be due to variable DNA samples as there was no positive control.

I observed untreated mice of various genotypes from this lineage until 1 year of age, at which point they were sacrificed if they had not already died naturally. Mice that were sacrificed for experiments were included in the study up until the time point they were sacrificed, at which point they were censored. In the six control genotypes, including $Smyd1^{KO/Flox}$, $Smyd1^{Flox/Flox}$, $Smyd1^{KO/WT}$, $Smyd1^{KO/WT}$; $\alpha MHC\text{-}CrePR1^{+/-}$,

*Smyd1*$^{Flox/WT}$, or *Smyd1*$^{Flox/WT}$*; αMHC-CrePR1*$^{+/-}$, over 75% of mice survived through the end of the study. However, only 4% of CKO mice, *αMHC-CrePR1*$^{+/-}$*; Smyd1*$^{Flox/Flox}$ or *αMHC-CrePR1*$^{+/-}$*; Smyd1*$^{Flox/KO}$, survived (Fig. 6.2). The average age of death in the CKO population was 196 days and the median age was 189 days (Fig. 6.3). Although the mean and median are very close, the distribution does not appear to be normal. Data from males and females were analyzed separately. The average female death occurred at 188 days, with a median of 193 days. The average male death occurred at 201 days, with a median of 186. The shape of the female curve had a left handed tail, whereas the shape of the male curve had a right handed tail (Fig. 6.4). Although this trend is obvious when analyzing the data, there is no obvious analysis that leads to statistically relevant numbers. While the differences in timing of death could be due to the effect of *Smyd1* on different pathways in opposing genders, I believe that the most likely explanation is a difference between the genders in the ability of the *Cre-PR1* fusion protein to enter the nucleus at different ages and knockdown *Smyd1* levels in the absence of the synthetic progesterone receptor, RU486. Regardless of the cause, all experiments were both age and gender matched to ensure that results would not be masked due to a possible difference in pathology between the two genders.

Despite this separation, there was still a lot of variability within each group for onset of disease and phenotype. Therefore, determining which ages to analyze was not a straight forward task. Initial autopsies and transcript analysis was performed at early stages, mid stages and late stages in order to determine whether the loss of *Smyd1* was occurring at early time points or whether it was likely that mice died shortly after the loss of the gene. Based on the Kaplan-Meyer plot (Fig 6.2), I chose to use mice at 16 weeks, before death occurred, at 24 weeks, an age to which approximately half of the mice survived, and 32 weeks, an age at which mice could still be obtained with some

confidence. Hearts for these animals were homogenized in *"Trizol"* and RNA isolates were reverse transcribed to analyze *Smyd1* levels (Fig. 6.5). There is a strong knockdown of *Smyd1* at 16 weeks, which continues to decrease through 32 weeks. Therefore, animals are not dying due to an acute loss of *Smyd1*, because there is a strong knock down well before the average age of death. Therefore, it is likely that the loss of *Smyd1* is causing chronic heart disease.

Although the age of phenotype onset was not predictable from mouse to mouse, CKO mice became outwardly distinguishable from their littermates before death. Common phenotypes included obesity, wasting, hunched backs, low range of motion, beady eyes, and quivering. Upon dissection, phenotypic CKO mice typically had lower body weights, but increased heart weights (Fig. 6.6). Although the increase percentage of heart weight was not statistically relevant, it is clear that the standard deviation of the conditional knockout data is increased (Fig. 6.6), suggesting that either body weight or heart weight may have a significant change. Variability, and therefore standard deviation, may be increased because both obesity and wasting was observed as a general phenotype. When the number of variables was reduced to analyze only heart weight, there was a significant increase in heart weight in Smyd1 conditional knockout mice (Fig. 6.7). Sectioning the hearts shows that the diameter of the chambers is increased, however the gross muscle structure does not appear to be affected (Fig. 6.8). An atrial thrombus was detected in the phenotypic mouse sectioned, although it is unclear if this is representative of all phenotypic CKOs.

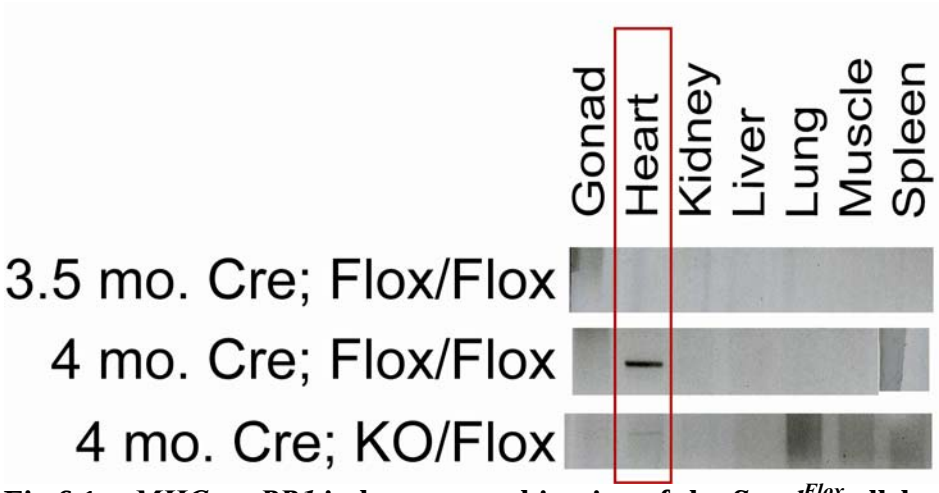


**Fig 6.1** ***αMHC-crePR1*** **induces recombination of the** ***Smyd*$^{Flox}$** **allele at varying levels specific to the heart.** Tissues were harvested from mice that had died naturally with no drug administration. DNA extractions show that recombination had occurred specifically in the heart in at least two of the three mice.

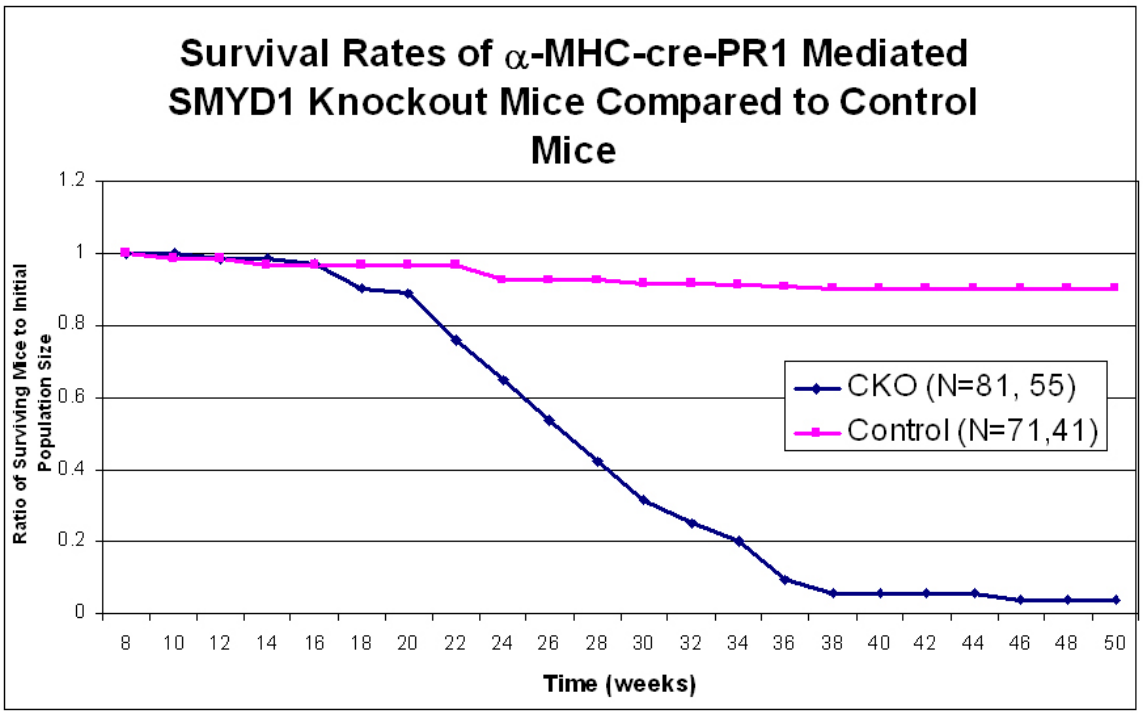


**Fig. 6.2 Mice with *αMHC-cre-PR1* mediated *Smyd1* deficiency have reduced life expectancies.** Mice were left untreated and observed for one year. Those that were removed from the untreated population and subjected to experimentation were censored. Therefore the first value for N indicates the initial population size and the second value for N reflects the final population size, the N at 50 weeks. Mice that died naturally at a midpoint were included in the total population size throughout the experiment.

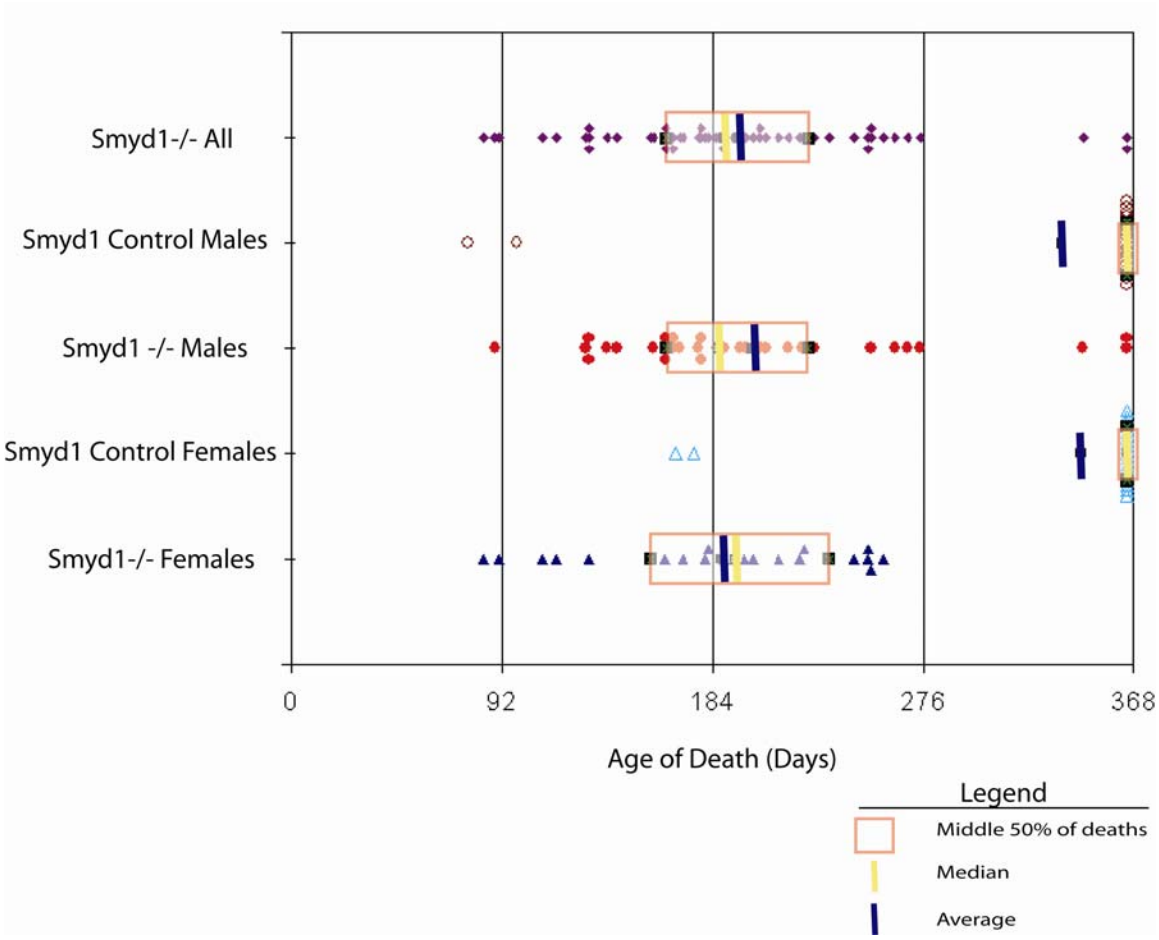


**Fig 6.3 The total population shows a normal Gaussian distribution for time of death, but when divided based on gender, the curves shift in opposite directions.** Each point represents one mouse fatality. Distribution of those clustered at the far right cannot be analyzed as they were artificially terminated at this point. The boundary between the first and second quartile is denoted by the left side of the box, the boundary between the second and third quartile is denoted by the yellow median line, and the boundary between the third and fourth quartile is denoted by the right side of the box. The average is also noted. Normal distribution is typically detected when both the average and median break the box into two equal segments. The total CKO population (purple) appears to have a normal distribution, however the male CKO population (red) peaks on the left with a right hand tail and the female CKO population (blue) peaks on the right with a right hand tail.

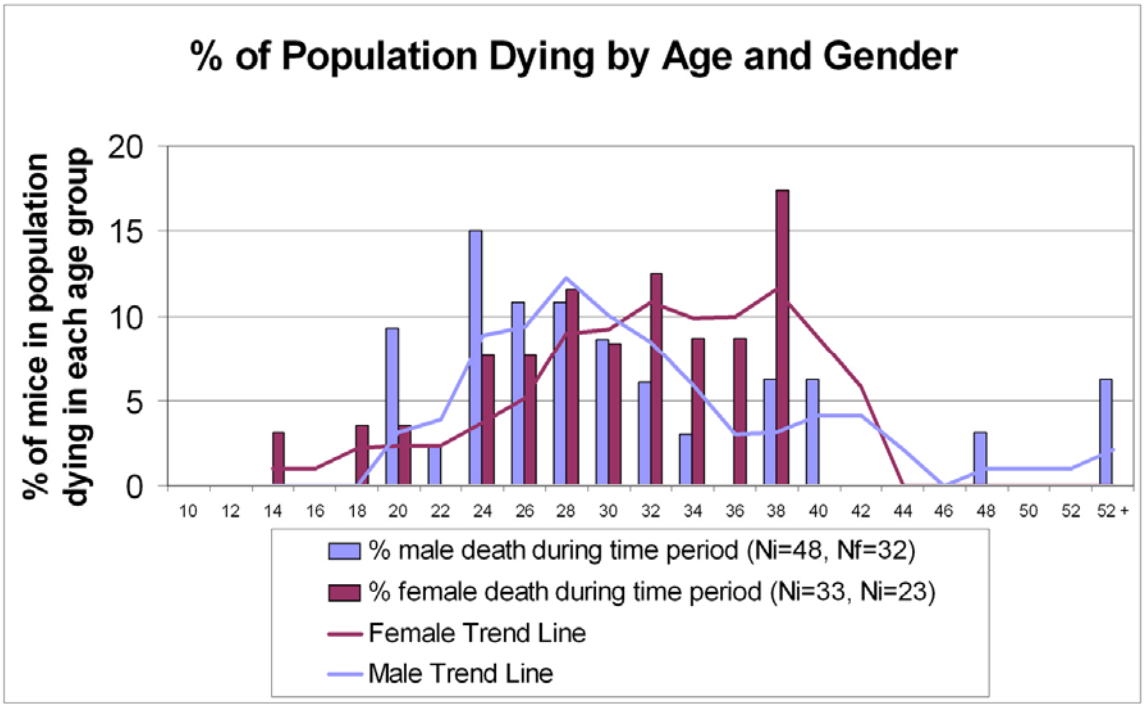


**Fig 6.4 The curves associated with male or female deaths do not show the same pattern.** A histogram was created for conditional knockouts of each gender based on two week intervals of age of death. Percentage was calculated based on the raw number of deaths and the N during that interval. Because N changes between intervals, percentages of each interval do not add up to 100%. Trend lines were then overlaid on the histogram to examine the shape of the curve. The highest percentages of males die at mid-stages, 24 weeks, whereas the highest percentages of females die at late stages, 38 weeks. Strikingly, no female has survived past 38 weeks whereas several males have survived past 52 weeks, the end of the study.

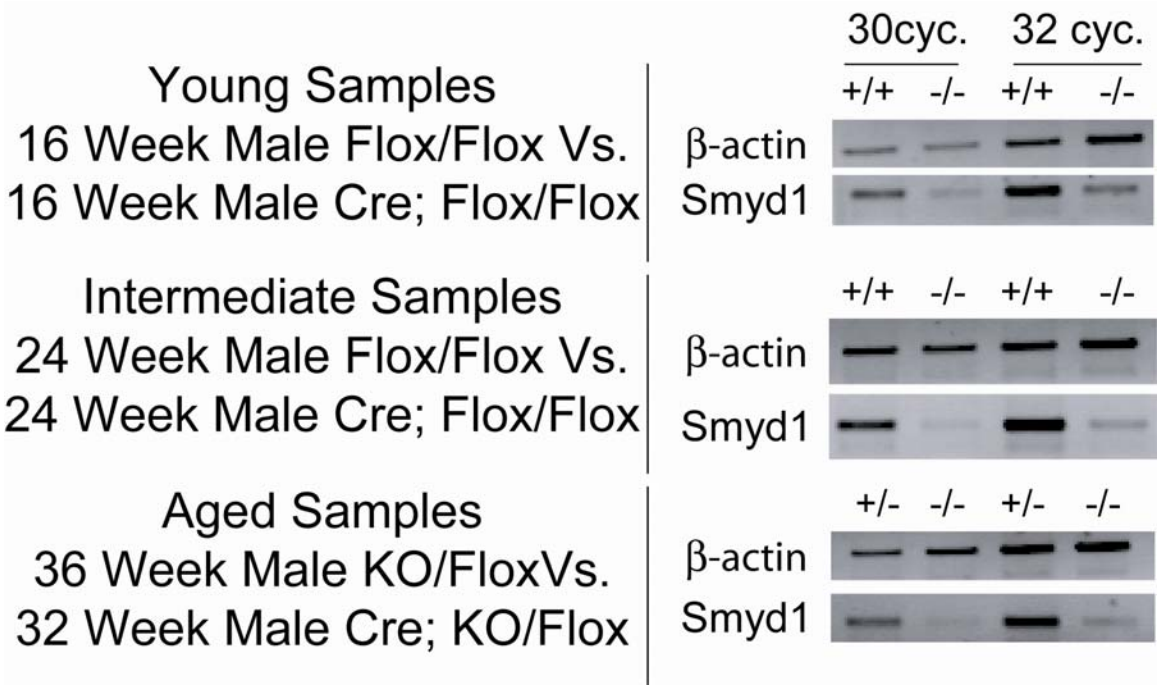


**Fig 6.5** ***Smyd1*** **is significantly knocked down before death occurs.** In order to compare the timing of death to the timing of the knockdown, animals were sacrificed to represent early, middle, and late stages of the plot depicted in Fig 6.2. At 16 weeks there is a *Smyd1* knockdown comparable to the knockdown seen in all other CKO systems. By 24 weeks this knockdown is very striking. Assays were done with multiple cycle number to ensure that conditions were in linear range and the band intensities were not maxed out.

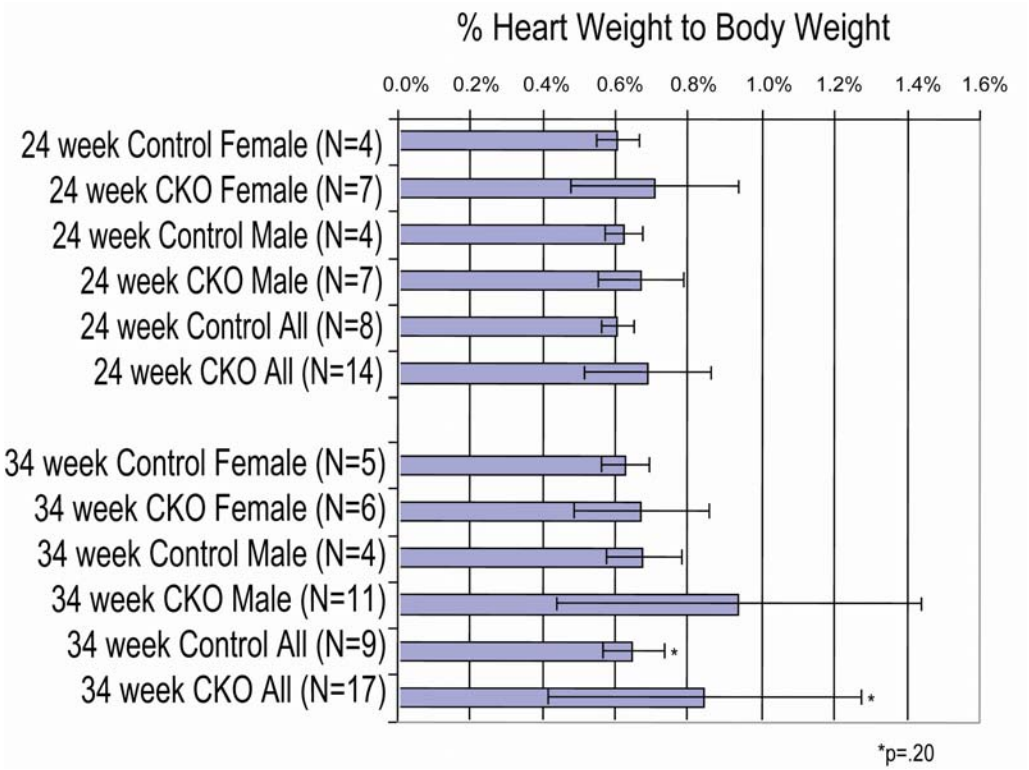


**Fig 6.6 The percent heart weight is not significantly changed, but is deregulated.** Mice were sacrificed in eight different groups based on age give or take 1 week (23 to 25 weeks and 33 to 35 weeks), gender and genotype. Prior to sacrifice mice were weighed. In most animals the weight of the atria are negligible. To allow for better blood removal, only the ventricles were isolated and weighed. Average percentage is reported and error bars represent 1 standard deviation. At 24 weeks, there is not a significant percentage change, but the error bars are increased. It does not seem that this is due to an increased N as pooling males and females doubles the N, but does not affect the error bars. Although it is not significant, there is a change in average weight of hearts in 34 week old males. When genders are pooled, this approaches significance. Again the error bars are greatly increased, but are not affected in the wild type by pooling genders. This suggests that regardless of whether there is a significant difference there is a deregulation of one or both of the variables. With a higher N, this could become significant.

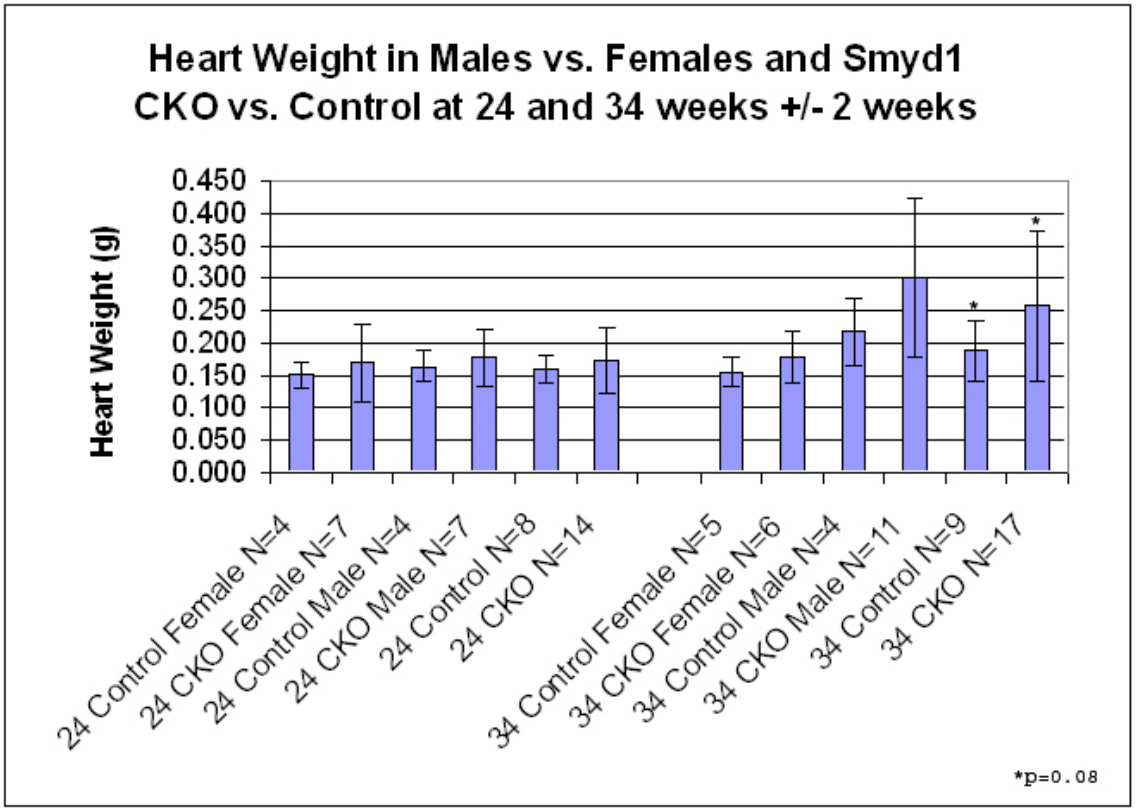


**Fig 6.7 The heart weight is significantly changed in pooled 34 week old mice.** Mice were sacrificed in eight different groups based on age give or take 1 week (23 to 25 weeks and 33 to 35 weeks), gender and genotype. Prior to sacrifice mice were weighed. In most animals the weight of the atria are negligible. To allow for better blood removal, only the ventricles were isolated and weighed. Average weight is reported and error bars represent 1 standard deviation. At 24 weeks, there is not a significant change, but the error bars are increased. It does not seem that this is due to an increased N as pooling wild type males and females doubles the N, but does not affect the error bars. There is a change in average weight of hearts in 34 week old males and females. When genders are pooled, this becomes significance. Again the error bars are greatly increased in the CKO, particularly in the males.

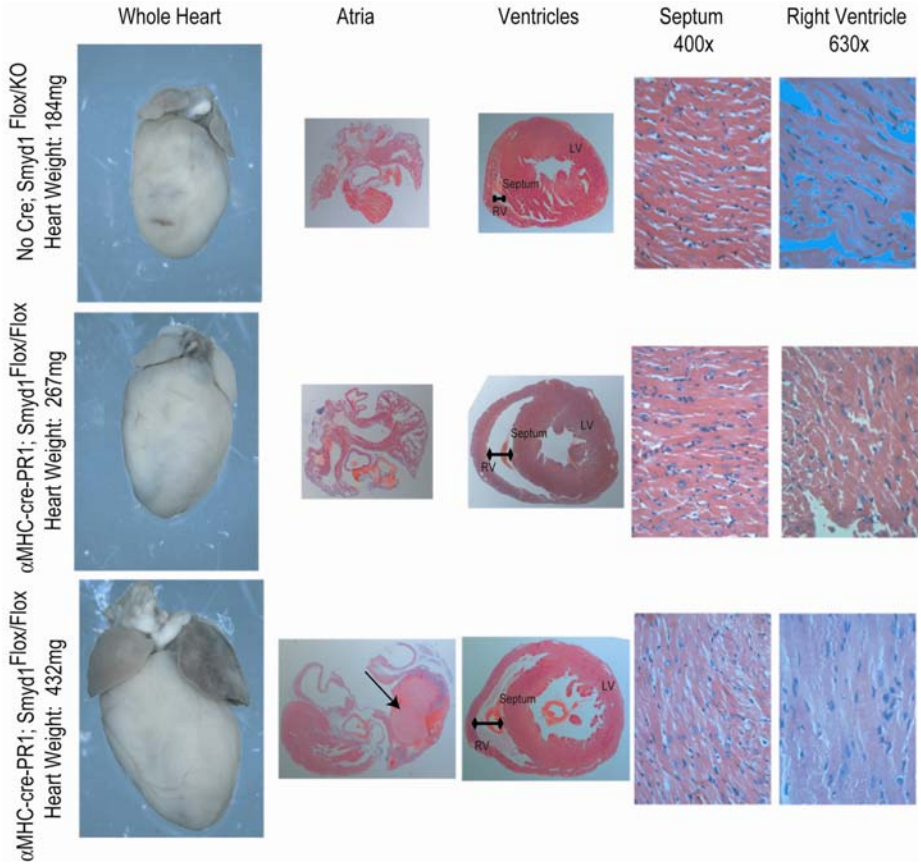


**Fig. 6.8 Hearts of conditional knockout mice are larger, with increased ventricular chamber diameters, and decreased wall thickness, but no defects in the gross muscle structure.** The top panel represents a control *Smyd1* heterozygote; the middle panel represents a non-phenotypic *Smyd1* conditional knockout, and the bottom panel represents a phenotypic *Smyd1* conditional knockout. The heart weights of the non-phenotypic conditional knockout are increased compared to the control and the heart weight of the phenotypic animal was increased compared to both of the others. Blood from the chambers was also more difficult to expel in the conditional knockouts evidenced by the blood pooling in the center of the ventricular chambers. There is an atrial thrombosis (arrow) in the phenotypic heart. The area within the right ventricular chamber is greatly increased in the both CKOs, the shape is of the left ventricle chamber is greatly circularized and the area is increased in the bottom panel. The septum is thinned in both CKOs. The left and right ventricular walls are thinned in the heart in the bottom panel. However, there does not appear to be an effect on the gross muscular structure evidenced in the 400x photographs of the septum and the 630x photographs of the right ventricle. RV: right ventricle, LV: left ventricle.

*Molecular Changes*

In order to determine which genes are affected by the knockdown and which genes are likely to be related to pathology, whole hearts from age matched males and females were subjected to RNA extraction, reverse transcription, and RT-PCR. This has revealed that there is no clear trend in fluctuations of family members, *Smyd2* and *Smyd3*. *Irx4* and *Hand2* were previously found to be deregulated in *Smyd1* deficient embryonic cardiac tissue. However, by RT-PCR analysis, only *Hand2* transcripts appear to be affected in *Smyd1* deficient adult heart tissues. *Hand2* is decreased in deficient hearts compared to controls (Fig. 6.9). Because *Mef2c* was previously linked to the *Smyd1* pathway, albeit upstream, we confirmed that its expression was unaltered by RT-PCR.

The same inflammatory cytokines discussed in chapter 6 are also known to have a role in adult cardiac disease, for example, *TNFα* has an immediate, early protective role against necrosis and apoptosis following induced cardiac injury [85], but over expression of *TNFα* leads to congestive heart failure [86] and knockouts of *TNFα* have an increased survival rate from induced myocardial infarction [87]. We found that while *TNFα* transcript levels were highly variant between individuals of this line, this did not appear to be dependent on *Smyd1* levels (Fig. 6.9). *IL6* is another cytokine linked to cardiac disease and is directly correlated with mortality after myocardial infarction [88]. *IL6* transcript levels appear to be tightly correlated to *Smyd1* levels when gender matching samples. Surprisingly, this affect is opposite in different genders: *Smyd1* deficient males have increased levels of *IL6* compared to controls while females have decreased levels of *IL6* compared to controls (Fig. 6.9).

Because there has recently been a connection between *Smyd1* and regulation of the voltage gated potassium channel alpha sub-unit, Kv4.2, we analyzed transcript levels of *KCND2*, which encodes for Kv4.2. Transcript levels of *Kcnd2* displayed a decrease in

*Smyd1* deficient adult heart tissue (Fig. 6.10). Effects on potassium gradients along with the appearance of seemingly normal cardiac musculature suggest the possibility that animals are dying from arrhythmogenesis, an inability to respond to electrical stimuli, and not the hypothesized ventricular heart failure. In order to address this, male littermates were sacrificed and cDNA pools were created from *"Trizol"* homogenates for the left ventricle and the right ventricle plus septum. Prior to sacrifice, it was noted which animals were phenotypic *Smyd1* CKOs, which were non-phenotypic CKOs, and which were $Smyd1^{flox/flox}$ controls. The phenotypic CKO had extremely low levels of *Smyd1* in both the left and right ventricle, whereas the non-phenotypic littermates have intermediate levels of *Smyd1* in at least one ventricular chamber (Fig. 6.10). Pools were analyzed for changes in atrial and brain natriuretic peptides, markers of hypertrophy and chronic heart failure. Transcript levels of *Atrial* and *Brain Natriuretic Peptides* (*ANP* and *BNP*) were elevated in all CKO left ventricles compared to the control whereas *ANP* and *BNP* transcripts were increased in right ventricles corresponding to the severity of *Smyd1* knockdown (Fig. 6.10). *Kcnd2* and *skNAC* transcripts were also analyzed in individual ventricles. Results from the *Kcnd2* experiment coincide with previously mentioned data and both transcripts are decreased in the absence of *Smyd1* (Fig. 6.10).

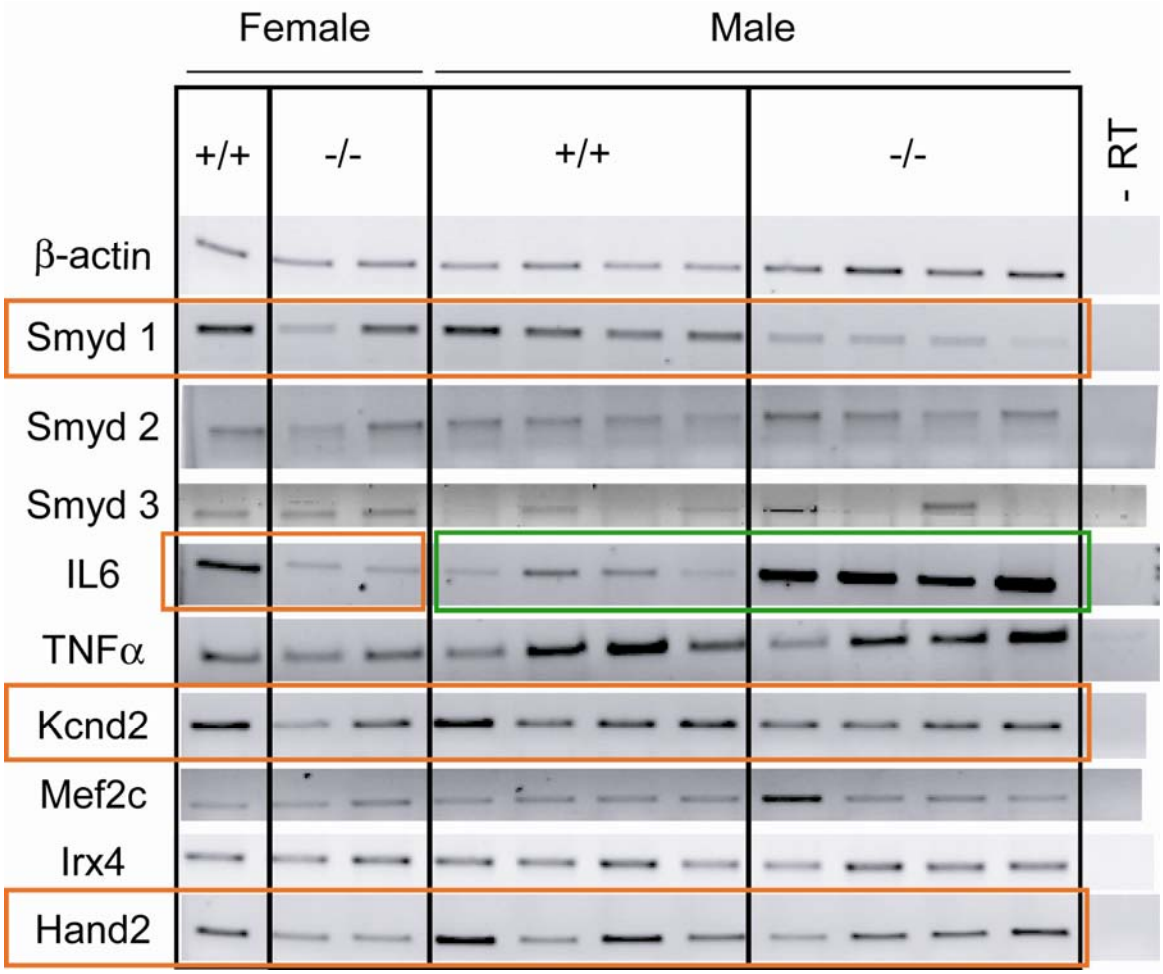


**Fig 6.9** ***Hand2, Kcnd2,*** **and** ***IL6*** **transcripts are deregulated in** ***Smyd1*** **deficient hearts.** RNA was isolated from hearts of 24 week old mice, reverse transcribed and analyzed via RT-PCR. Levels of *Smyd1* were decreased as expected in male mice. However the effect was much less striking in female animals. Levels of *Smyd2* and *Smyd3* were not altered in correlation with *Smyd1* expression level. As observed in embryonic skeletal muscle CKOs, levels of *IL6* were decreased in female CKOs and increased in male CKOs. *TNFα* levels varied, but did not correlate to *Smyd1* levels. *Kcnd2* and *Hand2* were moderately decreased in *Smyd1* deficient hearts.

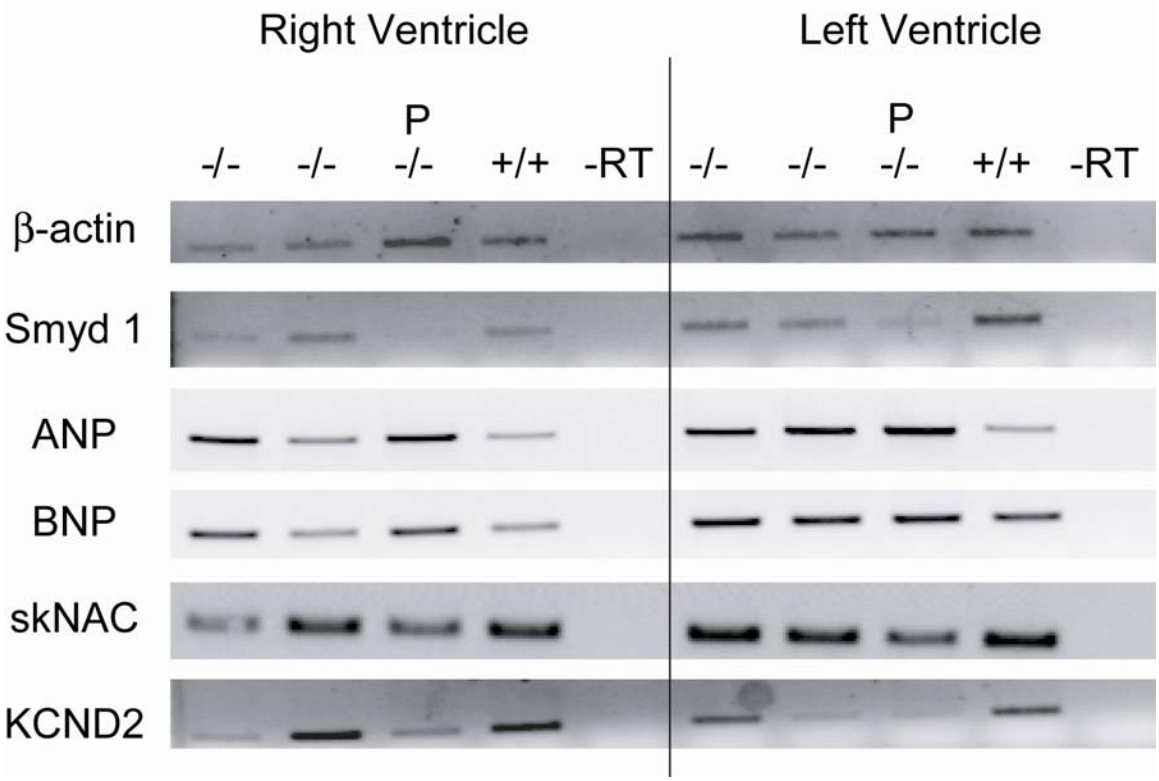


**Fig 6.10 Phenotypic *Smyd1* CKOs have significantly decreased levels of *Smyd1*, increased levels of natriuretic peptides, and decreased levels of *skNAC* and *KCND2* in both ventricles.** 33 week old littermates were sacrificed. The right ventricle and septum were homogenized together and the left ventricle was homogenized separately. RNA was isolated from both homogenates, reverse transcribed and analyzed via RT-PCR. -/- indicates animals that were *Cre* (+); $Smyd1^{Flox/Flox}$. +/+ indicates *Cre* (-); $Smyd1^{Flox/Flox}$. P indicates that the animal was phenotypic. Levels of *ANP* and *BNP* in both ventricles corresponded inversely to *Smyd1* levels. Levels of *skNAC* and *Kcnd2* corresponded directly to *Smyd1* levels. The phenotypic animal had the greatest reduction in *Smyd1* levels in both ventricles. The non-phenotypic have varying levels of *Smyd1* in the right and left ventricles compared to the phenotypic and control animal.

*Discussion*

Although this system was not as tightly controlled by drug administration as initially intended, the results do provide a cardiac specific *Smyd1* knockout that causes early lethality. Due to the delayed onset of phenotype from the point of loss of *Smyd1*, these animals are likely suffering from chronic heart problems, such as chronic heart failure or arrhythmogenesis. It is very rare for transgenic, knock-out, or knock-in mice to develop detectable chronic heart disease that ultimately results in early lethality without added stresses, such as electrical stimuli or invasive aortic banding procedures. With further characterization, this line of mice may provide a model for studying adult onset cardiovascular disease. I have shown that the heart weight in *Smyd1* deficient mice is significantly increased. I have also shown that *Atrial* and *Brain Natriuretic Peptides*, markers of hypertrophy and chronic heart failure, are increased in an inverse relationship with *Smyd1*.

If these mice indeed suffer from chronic heart failure, it mimics human heart disease in that it is highly variable. Like human populations, these mice have a varying mixed background which is likely to contribute to the variation in phenotype or timing of phenotype. Genomic or proteomic analysis may lead to identifying other genetic or expression variables that play a role in adult onset cardiovascular disease.

I have also shown that an alpha subunit of the potassium channel, which has previously been shown to be inhibited by cooperation of *Smyd1* and *Irx5*, is decreased in the absence of *Smyd1*. While this is counter-intuitive, the data shows that in the presence of only *Smyd1* or *Irx5*, *KCND2* is activated. *Irx5* is typically expressed in a gradient; therefore, much of the heart is void of *Irx5*. In this tissue *Smyd1* may typically activate *Kcnd2*. I predict that we are detecting the lack of this activation. *Irx5* and *Kcnd2* expression levels should be further analyzed by immunohistochemistry and in situ

hybridization. Sections of the cardiac muscle look normal, which suggests that mice may not be suffering from chronic heart failure. An alternative possibility is that mice die due to arrhythmogenesis, an inability to interpret the body's electrical signals that cause the heart to beat.

In order to determine which disease or diseases these mice suffer from, mice should be physiologically tested. Electrocardiograms can be performed before and after stimulus to determine if these mice have abnormal heart beats and if they exhibit abnormal tachycardia as seen in the *Irx5* knockout mice. Echocardiograms and invasive hemodynamics should also be used to determine whether the heart is beating and pumping blood through at normal pressures and volumes. Mice should be tested for susceptibility to lowered air oxygen content by using hypoxia chambers or susceptibility to exercise induced hypoxia by using treadmills.

## 6.2 *Smyd1* adult skeletal muscle inducible knockout mediated by *MLC-rtTA* and *Tet-On-Cre* is not effective in all skeletal muscle tissues.

Several inducible versions of transgenic recombinases have been created that are under control of drug sensitive promoters. For example, the bacterial protein *reverse tetracycline transactivator* (*rtTA*) can, when activated by tetracycline or doxycycline, bind to DNA in a sequence specific manner and activate expression of adjacent genes [89]. Therefore, if you create a transgene in which the bacterial *rtTA* coding sequence is fused to a tissue specific promoter and *Cre* recombinase is fused to the bacterial *rtTA* sensitive operon (Tet-O), you can induce tissue specific expression of *Cre* by systematically introducing doxycycline. Thus when both transgenes are combined with a

targeted homozygous floxed allele a tetracycline inducible conditional knockout is created.

*αMLC-rtTA* transgenic mice and *Tet-On-Cre* mice were obtained from the laboratory of Mendel Rimer. These were crossed onto the $Smyd1^{Flox/KO}$ line to obtain $MLC\text{-}rtTA^{+/?}$;$Tet\text{-}On\text{-}Cre^{+/?}$; $Smyd1^{Flox/Flox}$ and $MLC\text{-}rtTA^{+/?}$;$Tet\text{-}On\text{-}Cre^{+/?}$; $Smyd1^{Flox/KO}$ mice. Because mouse hind limbs are so much more muscular than forelimbs, RT-PCR was initially performed using hind limb tissue in mice that had received doxycycline vs. mice that had only received vehicle. However, *Smyd1* transcript levels were not decreased in the hind limbs from doxycycline treated mice. The Rimer lab observed varying levels of efficiency in different muscle groups, the most highly affected tissues were the forelimb and the tibialis anterior of the hind limb [90]. When forelimbs were analyzed for recombination, the knockdown was obvious (Fig. 6.11).

The transactivator of this *Cre* is not effective in all skeletal muscle groups. Therefore, the lack of a phenotype does not mean that *Smyd1* is not essential in adult skeletal muscle. Also, we were not able to analyze effects in slow vs. fast muscle because the soleus, the only primarily slow twitch skeletal muscle in the mouse, is not affected. Attempts at adapting the protocol to create primary myoblasts from forelimbs of individual neonates were not successful.

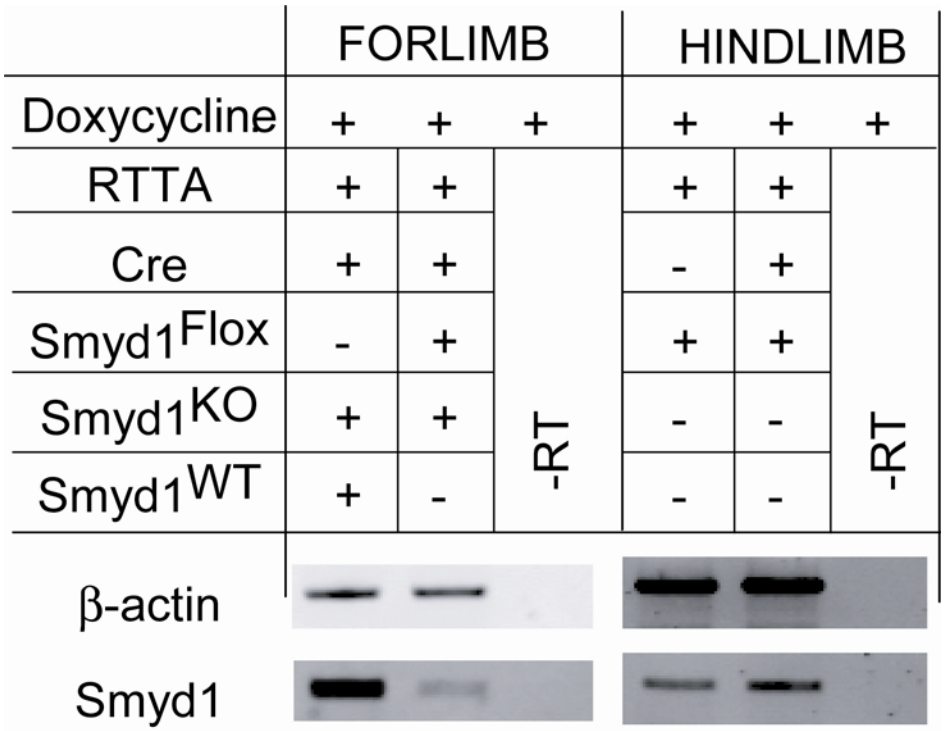


**Fig 6.11 Animals treated with Doxycycline have a significant *Smyd1* knockdown in the forelimb, but not the hind limb.** Mice were given food with doxycycline additives and no alternatives for five days. They were then returned to the regular food for a month before sacrifice. A significant decrease in *Smyd1* transcript was observed in cDNA made from the mouse forelimb of a *Smyd1* deficient animal compared to a homozygote. However, there appears to be an increase in *Smyd1* levels in cDNA made from the hind limb compared to a control that was *Cre* negative.

*Discussion*

While these mice had much potential for analyzing *Smyd1* in adults and for creating primary cultures that could serve as inducible knockdowns, the knockout was not ubiquitous within skeletal muscle. These complications retarded forward progress and the *MyoG-cre* induced adult *Smyd1* knockouts may be better suited for future experiments.

# Chapter 7. Identification of Previously Uncharacterized *Smyd1* Alternative Splice Isoforms

The knockout strategy used for the *Smyd1* conventional and conditional knockouts removes exons 2 and 3 as well as the intronic regions between and immediately adjacent. Previous studies have reported that this strategy was effective for knocking out all splice forms created by the *Smyd1* gene [10]. However, primers used to determine this were in exons one and four. I independently performed RT-PCR with primers in exons one and six and obtained a product that was significantly smaller than the expected product. By purifying and sequencing the observed product, I demonstrated that it represented a splice junction between *Smyd1* exons one and six. By PCR amplification on cDNA pools using primers that overlapped the start codon of *Smyd1a/b* and the 3' UTR of *Smyd1a/b*, a form of *Smyd1* that encodes exons 1, 6, 7, 8, 9, and 10 was purified and cloned into an expression vector. During the process of creating this clone, another alternative splice form was also discovered and cloned that encodes exons 1, 5, 6, 7, 8, 9, and 10.

In E15.5 embryonic limbs, I found a band that represents one or both of these transcripts in all samples that contained a *Smyd1* deficient allele, but not when there was simply a $Smyd1^{Flox}$ allele in the absence of *Cre* (Fig. 7.1). In an attempt to determine if this is an endogenous transcript, other cDNA pools were analyzed including normal and deficient, embryonic hearts, adult hearts, and adult limbs. I found that the alternative splice form(s) were introduced in skeletal muscle and embryonic heart after *Smyd1* knockdown, however, expression was detected in the adult heart without a knockdown (Fig. 7.1).

When the putative protein gene product from the exon one to exon six splice form was computationally analyzed, the domain blast program was able to detect a rudimentary SET domain, similar to an outlier yeast SET domain. If a protein is made by either of the newly discovered alternative splice products, they should be detectable by western blot with Smyd1 antibodies which were raised to the C-terminal portion of the protein. However, no evidence of this protein product has ever been detected (Fig. 7.2).

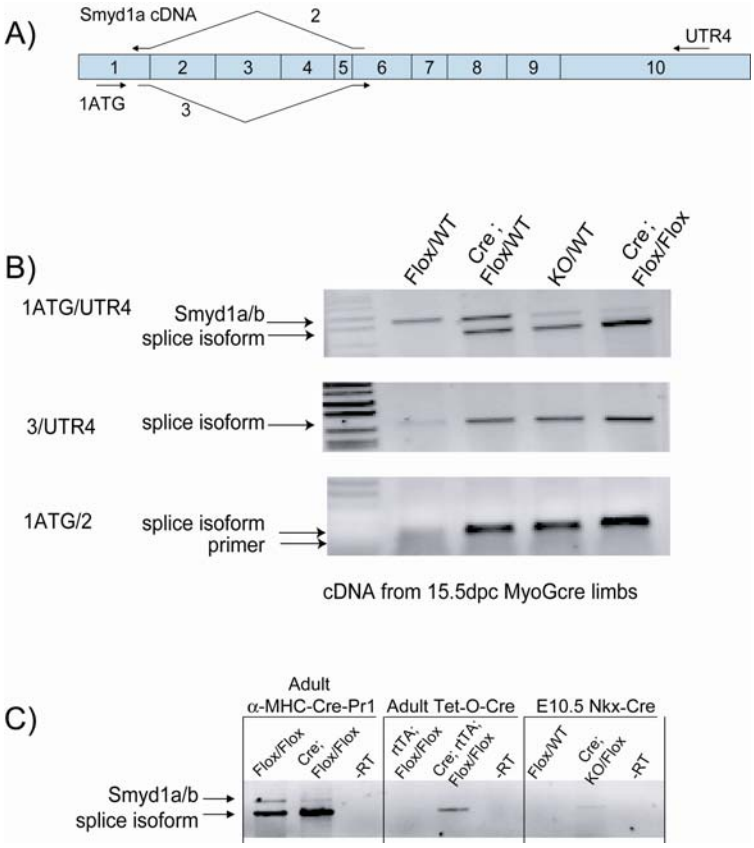


**Fig 7.1 With the exception of the adult heart, where *Smyd1* alternative splice forms exist endogenously, the previously undocumented transcripts exist only in the presence of the *Smyd1* deficient alleles. A)** *Smyd1a* cDNA representing exons and relevant primers. **B)** cDNA from several genotypes of *MyoG-cre* embryonic skeletal muscle**.** The top panel detects the full length version of all *Smyd1* isoforms. There is a 600 base pair difference between *Smyd1a* and the exon 1-6 splice variant, 1.8 and 1.2kb, respectively. The splice variant should be preferentially amplified due to a short (1') elongation period. There is no alternative splice form in the *Cre* (-) $Smyd1^{Flox/WT}$ sample, although there is a strong product in the *Cre*(+); $Smyd1^{Flox/WT}$, *Cre*(-); $Smyd1^{KO/WT}$, and *Cre*(+); $Smyd1^{Flox/Flox}$ samples. The middle and bottom panels should detect only the splice variant as the internal primer primes the splice junction. There could be minimal forward priming on the WT allele as the 3' end of both internal primers could anneal. All samples that contain a defective *Smyd1* allele are highly enriched for the splice junction product. **C)** 1ATG/UTR4 reaction in deficient and wild type adult heart, adult skeletal muscle and embryonic heart. The splice variant should be enriched due to having a limited elongation period and a shorter product. All tissues that contain one or more deficient *Smyd1* alleles are positive for the splice variant. The adult heart is the only place that the splice variant has been detected endogenously.

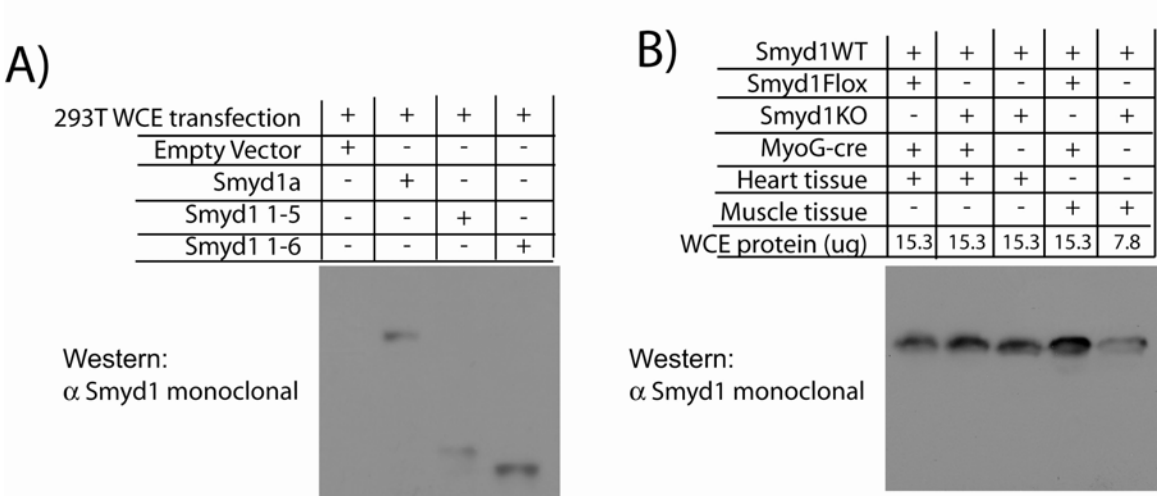


**Fig 7.2 Although, the *Smyd1* monoclonal antibody detects *Smyd1* splice variant encoded proteins, there is no evidence to suggest these proteins are expressed *in vivo*.** A) Western blot of transfected 293T cell whole cell extract. The *Smyd1* monoclonal antibody detects *Smyd1a* as well as both splice variants. B) Whole cell extracts from adult skeletal and adult cardiac tissue detect large amounts of *Smyd1a* and/or *Smyd1b*. However, there is no evidence that the smaller splice variants exist. RT-PCR data suggests that the transcripts would be present in all five of these tissue. Its possible that the protein exists in levels not detectable by western blot or that these transcripts do not encode a protein.

*Discussion*

It is possible that some of the observed phenotypes discussed above are due to the over expresssion of these splice isoforms instead of the knockdown of the previously characterized gene products. The fact that this isoform exists endogenously in the adult heart suggests that it functions in some manner. Although we have no evidence that this transcript is translated into protein, it is possible that it escaped detection by expressing at low levels, having a short half life, or requiring some sort of cues or stimulation prior to translation. If this protein is translated it is likely to have some methylation function as it does have a rudimentary SET domain. However, it does not contain the MYND domain, which is thought to be important for binding the methylation substrate. This gene product may act as a dominant negative as some protein-protein interactions are thought to occur in C-terminal tail, although it is unclear what the ultimate results are of these interactions. Over expressing a natural dominant negative would only enhance our visualization of the *Smyd1* deficient phenotype. However, it is possible that this protein isoform has a separate function or that the transcript has a function, perhaps encoding a short RNA. While the presence of these transcripts has been greatly ignored, they should continue to be considered as putatively playing a role in the observed phenotypes.

## Chapter 8. Discussion

Many mouse knockouts have effects only in certain backgrounds. Often a single cross to a different line will cause the phenotype to simply disappear. The fact that the knockouts of *Smyd1* have such pronounced effects on these highly variant backgrounds speaks that the role of *Smyd1* is very essential and non-redundant. However, this mixed background also adds many complications. For example, the modulation of *TNFα, Smyd2,* and *Smyd3* transcript levels was consistently seen. However, these changes seem to be completely independent of *Smyd1* levels. It is possible that changes in levels of these transcripts, while not caused by the lack of *Smyd1*, compound the observed phenotype; mice with lower levels may be able to survive longer or be less phenotypic than mice with higher levels, or vice versa. It is impossible to determine how many other variations exist from litter to litter or even within a single litter. Likewise, I was very surprised to consistently detect varying levels of *IL6* between males and females in both the adult heart and embryonic skeletal muscle. The ability to analyze age and gender matched litter mates is a limiting factor. Therefore, while the impact of this system is clear, increased numbers must be analyzed before drawing solid conclusions.

It is obvious that *Smyd1* is essential throughout the development of both cardiac and skeletal muscle. Even on mixed backgrounds mice are incapable of surviving through critical stages of development when they are deficient in *Smyd1* expression. Surprisingly, *Smyd1* seems to be very essential to cardiac performance in adult mice. However, it is unclear how *Smyd1* functions at the molecular level. It is likely that *Smyd1* functions through lysine methylation of histone or non-histone proteins, like the transcription factor *skNAC*. It is possible that *Smyd1* primarily acts in the canonical manner for SET domain containing proteins and is a histone methyltransferase.

However, although there is a change in global histone acetylation, it has not been possible to detect a global change in histone methylation in *Smyd1* deficient cells compared to controls (personal communication, L. Zhu and P. Tucker) and we have not been able to detect many alterations in the transcription of many genes in any knockout or knockdown system. Therefore it is becoming more and more suggestive that the primary role of *Smyd1* may to regulate protein activities post translationally via protein methyltransferase activity.

It is surprising that two different potassium channels have come into question in two different systems. It seems equally unlikely that Smyd1 has a direct role on two unrelated potassium channel subunits, *Kcne1* in embryonic heart and *Kcnd2* in adult heart, or that *Smyd1* has a global effect on the regulation of potassium channels, but both possibilities require further analysis. Regardless, it seems that *Smyd1* affects many independent pathways, the roles of the two characterized and two uncharacterized isoforms may need to be further separated before elucidation of the molecular mechanics can occur.

In order to accomplish this, over expression studies of the two novel *Smyd1* isoforms should be performed to determine if they affect the ability of C2C12 myoblasts to differentiate or proliferate. Over expression of all four isoforms can be studied in 10T1/2 cells, which can be pushed to differentiate into myoblasts. We should develop northern or in situ hybridization probes that are specific to each of the four isoforms. It will be easy to differentiate *Smyd1a* and the exon one to five splice isoform because they contain unique nucleotide and protein sequences. However, *Smyd1b* and the *1-6 splice isoform* only differ from *Smyd1a* and the *1-5 splice isoform*, respectively, by lacking the thirty-nine base pair exon five. These two isoforms do not have unique sequences. There are not currently antibodies that distinguish the individual isoforms. Specific primers and

probes are contingent on the splice junction sequences that will not hybridize to isoforms that contain the short insert.

Further studies should also incorporate the further characterization of the embryonic and adult *MyoG-cre* induced skeletal muscle specific conditional knockouts and the *αMHC-crePR1* induced adult heart specific conditional knockout. Further hypotheses on the molecular basis of the phenotypes would be well guided by knowing whether both differentiation and proliferation were blocked in the embryonic forms, whether the satellite cells or cardiac stem cells were affected by the loss of *Smyd1*, and whether the adults are capable of responding to the electrical stimuli that causes contraction of both muscle types.

## Appendix 1: Alphabetical Listing of all Transcription Factor Binding Site Hits on 1kb of mm*Smyd1* putative promoter

Definitions:

Opt: Threshold Matrix Similarity, optimized for reducing background of each gene

Position: column 1 represents the range of nucleotides corresponding to the matrix binding site, nucleotides numbered in reverse fashion where 1 is 1006bp upstream of the transcription initiation site and 1006 is equal to the transcription initiation site, +1.

Column 2 is the anchor point.

Str: DNA strand

Core Sim.: ratio of matching nucleotides in the core binding sequence

Matrix Sim: ratio of matching nucleotides in the entire binding sequence

Sequence: recognized DNA sequence, core sequence in capital letters

| Family/matrix | Further Information | Opt. | Position | | Str. | Core sim. | Matrix sim. | Sequence |
|---|---|---|---|---|---|---|---|---|
| V$AP1F/VMAF.01 | v-Maf | 0.82 | 386 - 406 | 396 | (+) | 1 | 0.826 | acagcTGACatgggag agaaa |
| V$AP4R/AP4.01 | Activator protein 4 | 0.97 | 308 - 324 | 316 | (-) | 1 | 0.97 | caagaCAGCtggattca |
| V$AP4R/AP4.01 | Activator protein 4 | 0.97 | 381 - 397 | 389 | (-) | 1 | 0.979 | catgtCAGCtgtgtctc |
| V$BCL6/BCL6.02 | POZ/zinc finger protein, transcriptional repressor, translocations observed in diffuse large cell lymphoma | 0.77 | 253 - 269 | 261 | (+) | 1 | 0.83 | agtttggTAGAaaggtc |
| V$CEBP/CEBP.02 | C/EBP binding site | 0.85 | 505 - 523 | 514 | (+) | 1 | 0.88 | ggggtgggGCAAgggg ccc |
| V$CLOX/CDPCR3 HD.01 | cut-like homeodomain protein | 0.94 | 760 - 776 | 768 | (-) | 1 | 0.958 | cctgacatgGATCctga |
| V$CMYB/CMYB.0 1 | c-Myb, important in hematopoesis, cellular equivalent to avian myoblastosis virus oncogene v-myb | 0.99 | 125 - 133 | 129 | (+) | 1 | 0.992 | caGTTGgcc |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| V$COMP/COMP1.01 | COMP1, cooperates with myogenic proteins in multicomponent complex | 0.76 | 283 - 303 | 293 | (-) | 1 | 0.811 | aggcaggATTGaaaagaagag |
| V$COMP/COMP1.01 | COMP1, cooperates with myogenic proteins in multicomponent complex | 0.76 | 616 - 636 | 626 | (+) | 1 | 0.827 | ccttttgATTGtcagaggcac |
| V$EBOX/MYCMAX.01 | c-Myc/Max heterodimer | 0.83 | 658 - 674 | 666 | (+) | 0.75 | 0.837 | ctgacCAAGtgctcttg |
| V$EBOX/MYCMAX.02 | c-Myc/Max heterodimer | 0.92 | 711 - 727 | 719 | (+) | 0.895 | 0.928 | cacagCACAtgctggga |
| V$EBOX/MYCMAX.03 | MYC-MAX binding sites | 0.92 | 480 - 496 | 488 | (+) | 1 | 0.926 | cttgcCACGgggagaca |
| V$ECAT/NFY.01 | nuclear factor Y (Y-box binding factor) | 0.90 | 850 - 864 | 857 | (-) | 1 | 0.93 | aaaggCCAAtgaaga |
| V$EGRF/WT1.01 | Wilms Tumor Suppressor | 0.88 | 498 - 512 | 505 | (+) | 1 | 0.945 | tagggTGGGggtggg |
| V$ETSF/FLI.01 | ETS family member FLI | 0.81 | 34 - 50 | 42 | (+) | 0.75 | 0.817 | ataaCCAGaaatttaaa |
| V$EVI1/EVI1.02 | Ecotropic viral integration site 1 encoded factor | 0.83 | 808 - 824 | 816 | (-) | 1 | 0.851 | cacccagaaAAGAcacc |
| V$EVI1/EVI1.04 | Ecotropic viral integration site 1 encoded factor | 0.77 | 23 - 39 | 31 | (-) | 0.8 | 0.804 | gGTTAtgtttagattct |
| V$EVI1/EVI1.04 | Ecotropic viral integration site 1 encoded factor | 0.77 | 742 - 758 | 750 | (-) | 0.8 | 0.817 | aGACAggatttgaaaac |
| V$EVI1/EVI1.06 | Ecotropic viral integration site 1 encoded factor | 0.83 | 316 - 332 | 324 | (-) | 0.75 | 0.835 | tggtgatacaAGACagc |
| V$EVI1/EVI1.06 | Ecotropic viral integration site 1 encoded factor | 0.83 | 782 - 798 | 790 | (-) | 0.75 | 0.835 | aatgaccacaAGCTagt |
| V$FKHD/HFH1.01 | HNF-3/Fkh Homolog 1 | 0.85 | 397 - 413 | 405 | (+) | 1 | 0.857 | gggagagAAACaatgtc |
| V$FKHD/HNF3B.01 | Hepatocyte Nuclear Factor 3beta | 0.95 | 46 - 62 | 54 | (+) | 1 | 0.955 | ttaaaaaaAATActaca |
| V$FKHD/XFD2.01 | Xenopus fork head domain factor 2 | 0.89 | 23 - 39 | 31 | (+) | 1 | 0.912 | agaatcTAAAcataacc |
| V$GATA/LMO2COM.02 | complex of Lmo2 bound to Tal-1, E2A proteins, and GATA-1, half-site 2 | 0.96 | 114 - 126 | 120 | (+) | 1 | 0.961 | gtaaGATAggcca |
| V$GFI1/GFI1.01 | Growth factor independence 1 zinc finger protein acts as transcriptional repressor | 0.97 | 666 - 680 | 673 | (-) | 1 | 0.972 | tgaAATCaagagcac |
| V$GFI1/GfI1B.01 | Growth factor independence 1 zinc finger protein Gfi-1B | 0.82 | 690 - 704 | 697 | (+) | 0.75 | 0.82 | gaaTATGaccgcagt |
| V$HAML/AML1.01 | runt-factor AML-1 | 1.00 | 772 - 786 | 779 | (-) | 1 | 1 | ctagtgTGGTcctga |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| V$HAML/AML1.01 | runt-factor AML-1 | 1.00 | 785 - 799 | 792 | (+) | 1 | 1 | agcttgTGGTcattt |
| V$HEAT/HSF1.01 | heat shock factor 1 | 0.93 | 261 - 271 | 266 | (+) | 1 | 0.933 | AGAAaggtctg |
| V$HMTB/MTBF.01 | muscle-specific Mt binding site | 0.90 | 468 – 476 | 472 | (+) | 1 | 0.932 | tgctATTTa |
| V$HMTB/MTBF.01 | muscle-specific Mt binding site | 0.90 | 647 - 655 | 651 | (-) | 1 | 0.901 | ggctATTTt |
| V$HMTB/MTBF.01 | muscle-specific Mt binding site | 0.90 | 978 – 986 | 982 | (-) | 1 | 0.932 | tgctATTTa |
| V$HNF4/HNF4.01 | Hepatic nuclear factor 4 | 0.82 | 855 - 871 | 863 | (-) | 0.75 | 0.849 | ggagggAAAAggccaa t |
| V$HNF4/HNF4.02 | Hepatic nuclear factor 4 | 0.77 | 256 - 272 | 264 | (+) | 1 | 0.815 | ttggtagAAAGgtctgg |
| V$HOXF/EN1.01 | Homeobox protein engrailed (en-1) | 0.77 | 729 - 745 | 737 | (+) | 1 | 0.807 | tatcTTTAagtgggttt |
| V$HOXF/HOX1-3.01 | Hox-1.3, vertebrate homeobox protein | 0.83 | 567 - 583 | 575 | (+) | 1 | 0.83 | gggcagccATTAttact |
| V$HOXF/HOX1-3.01 | Hox-1.3, vertebrate homeobox protein | 0.83 | 570 - 586 | 578 | (+) | 1 | 0.864 | cagccattATTActttg |
| V$HOXF/PTX1.01 | Pituitary Homeobox 1 (Ptx1) | 0.79 | 62 - 78 | 70 | (-) | 1 | 0.845 | tgtggcTTAGcttctgt |
| V$HOXF/PTX1.01 | Pituitary Homeobox 1 (Ptx1) | 0.79 | 637 - 653 | 645 | (-) | 1 | 0.793 | ctatttTTAGcctccag |
| V$IKRS/IK1.01 | Ikaros 1, potential regulator of lymphocyte differentiation | 0.92 | 96 - 108 | 102 | (-) | 1 | 0.926 | ttgaGGGAatggc |
| V$IKRS/IK2.01 | Ikaros 2, potential regulator of lymphocyte differentiation | 0.98 | 917 - 929 | 923 | (+) | 1 | 0.982 | agttGGGAagggg |
| V$LEFF/LEF1.01 | TCF/LEF-1, involved in the Wnt signal transduction pathway | 0.86 | 831 - 847 | 839 | (-) | 1 | 0.878 | ggccttaCAAAggtctc |
| V$LTUP/TAACC.01 | Lentiviral TATA upstream element | 0.71 | 23 - 45 | 34 | (+) | 1 | 0.737 | agaatctaaacatAACC agaaat |
| V$LTUP/TAACC.01 | Lentiviral TATA upstream element | 0.71 | 786 - 808 | 797 | (-) | 0.773 | 0.79 | caagacagaaaatGAC Cacaagc |
| V$MAZF/MAZ.01 | Myc associated zinc finger protein (MAZ) | 0.90 | 182 - 194 | 188 | (-) | 1 | 0.903 | tcagGAGGggcga |
| V$MAZF/MAZ.01 | Myc associated zinc finger protein (MAZ) | 0.90 | 887 - 899 | 893 | (+) | 1 | 0.949 | ggggGAGGtgagg |
| V$MAZF/MAZR.01 | MYC-associated zinc finger protein related transcription factor | 0.88 | 504 - 516 | 510 | (+) | 1 | 0.938 | gggggtGGGGcaa |
| V$MEF2/AMEF2.01 | myocyte enhancer factor | 0.80 | 21 - 43 | 32 | (+) | 1 | 0.847 | ccagaatcTAAAcataa ccagaa |
| V$MEF2/MEF2.02 | myogenic MADS factor MEF-2 | 0.89 | 637 - 659 | 648 | (+) | 1 | 0.986 | ctggaggctaaaAATAg cctact |
| V$MEF2/MEF2.05 | MEF2 | 0.96 | 37 - 59 | 48 | (+) | 1 | 0.969 | accagaaattTAAAaaa aatact |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| V$MEF2/MMEF2.01 | myocyte enhancer factor | 0.90 | 39 - 61 | 50 | (+) | 1 | 0.904 | cagaaattTAAAaaaaatactac |
| V$MEIS/MEIS1.01 | Binding site for monomeric Meis1 homeodomain protein | 0.95 | 547 - 555 | 551 | (+) | 1 | 0.991 | ctGACAgct |
| V$MOKF/MOK2.02 | Ribonucleoprotein associated zinc finger protein MOK-2 (human) | 0.98 | 446 - 466 | 456 | (-) | 1 | 0.993 | aagctaaaatgagCCTTttgc |
| V$MYOD/LMO2COM.01 | complex of Lmo2 bound to Tal-1, E2A proteins, and GATA-1, half-site 1 | 0.98 | 629 - 643 | 636 | (-) | 1 | 0.98 | cctcCAGGtgcctct |
| V$MYOD/MYF5.01 | Myf5 myogenic bHLH protein | 0.90 | 602 - 616 | 609 | (-) | 0.963 | 0.907 | gagaCAACtggagtg |
| V$MYOD/MYOD.02 | myoblast determining factor | 0.98 | 310 - 324 | 317 | (+) | 0.914 | 0.983 | aatcCAGCtgtcttg |
| V$MYOF/MYOGNF1.01 | Myogenin / nuclear factor 1 or related factors | 0.71 | 844 - 872 | 858 | (+) | 1 | 0.714 | ggccactcttcaTTGGcctttccctcct |
| V$NEUR/NEUROD1.01 | DNA binding site for NEUROD1 (BETA-2 / E47 dimer) | 0.83 | 122 - 134 | 128 | (-) | 0.767 | 0.886 | tggcCAACtggcc |
| V$NF1F/NF1.01 | Nuclear factor 1 | 0.94 | 854 - 872 | 863 | (+) | 1 | 0.942 | catTGGCcttttccctcct |
| V$NFKB/CREL.01 | c-Rel | 0.91 | 92 - 106 | 99 | (+) | 1 | 0.947 | aagggccaTTCCctc |
| V$NKXH/HMX3.01 | H6 homeodomain HMX3/Nkx5.1 transcription factor | 0.89 | 732 - 744 | 738 | (+) | 1 | 0.908 | ctttAAGTgggtt |
| V$NKXH/NKX25.01 | homeo domain factor Nkx-2.5/Csx, tinman homolog, high affinity sites | 1.00 | 524 - 536 | 530 | (-) | 1 | 1 | tctcAAGTgaggt |
| V$NOLF/OLF1.01 | olfactory neuron-specific factor | 0.82 | 237 - 259 | 248 | (-) | 1 | 0.869 | ccaaacTCCCtcaaggatgtgtc |
| V$NRSF/NRSF.01 | neuron-restrictive silencer factor | 0.69 | 759 - 779 | 769 | (+) | 0.756 | 0.693 | gtcAGGAtccatgtcaggacc |
| V$OAZF/ROAZ.01 | Rat C2H2 Zn finger protein involved in olfactory neuronal differentiation | 0.73 | 491 - 507 | 499 | (-) | 0.75 | 0.733 | ccCCACcctagtgtctc |
| V$OCT1/OCT1.05 | octamer-binding factor 1 | 0.90 | 347 - 361 | 354 | (-) | 1 | 0.917 | gGCATggaagtcaga |
| V$OCTP/OCT1P.01 | octamer-binding factor 1, POU-specific domain | 0.86 | 687 - 699 | 693 | (-) | 0.98 | 0.905 | ggtcatATTCagc |
| V$PAX1/PAX1.01 | Pax1 paired domain protein, expressed in the developing vertebral column of mouse embryos | 0.61 | 860 - 878 | 869 | (+) | 0.75 | 0.616 | CCTTttccctcctgacaca |
| V$PAX5/PAX9.01 | zebrafish PAX9 binding sites | 0.78 | 838 - 866 | 852 | (-) | 0.8 | 0.837 | gaaaagGCCAatgaagagtggccttacaa |
| V$PAX8/PAX8.01 | PAX 2/5/8 binding site | 0.88 | 1028 - 1040 | 1034 | (-) | 0.85 | 0.927 | ctgTCAGgccttc |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| V$PBXC/PBX1_MEIS1.02 | Binding site for a Pbx1/Meis1 heterodimer | 0.77 | 617 - 633 | 625 | (+) | 1 | 0.883 | ctttTGATtgtcagagg |
| V$PBXC/PBX1_MEIS1.03 | Binding site for a Pbx1/Meis1 heterodimer | 0.76 | 188 - 204 | 196 | (+) | 0.75 | 0.774 | ctcctgagTGAGaggag |
| V$PBXC/PBX1_MEIS1.03 | Binding site for a Pbx1/Meis1 heterodimer | 0.76 | 750 - 766 | 758 | (-) | 0.75 | 0.781 | atcctgacAGACaggat |
| V$PBXF/PBX1.01 | homeo domain factor Pbx-1 | 0.78 | 618 - 630 | 624 | (-) | 1 | 0.789 | ctgaCAATcaaaa |
| V$PCAT/ACAAT.01 | Avian C-type LTR CCAAT box | 0.86 | 876 - 886 | 881 | (-) | 0.75 | 0.866 | tccaCCACtgt |
| V$PCAT/CAAT.01 | cellular and viral CCAAT box | 0.90 | 853 - 863 | 858 | (-) | 1 | 0.95 | aaggCCAAtga |
| V$PERO/PPARA.01 | PPAR/RXR heterodimers | 0.70 | 856 - 876 | 866 | (-) | 1 | 0.701 | tgtcaggagggaAAAGgccaa |
| V$RBPF/RBPJK.01 | Mammalian transcriptional repressor RBP-Jkappa/CBF1 | 0.84 | 719 - 733 | 726 | (+) | 1 | 0.84 | atgcTGGGagtatct |
| V$RORA/NBRE.01 | Monomers of the nur subfamily of nuclear receptors (nur77, nurr1, nor-1) | 0.89 | 259 - 275 | 267 | (+) | 1 | 0.93 | gtagaAAGGtctgggag |
| V$RORA/NBRE.01 | Monomers of the nur subfamily of nuclear receptors (nur77, nurr1, nor-1) | 0.89 | 827 - 843 | 835 | (-) | 1 | 0.93 | ttacaAAGGtctcgttc |
| V$RREB/RREB1.01 | Ras-responsive element binding protein 1 | 0.79 | 499 - 513 | 506 | (-) | 1 | 0.853 | cCCCAcccccaccct |
| V$SORY/SOX5.01 | Sox-5 | 0.87 | 402 - 418 | 410 | (+) | 1 | 0.984 | agaaaCAATgtcccttg |
| V$SORY/SOX9.01 | SOX (SRY-related HMG box) | 0.90 | 1048 - 1064 | 1056 | (+) | 1 | 0.912 | gatgaCAATaggcagca |
| V$SP1F/GC.01 | GC box elements | 0.88 | 503 - 517 | 510 | (+) | 0.872 | 0.934 | tgggGGTGgggcaag |
| V$SRFF/SRF.02 | serum response factor | 0.83 | 432 - 450 | 441 | (-) | 0.861 | 0.831 | tttgCCACatatgtagcag |
| V$STAT/STAT.01 | signal transducers and activators of transcription | 0.87 | 95 - 113 | 104 | (-) | 1 | 0.904 | taaggttgagGGAAtggcc |
| V$STAT/STAT5.01 | STAT5: signal transducer and activator of transcription 5 | 0.89 | 298 - 316 | 307 | (-) | 0.845 | 0.941 | ctggaTTCAgagaaggcag |
| V$TALE/TGIF.01 | TG-interacting factor belonging to TALE class of homeodomain factors | 1.00 | 546 - 552 | 549 | (-) | 1 | 1 | tGTCAga |
| V$TALE/TGIF.01 | TG-interacting factor belonging to TALE class of homeodomain factors | 1.00 | 625 - 631 | 628 | (+) | 1 | 1 | tGTCAga |
| V$TBPF/ATATA.01 | Avian C-type LTR TATA box | 0.81 | 728 - 744 | 736 | (+) | 1 | 0.816 | gtatcttTAAGtgggtt |
| V$TBPF/MTATA.01 | Muscle TATA box | 0.84 | 37 - 53 | 45 | (-) | 1 | 0.853 | tttttTAAAtttctggt |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| V$TBPF/TATA.01 | cellular and viral TATA box elements | 0.90 | 936 - 952 | 944 | (-) | 1 | 0.967 | gactaTAAAtggactgg |
| V$TCFF/TCF11.01 | TCF11/KCR-F1/Nrf1 homodimers | 1.00 | 793 - 799 | 796 | (+) | 1 | 1 | GTCAttt |
| V$TTFF/TTF1.01 | Thyroid transcription factor-1 (TTF1) binding site | 0.92 | 524 - 538 | 531 | (-) | 1 | 0.925 | tttctCAAGtgaggt |
| V$TTFF/TTF1.01 | Thyroid transcription factor-1 (TTF1) binding site | 0.92 | 665 - 679 | 672 | (-) | 1 | 0.924 | gaaatCAAGagcact |
| V$ZBPF/ZBP89.01 | Zinc finger transcription factor ZBP-89 | 0.93 | 501 - 513 | 507 | (-) | 1 | 0.976 | ccccaCCCCcacc |
| V$ZBPF/ZBP89.01 | Zinc finger transcription factor ZBP-89 | 0.93 | 884 - 896 | 890 | (-) | 1 | 0.966 | cacctCCCCctcc |
| | | | from – to | anchor | | | | |

## Appendix 2: RT PCR primer pairs and conditions

| Gene | Forward Primer | Reverse Primer | Annealing Temp. | Cycle# (range) | Product Size(bp) |
|---|---|---|---|---|---|
| 1700129I15Rik | ggccctgaagctaactgaagac | ggtttccttggagggcagag | 65 | 36-41 | 169 |
| Actb | tacgagggctatgctctc | cgcagctcagtaacagtc | 58-62 | 19-30 | 651 |
| Akt1 | gccctggactacttgcac | cgaacagcttctcgtggtcc | 61 | 27-43 | 300 |
| AR | caagctggagaacccattggac | ggcagcaaaggaatctggttgg | 55 | 35 | 897 |
| Calm4 | gcagcagactgacatccagttg | cacacccaaggaaaccatcctc | 65 | 30-35 | 745 |
| CideA | gcctgcaggaacttatcagc | cacggccttgaagcttgtg | 58 | 30 | 347 |
| Cre | ggacatgttcagggatcgccaggcg | gcataaccagtgaaacagcattgctg | 60-62 | 20-35 | 268 |
| Cyr61 | ggcatctccacacgagttacc | ccgcatcttcacagttctggtc | 60-65 | 30-40 | 268 |
| Gyltl1B | cgtcgccagtgctctctatg | cagctgaccactgcacactg | 60 | 30-32 | 414 |
| Hand2 | tacttccacggctggcttattg | tcgttgctgctcactgtgcttt | 62 | 35-36 | 467 |
| Igf1 | cttctacctggcgctctgc | gtcttgggcatgtcagtgtgg | 61-61.4 | 27-43 | 300 |
| IL6 | gccttcttgggactgatgc | cgcactaggtttgccgag | 60 | 27-35 | 604 |
| Irx4 | tgctggccatcatcaccaagat | ttggaccttggcgtttcagcat | 62 | 35-36 | 666 |
| Irx5-2 | caccaagatgaccctcaccc | gtaataaagaggccgggacagc | 60-62 | 34-36 | 763 |
| Isl1 | tttctccggatttggagtggca | atgcaagggactgagagggtct | 62 | 35-40 | 359 |
| Kcnd2 | tttcgtgcgcagtgtcatgagt | aagttcgacacgatcacaggca | 62 | 28-35 | 459 |
| Mafbx | gtctgtgctggtgggcaac | gcaaagctgcagggtgac | 61 | 30-43 | 394 |
| Mdfi | cggaagttgcagacgcatcc | gagtccacagcactccatg | 60-65 | 30-40 | 342 |
| Mef2c | acagcaccaacaagctgttcca | tgcttaaagagcaaccggagga | 62 | 30-35 | 549 |
| MK2 | gtgcctgctgattgtcatgg | gatgacacccaaggaccac | 60 | 30-32 | 361 |
| Murf1 | ctcctgcagagtgaccaag | gtcccaaagtcaatggccc | 61 | 30-43 | 399 |
| Myf5 | ttcggagcacacaaagctga | aagctgctgttctttcggga | 60 | 30-35 | 431 |
| Myf6 | tgcgcgaaaggaggagac | gaaaggcgctgaagactgc | 60 | 32-35 | 379 |
| MyHCIIa (Myh2) | cgaccctgaccgtgaaag | tccggtgatcaggattga | 60 | 38-40 | 335 |
| Myo6 | gtggagcaggcaaacaatgctc | cttcattaacacccaggccctc | 60-65 | 36-40 | 280 |
| MyoD1 | tgatgacccgtgtttcgact | ttccctgttctgtgtcgctt | 60 | 32-37 | 831 |
| MyoG | aagcgcaggctcaagaaagt | tcagttgggcatggtttcgt | 60 | 30-35 | 426 |
| Mb | ctcaccctgagaccctgg | gtcattccggaagagctcc | 60-65 | 30-37 | 330 |
| Myostatin | gctggcccagtggatctaaatg | gggtgtgtctgtcaccttgac | 60 | 30-32 | 362 |
| p21 | gcagatccacagcgatatccag | ctaagggccctaccgtcctac | 60 | 28-30 | 768 |
| p38 | gagaacgtcattgggcttctgg | gtgttgtttccattcgtccacgc | 60 | 30-32 | 780 |
| PW1 | catgagccgaagtgggagag | gtggccttgtgtcatgtgag | 60 | 25-32 | 399 |
| Serca2 | ggcatctttgggcaggatg | cctcaacagcagccacaatg | 60 | 30-32 | 1239 |
| Serpini1 | gtgatggatccaatgaggctgg | gcacagccatcctactaatcgc | 60-65 | 40-41 | 411 |
| skNAC | ccagccgatcctcgtcaac | ccttccttgccttcttctcac | 60-65 | 29-37 | 260 |
| Smyd1(1-6) | atggagaacgtggaggtcttc | gctcacaggagcagtcaaagtag | 60-62 | 25-29 | 816/424 |
| Smyd1(2-6) | ctgcgggcagtgcaagtt | gctcacaggagcagtcaaagtag | 60-62 | 25-37 | 640 |
| Smyd1(4-6) | caactgcaacggtttcactctcag | gctcacaggagcagtcaaagtag | 60-62 | 30 | 260/220 |
| Smyd1(6-9) | acatagacttcctgcacctcagtg | catgggtcaccaggagaatagcat | 60-62 | 30 | 504 |
| Smyd1(7-9) | cgagggtttgtatcacgaggttgt | catgggtcaccaggagaatagcat | 60-62 | 30 | 300 |
| Smyd2 | ctggacaagctagacaacg | tgatcttctgcccgtatttcaggg | 62 | 28-32 | 700 |
| Smyd3 | gagtcgtggcgtagtctgtgatc | cagctgagacgcatcctggatc | 62 | 28-32 | 360 |

| | | | | | |
|---|---|---|---|---|---|
| Sprrl3 | cactcaccttccgaggtatcc | ctgccaccactgctacaac | 65 | 36-41 | 365 |
| TNFa | ggcctccctctcatcagttc | gagatagcaaatcggctgacgg | 60 | 25-35 | 293 |
| Tnni1 | ggagcaggaacacgagga | actggccgctccttttct | 60 | 35-37 | 377 |
| UCP-1 | gaaggtcagaatgcaagccc | catggacatcgcacagcttg | 58 | 30 | 364 |
| Zfp36 | ccatggatctctctgccatctacg | caggcctggttagggtctc | 60 | 25-35 | 724 |
| Zfy | tcctgagcatgagcaacagatgga | ctgactggtactgtttggattcagg | 60 | 31 | 217 |

## Bibliography


1. Shi, Y., *Histone lysine demethylases: emerging roles in development, physiology and disease.* Nat Rev Genet, 2007. **8**(11): p. 829-33.
2. Terranova, R., et al., *The reorganisation of constitutive heterochromatin in differentiating muscle requires HDAC activity.* Exp Cell Res, 2005. **310**(2): p. 344-56.
3. Zhu, L., *Functional characterization of Smyd1, a methyltransferase essential for heart and skeletal muscle development* in *Microbiology*. 2006, The University of Texas Austin. p. 117.
4. Abu-Farha, M., et al., *The tale of two domains: Proteomic and genomic analysis of SMYD2, a new histone methyltransferase.* Mol Cell Proteomics, 2007.
5. Nishioka, K., et al., *PR-Set7 is a nucleosome-specific methyltransferase that modifies lysine 20 of histone H4 and is associated with silent chromatin.* Mol Cell, 2002. **9**(6): p. 1201-13.
6. Sims, R.J., *Functional characterization of m-Bop, a transcriptional repressor essential for heart development*, in *Cell and Molecular Biology*. 2002, University of Texas: Austin. p. 122.
7. Pierce, S.A., *Transcriptional regulation of cardiac ventricular development*. 2004, The University of Texas Southwestern Medical Center: Dallas.
8. Sims, R.J., 3rd, et al., *m-Bop, a repressor protein essential for cardiogenesis, interacts with skNAC, a heart- and muscle-specific transcription factor.* J Biol Chem, 2002. **277**(29): p. 26524-9.
9. Yotov, W.V. and R. St-Arnaud, *Differential splicing-in of a proline-rich exon converts alphaNAC into a muscle-specific transcription factor.* Genes Dev, 1996. **10**(14): p. 1763-72.
10. Gottlieb, P.D., et al., *Bop encodes a muscle-restricted protein containing MYND and SET domains and is essential for cardiac differentiation and morphogenesis.* Nat Genet, 2002. **31**(1): p. 25-32.
11. Rosamond, W., et al., *Heart Disease and Stroke Statistics 2008 Update. A Report From the American Heart Association Statistics Committee and Stroke Statistics Subcommittee.* Circulation, 2007.
12. Waldo, K.L., et al., *Conotruncal myocardium arises from a secondary heart field.* Development, 2001. **128**(16): p. 3179-88.
13. Mjaatvedt, C.H., et al., *The outflow tract of the heart is recruited from a novel heart-forming field.* Dev Biol, 2001. **238**(1): p. 97-109.
14. Kelly, R.G., N.A. Brown, and M.E. Buckingham, *The arterial pole of the mouse heart forms from Fgf10-expressing cells in pharyngeal mesoderm.* Dev Cell, 2001. **1**(3): p. 435-40.
15. Chen, H., et al., *BMP10 is essential for maintaining cardiac growth during murine cardiogenesis.* Development, 2004. **131**(9): p. 2219-31.
16. Bruneau, B.G., *Transcriptional regulation of vertebrate cardiac morphogenesis.* Circ Res, 2002. **90**(5): p. 509-19.

17. Latinkic, B.V., S. Kotecha, and T.J. Mohun, *Induction of cardiomyocytes by GATA4 in Xenopus ectodermal explants.* Development, 2003. **130**(16): p. 3865-76.
18. Lien, C.L., et al., *Control of early cardiac-specific transcription of Nkx2-5 by a GATA-dependent enhancer.* Development, 1999. **126**(1): p. 75-84.
19. Saga, Y., et al., *MesP1 is expressed in the heart precursor cells and required for the formation of a single heart tube.* Development, 1999. **126**(15): p. 3437-47.
20. Nemer, M., *Genetic insights into normal and abnormal heart development.* Cardiovasc Pathol, 2008. **17**(1): p. 48-54.
21. Bruneau, B.G., et al., *A murine model of Holt-Oram syndrome defines roles of the T-box transcription factor Tbx5 in cardiogenesis and disease.* Cell, 2001. **106**(6): p. 709-21.
22. Klaus, A., et al., *Distinct roles of Wnt/beta-catenin and Bmp signaling during early cardiogenesis.* Proc Natl Acad Sci U S A, 2007. **104**(47): p. 18531-6.
23. O'Brien, T.X., K.J. Lee, and K.R. Chien, *Positional specification of ventricular myosin light chain 2 expression in the primitive murine heart tube.* Proc Natl Acad Sci U S A, 1993. **90**(11): p. 5157-61.
24. Cai, C.L., et al., *Isl1 identifies a cardiac progenitor population that proliferates prior to differentiation and contributes a majority of cells to the heart.* Dev Cell, 2003. **5**(6): p. 877-89.
25. Srivastava, D., et al., *Regulation of cardiac mesodermal and neural crest development by the bHLH transcription factor, dHAND.* Nat Genet, 1997. **16**(2): p. 154-60.
26. Benchabane, H. and J.L. Wrana, *GATA- and Smad1-dependent enhancers in the Smad7 gene differentially interpret bone morphogenetic protein concentrations.* Mol Cell Biol, 2003. **23**(18): p. 6646-61.
27. Liu, W., et al., *Bmp4 signaling is required for outflow-tract septation and branchial-arch artery remodeling.* Proc Natl Acad Sci U S A, 2004. **101**(13): p. 4489-94.
28. Qi, X., et al., *Essential role of Smad4 in maintaining cardiomyocyte proliferation during murine embryonic heart development.* Dev Biol, 2007. **311**(1): p. 136-46.
29. van Wijk, B., A.F. Moorman, and M.J. van den Hoff, *Role of bone morphogenetic proteins in cardiac differentiation.* Cardiovasc Res, 2007. **74**(2): p. 244-55.
30. Hasty, P., et al., *Muscle deficiency and neonatal death in mice with a targeted mutation in the myogenin gene.* Nature, 1993. **364**(6437): p. 501-6.
31. Nabeshima, Y., et al., *Myogenin gene disruption results in perinatal lethality because of severe muscle defect.* Nature, 1993. **364**(6437): p. 532-5.
32. Rudnicki, M.A., et al., *MyoD or Myf-5 is required for the formation of skeletal muscle.* Cell, 1993. **75**(7): p. 1351-9.
33. Arnold, H.H. and T. Braun, *Targeted inactivation of myogenic factor genes reveals their role during mouse myogenesis: a review.* Int J Dev Biol, 1996. **40**(1): p. 345-53.
34. Kassar-Duchossoy, L., et al., *Mrf4 determines skeletal muscle identity in Myf5:Myod double-mutant mice.* Nature, 2004. **431**(7007): p. 466-71.

35. Borlak, J. and T. Thum, *Hallmarks of ion channel gene expression in end-stage heart failure.* FASEB J, 2003. **17**(12): p. 1592-608.
36. Hama, N., et al., *Rapid ventricular induction of brain natriuretic peptide gene expression in experimental acute myocardial infarction.* Circulation, 1995. **92**(6): p. 1558-64.
37. McDonagh, T.A., et al., *Biochemical detection of left-ventricular systolic dysfunction.* Lancet, 1998. **351**(9095): p. 9-13.
38. Uusimaa, P., et al., *Plasma vasoactive peptides after acute myocardial infarction in relation to left ventricular dysfunction.* Int J Cardiol, 1999. **69**(1): p. 5-14.
39. Daniels, L.B. and A.S. Maisel, *Natriuretic peptides.* J Am Coll Cardiol, 2007. **50**(25): p. 2357-68.
40. Li, P., et al., *Atrial Natriuretic Peptide Inhibits Transforming Growth Factor {beta} Induced Smad Signaling and Myofibroblast Transformation in Mouse Cardiac Fibroblasts.* Circ Res, 2007.
41. de Boer, R.A., et al., *Early expression of natriuretic peptides and SERCA in mild heart failure: association with severity of the disease.* Int J Cardiol, 2001. **78**(1): p. 5-12.
42. Costantini, D.L., et al., *The homeodomain transcription factor Irx5 establishes the mouse cardiac ventricular repolarization gradient.* Cell, 2005. **123**(2): p. 347-58.
43. McFadden, D.G., et al., *A GATA-dependent right ventricular enhancer controls dHAND transcription in the developing heart.* Development, 2000. **127**(24): p. 5331-41.
44. Cheng, T.C., et al., *Separable regulatory elements governing myogenin transcription in mouse embryogenesis.* Science, 1993. **261**(5118): p. 215-8.
45. Rugh, R., *The mouse; its reproduction and development*. 1968, Minneapolis: Burgess Pub. Co. iv, 430.
46. Cartharius, K., et al., *MatInspector and beyond: promoter analysis based on transcription factor binding sites.* Bioinformatics, 2005. **21**(13): p. 2933-42.
47. Gao, J., Z. Li, and D. Paulin, *A novel site, Mt, in the human desmin enhancer is necessary for maximal expression in skeletal muscle.* J Biol Chem, 1998. **273**(11): p. 6402-9.
48. Xu, Q., et al., *Identification of zinc finger binding protein 89 (ZBP-89) as a transcriptional activator for a major bovine growth hormone receptor promoter.* Mol Cell Endocrinol, 2006. **251**(1-2): p. 88-95.
49. Song, J., et al., *Transcriptional regulation by zinc-finger proteins Sp1 and MAZ involves interactions with the same cis-elements.* Int J Mol Med, 2003. **11**(5): p. 547-53.
50. Johnsen, O., et al., *Interaction of the CNC-bZIP factor TCF11/LCR-F1/Nrf1 with MafG: binding-site selection and regulation of transcription.* Nucleic Acids Res, 1998. **26**(2): p. 512-20.
51. Azcoitia, V., et al., *The homeodomain protein Meis1 is essential for definitive hematopoiesis and vascular patterning in the mouse embryo.* Dev Biol, 2005. **280**(2): p. 307-20.

52. Knoepfler, P.S., et al., *A conserved motif N-terminal to the DNA-binding domains of myogenic bHLH transcription factors mediates cooperative DNA binding with pbx-Meis1/Prep1.* Nucleic Acids Res, 1999. **27**(18): p. 3752-61.
53. Heidt, A.B., et al., *Determinants of myogenic specificity within MyoD are required for noncanonical E box binding.* Mol Cell Biol, 2007. **27**(16): p. 5910-20.
54. Bartholin, L., et al., *TGIF inhibits retinoid signaling.* Mol Cell Biol, 2006. **26**(3): p. 990-1001.
55. Funk, W.D. and W.E. Wright, *Cyclic amplification and selection of targets for multicomponent complexes: myogenin interacts with factors recognizing binding sites for basic helix-loop-helix, nuclear factor 1, myocyte-specific enhancer-binding factor 2, and COMP1 factor.* Proc Natl Acad Sci U S A, 1992. **89**(20): p. 9484-8.
56. McKinsey, T.A., C.L. Zhang, and E.N. Olson, *MEF2: a calcium-dependent regulator of cell division, differentiation and death.* Trends Biochem Sci, 2002. **27**(1): p. 40-7.
57. Zhang, S.X., et al., *Identification of direct serum-response factor gene targets during Me2SO-induced P19 cardiac cell differentiation.* J Biol Chem, 2005. **280**(19): p. 19115-26.
58. Dong, K.W., et al., *The POU homeodomain protein Oct-1 binds Cis-regulatory element essential for the human GnRH upstream promoter activity in JEG-3 cells.* J Clin Endocrinol Metab, 2001. **86**(6): p. 2838-44.
59. Fukuda, T., et al., *The murine BCL6 gene is induced in activated lymphocytes as an immediate early gene.* Oncogene, 1995. **11**(8): p. 1657-63.
60. Kuisk, I.R., et al., *A single MEF2 site governs desmin transcription in both heart and skeletal muscle during mouse embryogenesis.* Dev Biol, 1996. **174**(1): p. 1-13.
61. Chen, F., et al., *Hop is an unusual homeobox gene that modulates cardiac development.* Cell, 2002. **110**(6): p. 713-23.
62. Berkes, C.A. and S.J. Tapscott, *MyoD and the transcriptional control of myogenesis.* Semin Cell Dev Biol, 2005. **16**(4-5): p. 585-95.
63. Lim, C.P. and X. Cao, *Structure, function, and regulation of STAT proteins.* Mol Biosyst, 2006. **2**(11): p. 536-50.
64. Weisser, M., et al., *Reverse transcriptase-polymerase chain reaction based quantification of the combined MDS-EVI1/EVI1 gene in acute myeloid leukemia.* Leuk Lymphoma, 2006. **47**(12): p. 2645-7.
65. Schwartz, R.J. and E.N. Olson, *Building the heart piece by piece: modularity of cis-elements regulating Nkx2-5 transcription.* Development, 1999. **126**(19): p. 4187-92.
66. Potthoff, M.J., et al., *Regulation of skeletal muscle sarcomere integrity and postnatal muscle function by Mef2c.* Mol Cell Biol, 2007. **27**(23): p. 8143-51.
67. Phan, D., et al., *BOP, a regulator of right ventricular heart development, is a direct transcriptional target of MEF2C in the developing heart.* Development, 2005. **132**(11): p. 2669-78.

68. Wadman, I.A., et al., *The LIM-only protein Lmo2 is a bridging molecule assembling an erythroid, DNA-binding complex which includes the TAL1, E47, GATA-1 and Ldb1/NLI proteins.* EMBO J, 1997. **16**(11): p. 3145-57.
69. Jacks, T., et al., *Tumour predisposition in mice heterozygous for a targeted mutation in Nf1.* Nat Genet, 1994. **7**(3): p. 353-61.
70. Levin, S.I. and M.H. Meisler, *Floxed allele for conditional inactivation of the voltage-gated sodium channel Scn8a (NaV1.6).* Genesis, 2004. **39**(4): p. 234-9.
71. Sauer, B. and N. Henderson, *Site-specific DNA recombination in mammalian cells by the Cre recombinase of bacteriophage P1.* Proc Natl Acad Sci U S A, 1988. **85**(14): p. 5166-70.
72. Sauer, B., *Inducible gene targeting in mice using the Cre/lox system.* Methods, 1998. **14**(4): p. 381-92.
73. Dymecki, S.M., *Flp recombinase promotes site-specific DNA recombination in embryonic stem cells and transgenic mice.* Proc Natl Acad Sci U S A, 1996. **93**(12): p. 6191-6.
74. Lakso, M., et al., *Efficient in vivo manipulation of mouse genomic sequences at the zygote stage.* Proc Natl Acad Sci U S A, 1996. **93**(12): p. 5860-5.
75. McFadden, D.G., et al., *The Hand1 and Hand2 transcription factors regulate expansion of the embryonic cardiac ventricles in a gene dosage-dependent manner.* Development, 2005. **132**(1): p. 189-201.
76. Agah, R., et al., *Gene recombination in postmitotic cells. Targeted expression of Cre recombinase provokes cardiac-restricted, site-specific rearrangement in adult ventricular muscle in vivo.* J Clin Invest, 1997. **100**(1): p. 169-79.
77. Li, S., et al., *Requirement for serum response factor for skeletal muscle growth and maturation revealed by tissue-specific gene deletion in mice.* Proc Natl Acad Sci U S A, 2005. **102**(4): p. 1082-7.
78. Traut, W., et al., *The temporal and spatial distribution of the proliferation associated Ki-67 protein during female and male meiosis.* Chromosoma, 2002. **111**(3): p. 156-64.
79. Bruce, C.R. and D.J. Dyck, *Cytokine regulation of skeletal muscle fatty acid metabolism: effect of interleukin-6 and tumor necrosis factor-alpha.* Am J Physiol Endocrinol Metab, 2004. **287**(4): p. E616-21.
80. Chen, S.E., B. Jin, and Y.P. Li, *TNF-alpha regulates myogenesis and muscle regeneration by activating p38 MAPK.* Am J Physiol Cell Physiol, 2007. **292**(5): p. C1660-71.
81. Kandarian, S.C. and R.W. Jackman, *Intracellular signaling during skeletal muscle atrophy.* Muscle Nerve, 2006. **33**(2): p. 155-65.
82. Baeza-Raja, B. and P. Munoz-Canoves, *p38 MAPK-induced nuclear factor-kappaB activity is required for skeletal muscle differentiation: role of interleukin-6.* Mol Biol Cell, 2004. **15**(4): p. 2013-26.
83. Teboul, L., et al., *Thiazolidinediones and fatty acids convert myogenic cells into adipose-like cells.* J Biol Chem, 1995. **270**(47): p. 28183-7.
84. Minamino, T., et al., *Inducible gene targeting in postnatal myocardium by cardiac-specific expression of a hormone-activated Cre fusion protein.* Circ Res, 2001. **88**(6): p. 587-92.

85. Kurrelmeyer, K.M., et al., *Endogenous tumor necrosis factor protects the adult cardiac myocyte against ischemic-induced apoptosis in a murine model of acute myocardial infarction.* Proc Natl Acad Sci U S A, 2000. **97**(10): p. 5456-61.
86. Funakoshi, H., et al., *Involvement of inducible nitric oxide synthase in cardiac dysfunction with tumor necrosis factor-alpha.* Am J Physiol Heart Circ Physiol, 2002. **282**(6): p. H2159-66.
87. Sun, M., et al., *Excessive tumor necrosis factor activation after infarction contributes to susceptibility of myocardial rupture and left ventricular dysfunction.* Circulation, 2004. **110**(20): p. 3221-8.
88. Fuchs, M., et al., *Role of interleukin-6 for LV remodeling and survival after experimental myocardial infarction.* FASEB J, 2003. **17**(14): p. 2118-20.
89. Watsuji, T., et al., *Controlled gene expression with a reverse tetracycline-regulated retroviral vector (RTRV) system.* Biochem Biophys Res Commun, 1997. **234**(3): p. 769-73.
90. Ponomareva, O.N., et al., *Defective neuromuscular synaptogenesis in mice expressing constitutively active ErbB2 in skeletal muscle fibers.* Mol Cell Neurosci, 2006. **31**(2): p. 334-45.

# VITA

Tara L Rasmussen was born in Peoria, Illinois on October 22, 1979, the daughter of Ginger and Peter Rasmussen. After graduating Morton High School in May 1997, she attended the University of Illinois. She received a Howard Hughes Undergraduate Research Fellowship at the University of Illinois during the summer of 1999, which extended through the spring of 2000. In May of 2001, she received the degree of Bachelor of Science from the school of Liberal Arts and Sciences at the University of Illinois. In August 2001, she entered the Graduate School at the University of Texas at Austin in the Institute of Cellular and Molecular Biology. She rotated in the lab of Paul D. Gottlieb in the Spring 2002 and officially joined his lab in May 2002. In December 2003, she continued her project under the guidance of Phillip W. Tucker. Tara was married to Todd Tibbs in March, 2008. During graduate school she worked as a teaching assistant for the undergraduate virology laboratory. The author has 1 publication.

Permanent Address: 6307 Shadow Valley Dr. Unit A, Austin, TX 78731

This dissertation was typed by the author.