Building A Gene Regulatory Network in Adult Mouse Skeletal Muscle Following Nerve Injury: Transcriptome Characterization and New Model for Functional *cis*-Regulatory Analysis *In Vivo*

Thesis by

Gilberto Hernandez, M.D.

In Partial Fulfillment of the Requirements
for the Degree of
Doctor of Philosophy

California Institute of Technology
Pasadena, California

2013

(Defended July 23, 2012)

© 2013
Gilberto Hernandez, M.D.
All Rights Reserved

To my parents, Tomasa and Gilberto Hernandez, who have provided me with the ability and motivation to always strive to live life beautifully.

Acknowledgments

I would like to thank my advisor, Barbara J. Wold, for supporting my research, despite it not being directly in accordance with the lab's focus. I have learned a great deal from our partnership. I would also like to thank, Eric H. Davidson, for literally providing me with a key to his laboratory and, by default, granting me access to the world of Gene Regulatory Networks and his team of highly talented scientists from which to learn. Special thanks to Marianne Bronner, who was my first mentor at Caltech and not only welcomed me warmly to the world of Caltech research, but provided invaluable understanding and guidance during my early graduate years in my quest to find a research focus that resonated best with my sensibilities. Naturally, I thank Professors Alexander Varshavsky and Paul Sternberg for guidance in their roles as members of my thesis committee.

I have had the privilege of learning from very talented scientists, including my first teacher of molecular biology: Leslie Dunipace; my current teacher of molecular biology: Brian Williams; my bioinformatics tutors: Christopher Hart, Henry Amrhein, Diane Trout, Anna Cecilia Therese Abelin and the rest of the Wold lab; my GRN instructors: the Davidson lab for their excellent INFOs/research/and discussions; Roger Revilla-i-Domingo, Qiang Tu, Jongmin Nam, Emmanuel Faure, Julie Hahn, Erika Vielmas, and Ping Dong all took time from their work to help me with mine. To Smadar Ben-Tabou De-Leon, Jane Rigg and Deanna Thomas for their strong and positive support throughout the years. To Miao Cui without whom I would never have arrived to any Davidson lab social event. My collaborators from the Pierce lab: Harry Choi and Colby Calvert; and from the Fraser lab: Periklis Pantazis, Thai

Truong, and Sarah Sweeney for guidance with confocal and two-photon imaging. To Igor Antoshechkin and Vijaya Kumar for assistance with RNAseq library construction and sequencing.

My *in vivo* gene transfer and skeletal muscle physiology mentors from UCLA: Julio Vergara, Marino DiFranco, and Marbella Quinonez. This project would not have been possible without the expertise provided by Professor Julio Vergara. Thanks also to Raul Serrano for his electrical engineering prowess and Phillip Tran for his technical assistance with electroporation into flexor digitorum brevis and interosseous as well as muscle fiber dissociation.

Special thanks to Libera Berghella, whose research on the *myogenin*-highly conserved element (HCE) and Msy3 protein set the stage for my dissertation work. Her patient guidance and generosity with knowledge and reporter constructs greatly assisted me throughout the development of my project. To Edoardo "Dado" Marcora for expert assistance in every possible scientific area that my project entered. To Holly Carlisle for serving as my cultural attache to the scientific community of Caltech. To Sagar Damle for his assistance in helping me overcome countless bioinformatics and experiment-related obstacles in these last few months, including defense seminar organization and for being pretty much my only source of human contact in the last 8 weeks.

I am forever grateful to Drs. Eugene Seymour, Gerald S. Levey, and Jennie B. Dorman. Dr. Seymour took the time to get to know me during a morning bike ride when I was an undergraduate at UCLA and kick started my journey towards becoming a physican-scientist through his mentorship and unconditional support. Dr. Levey took the time to get to know me during a medical student welcome BBQ and was the first person to actively stand behind my desire to explore my research interests as a member of the Caltech/UCLA Medical Scientist Training Program. Dr. Dorman has been a part of my scientific journey from even before Dr. Seymour, in the days when I studied skeletal muscle physiology by seeing how strong and fast I could make my

own muscles function. I have received nothing but support and encouragement from her through the years and her help in the construction of this written dissertation, from editing to bringing my hands to the keyboard every time I was running out of steam physically and emotionally, has been invaluable and awe inspiring.

Finally, thanks to my close family and friends (many of whom have already been mentioned above) for believing in me, appreciating my effort and for being at the finish line of this most challenging event.

Abstract

The essential functional linkages of gene regulatory networks (GRNs) consist of the interactions between *cis*-regulatory DNA sequences and *trans*-acting regulatory factors. These genomically encoded regulatory interactions govern the differential gene expression programs which direct specific biological processes during development and adulthood. Detailed analysis of GRNs during development has yielded important insights regarding the structural and functional dynamics of *cis*-regulatory modules (CRMs) and *cis*-regulatory elements (CREs). Indeed, the comprehensive GRNs that have been characterized for various developmental processes provide a model for both the methodological approach and the intellectual understanding required to explore *cis*-regulatory architecture in other biological contexts. The present study focuses on the physiological context, investigating the GRNs that govern the molecular response to nerve injury in adult mouse skeletal muscle. Until now, high-quality GRN investigations in this context have been hampered by the absence of two fundamental components: a comprehensive catalog of genes differentially expressed after nerve injury, and an effective *in vivo* gene transfer technique to functionally test putative *cis*-regulatory modules. Using RNAseq, we have compiled a comprehensive list of all differentially expressed genes at 6.0, 12.0, 24.0, and 168.0 hours following nerve injury. This data has validated previously known differentially expressed genes, as well as identified novel candidates for *cis*-regulatory analysis. The *in vivo* gene transfer technique I have adapted and advanced targets an easily accessible muscle group for minimally invasive injection and electroporation of DNA; with it, I demonstrate highly efficient, reproducible, and stable gene transfer in mouse skeletal muscle. In addition,

I have optimized the gene transfer technique not only for plasmid DNA reporter vectors, but also for BAC DNA reporter vectors, thus enabling *cis*-regulatory modules to be tested in a broad chromosomal environment. Finally, I have validated the capacity of this gene transfer method to functionally test CRMs, by identifying a nerve injury-associated CRM of the skeletal muscle-specific *myogenin* gene. The enhanced resolution provided by this technique allowed for qualitative and quantitative detection of increased reporter signal from a mutated version the nerve injury-associated CRM at ten days following denervation, when compared to the wild-type CRM, implicating it as the *cis*-acting regulatory sequence responsible for mediating the down-regulation of *myogenin* during late phase neurogenic skeletal muscle atrophy. This work lays the foundation from which a high-quality adult skeletal muscle GRN can be constructed for nerve injury and other muscle-associated disease states.

Contents

Acknowledgments **iv**

Abstract **vii**

1 Introduction **1**

1.1 FDB and IO as a model system for investigation of GRNs in adult skeletal muscle 4

1.2 Gene regulatory networks of limb skeletal muscle in the embryo and adult 7

1.2.1 Regulatory circuitry during delamination and migration of progenitor cells to the limb bud 8

1.2.2 Regulatory circuitry of hind limb muscle development in the embryo 10

1.2.3 Regulatory circuitry of adult hind limb muscle 14

1.2.4 Regulatory circuitry of denervation response in adult fast-twitch hind limb muscle 15

2 Functional testing of the *myogenin*-highly conserved *cis*-acting element (HCE) during late phase neurogenic atrophy using a new model for functional *cis*-regulatory analysis *in vivo* **19**

2.1 Introduction 19

2.2 Results 21

2.2.1 Gene transfer of plasmid and BAC DNA towards functional testing of *cis*-regulatory modules 21

2.2.2 Testing the myogenin-highly conserved *cis*-regulatory element (HCE) for regulatory activity at 10.0 days following denervation 25

2.2.3 Denervation of of hind limb in the absence of Msy3 at 17.0 days after denervation 27

2.3 Discussion 29

2.4 Materials and Methods 33

2.4.1 *Plasmid and BAC DNA constructs* 33

2.4.2 *Gene transfer* 33

2.4.3 *Sciatic nerve denervation and mock denervation surgery* . . . 35

2.4.4 *Total RNA isolation* 35

2.4.5 *Imaging* 36

3 High-resolution transcriptome in adult skeletal muscle following nerve injury 37

3.1 Introduction 37

3.2 Results 39

3.2.1 Differential expression in control vs. mock denervated muscle (at 6.0, 12.0, 24.0, and 168.0 hours), using RNAseq 39

3.2.2 Differential expression in mock denervated vs. denervated muscle (at 6.0, 12.0, 24.0, and 168.0 hours), using RNAseq 43

3.3 Discussion 48

3.4 Materials and Methods 51

3.4.1 *RNAseq* 51

4 Conclusion 53

Bibliography 55

Appendices 75

A Control vs. mock denervated muscle: top 100 differentially expressed genes 76

B Mock denervated vs. denervated muscle: top 100 differentially expressed genes 105

Chapter 1

Introduction

Comprehensive, validated gene regulatory network (GRN) models during early sea urchin development have provided an unprecedented perspective on how the animal body plan is encoded in the DNA [Davidson et al., 2002, Su et al., 2009, Peter and Davidson, 2009, Oliveri et al., 2008]. To date, GRNs have been implicated in controlling a wide range of biological processes, including morphogenesis, differentiation, and physiological response in various model systems [Davidson, 2006]. The essential functional linkages between *cis*-acting regulatory DNA sequences, on the one hand, and factors acting in *trans*, serve as the 'weight bearing columns' of a Gene Regulatory Network (GRN). These are the causal interactions that respond to a given regulatory state and, in turn, determine subsequent regulatory states, all of which contribute to the dynamic nature of biological responses. Functional testing of *cis*-regulatory modules (CRMs) via gene transfer of candidate *cis*-regulatory DNA sequence-driven reporter genes, reveals the general *cis*-regulatory architecture encoded in the DNA. Perturbation of validated CRMs provides further resolution of the *cis*-regulatory architecture by identifying the *cis*-regulatory elements (CREs) within each CRM. In all cases where the GRNs controlling a biological response have been well characterized, the validation of genomic network architecture and dynamics has required effective application of GRN bioscience methodologies. Central to these methodologies is the experimental validation of *cis*-acting regulatory DNA sequences via reporter gene

transfer [Revilla-i Domingo et al., 2004, Davidson et al., 2002]. To date, however, the application of GRN bioscience to adult skeletal muscle biology has been limited by existing *in vivo* gene transfer techniques, which suffer from poor or variable transfection efficiencies and often result in excessive tissue trauma. Researchers interested in adult skeletal muscle biology at the molecular-level have made many attempts at *in vivo* gene transfer into adult skeletal muscle, including both trans-epidermal and direct contact muscle-to-electrode electroporation, as well as ultrasound and viral vector infection [Taniyama et al., 2002, Mir et al., 1999, Wolff et al., 1990]. All these methods introduce confounding variables when attempted in adult skeletal muscle. In the case of electroporation, high voltage and/or the invasive survival surgery that is required to access muscle tissue result in significant trauma to target muscles and neighboring tissues [Lefesvre et al., 2002, Hartikka et al., 2001]. Viral-mediated transfection has proven effective for making transgenic mice, but its application in transient transfection experiments suffers from labor-intensive viral preparations and nonspecific infection of nontarget tissues, with resulting immune response [Lefesvre et al., 2002, Dai et al., 1995].

Consequently, the majority of the data on the *cis*-regulatory architecture of skeletal muscle centers around muscle cell differentiation, from *in vitro* experiments where gene transfer is most efficient and from *in vivo* transgenic experiments in the developing mouse embryo where reporter gene spatial and temporal expression is most easily observable [Weintraub et al., 1991b,a, Tapscott et al., 1988, Edmondson and Olson, 1989, Miner and Wold, 1990, Braun et al., 1989]. Otherwise, the majority of the regulatory network data data for adult skeletal muscle in mouse, comes from transgenic loss of function or gain of function experiments. Indeed, transgenesis has proven to be the most successful and reliable method for reporter gene transfer in adult mouse muscle. Currently, the experimental path towards functional *cis*-regulatory analysis in adult mouse skeletal muscle begins with a set of differentially expressed target genes, followed by inference of CRM function from differentiation-related *in vitro* experiments or poorly resolved *in vivo* transient transfections. Next, researchers

create a transgenic mouse to test CRM function in a physiologically relevant context. However, the cost and time required to produce a transgenic mouse is significant, especially considering that only one or two *cis*-regulatory modules can be tested at a time in this way. Thus, although genetic studies have identified many of the key regulatory and effector molecules involved in various adult skeletal muscle physiological events, how these molecules interact at the *cis* and *trans* level is still largely unknown.

My objective is to determine the genomically encoded regulatory interactions that control the physiological response capabilities of the adult skeletal muscle nerve injury-response. The first step necessary to meet this objective, is a more efficient and cost effective experimental approach to uncovering the GRNs that underlie the physiological response in adult mouse skeletal muscle. An ideal experimental pipeline would begin with a robust, stable, efficient, cost effective, short time-to-assay, and *in vivo* method for reporter gene transfer. This would allow for functional testing and reliable measurement of the reporter signals of candidate CRMs and CREs, all while in a biologically relevant context. I propose electroporation-mediated gene transfer into flexor digitorum brevis (FDB) and interosseous (IO) as an ideal *in vivo* experimental model for functional testing of candidate CRMs/CREs. In Chapter 2, I apply this experimental model to identify a functional CRE of *myogenin* that contributes to the down-regulation of *myogenin* at 10.0 days following denervation [Moresi et al., 2010]. Once a method to experimentally validate the functional CRMs and CREs which interpret the dynamic regulatory states present in skeletal muscle following nerve injury is in place, the next step is to focus on the regulatory molecules that comprise these denervation-associated regulatory states. Of particular importance are the transcription factors and signaling molecules which determine gene expression. GRN studies during embryological development in a range of model systems have provided clear evidence that transcription factors and signaling molecules are the key molecular drivers for establishing, maintaining, or transitioning between different regulatory states [Davidson, 2006]. Thus, by knowing the identity of the full

repertoire of candidate regulatory molecules, one can track their temporal and spatial expression to deduce the kinetics of their regulatory activity. In Chapter 3, I present a high-resolution transcriptome-wide RNAseq analysis for flexor digitorum brevis (FDB) and interosseous (IO) muscles at 6.0, 12.0, 24.0, and 168.0 hrs post denervation. This work moves the field forward in two important ways. First, it represents the first transcriptome-wide expression analysis for denervated FDB and IO muscles using RNAseq. Prior to this study, transcriptome-wide analysis of denervated adult mouse hind limb skeletal muscles were done using microarray assays. Tibialis anterior, extensor digitorum longus, gastrocnemius and soleus have all been assayed by microarray following sciatic nerve resection [Batt et al., 2006, Raffaello et al., 2006, Kostrominova et al., 2005, Tang et al., 2009]. RNAseq-generated data, however, has significant advantages over microarray data, including very low background, a much wider dynamic range for detecting fold changes, and higher accuracy in terms of absolute quantitation of expression across technical and biological replicates [Wang et al., 2009, Fu et al., 2009].

1.1 FDB and IO as a model system for investigation of GRNs in adult skeletal muscle

The current repertoire of skeletal muscles selected for denervation-associated studies have proven to be non ideal for gene transfer via electroporation, despite the known effectiveness of electroporation for introducting DNA or other molecules into cells [Sugar and Neumann, 1984]. A muscle group more amenable to electroporation mediated gene transfer, which also avoids the confounding trauma associated with current electroporation into adult skeletal muscle, would overcome a significant limiter to *in vivo* functional testing of candiate CRMs/CREs. Since the pioneering studies of Galvani on frog legs in the 18th century, the skeletal muscles of the hind limb have served as the premier model system for research regarding nerve-muscle connec-

tivity in vertebrates (Figure 1.1A). Only as recently as the mid-20th century has technology advanced enough to successfully record and characterize the excitation-contraction coupling effect first observed by Galvani in 1771 [Galvani, 1791, Sandow, 1952].

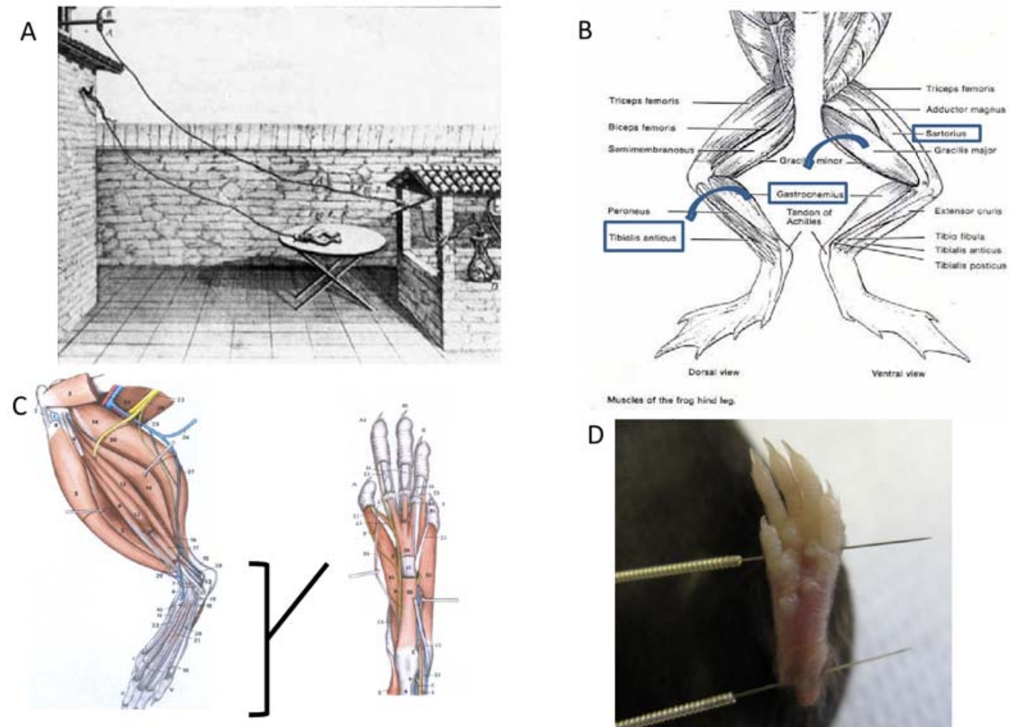

Figure 1.1: Experimental 'model muscle' groups. (A) Illustration of Galvani's experimental design for elliciting frog leg skeletal muscle contraction [Galvani, 1791]. (B) Anatomical distal trend to smaller muscles more ideal for advancing experimental technologies (sketch from Fastol.com). (C) Anatomical sketch of mouse hind limb and plantar surface views of FDB. Interosseous lies underneath the FDB and is not shown [Popesko P, Rajtova V, Horak J, Kapeller K, 1992]. (D) *in vivo* demonstration of sub-cutaneous electrode placement.

The longest muscle in the frog leg, sartorius, was the most amenable to the experimental method used in this seminal work. Since then, there has been an anatomical trend in the distal direction to the smaller skeletal muscles of the lower hind limb,

as model muscles for investigation (Figure 1.1B). Mouse gastrocnemius, plantaris, soleus, extensor digitorum longus, and tibialis anterior proved most amenable to the technological advances in experimental modalities of the second half of the 20th century. However, *in vivo* electroporation into these larger muscles of the lower limb requires surgical exposure of underlying muscle and/or high voltages to penetrate the thick muscle bundles. This results in inefficient gene transfer and confounding trauma in the tissue under study.

The turn of the century marked the experimental debut of another muscle group that proved even more amenable to voltage clamp and current studies [Friedrich et al., 1999]. This group, consisting of the flexor digitorum brevis (FDB) and interosseous (IO) muscles, is located on the plantar surface of the mouse paw, and the use of FDB and IO continues the trend of employing more distal and smaller muscles (Figure 1.1C). In 2009, DiFranco et al. [2009] took advantage of the smaller size and superficial anatomical location of FDB and IO by engineering an *in vivo* electroporation method for gene transfer into this muscle group (Figure 1.1D). This highly efficient gene transfer technique has enabled them to perform protein-level analysis of key components of excitation-contraction coupling and represents a significant step towards understanding the protein-level dynamics of electrophysiological events. In the present study, we have advanced this methodology to enable transcript-level GRN experimentation, in order to test putative *cis*-regulatory modules of key genes controlling the response to nerve injury in adult skeletal muscle.

The relatively superficial location of the FDB and IO muscle groups allows for subcutaneous access to the muscle tissue, via small-diameter needles, for both introduction of DNA solution and electrode placement. Furthermore, the compartment of the plantar surface of the mouse paw can accept a volume up to 25 µL; we have confirmed that a subcutaneous injection of 10 µL of a reporter DNA solution is sufficient to fully bathe the FDB/IO muscles. Small-diameter acupuncture needles serve as electrodes, which are placed subcutaneously at the toe line and at the heel at the plantar surface of the mouse paw. Much like the gel in an electrophoresis assay, the

FDB and IO muscles lie longitudinally between the two electrodes (Figure 1.1D). Once voltage is applied, the current travels along the entire longitudinal plane of the FDB/IO muscle bundles. The result is uniform muscle coverage of both DNA solution and the applied current. We have found that the FDB and IO of one mouse paw from a 2.0 to 4.0-month-old mouse consistently yields at least 0.5 µg of total RNA, which is sufficient for both RNAseq or QRT-PCR assays.

The FDB and IO muscles are innervated exclusively by tributaries of the sciatic nerve, as are the classically studied skeletal muscles of the hind limb used for denervation experiments. Therefore, sciatic nerve resection can remain as the method of choice to induce the denervation response. In addition, FDB and IO muscles share the same fiber type composition as the classically studied muscles used for denervation studies. They are all typeIIB-fiber-type dominant or fast-twitch muscles. This is important to ensure that we can build upon prior knowledge obtained from denervation studies in the classically studied fast-twitch muscles.

1.2 Gene regulatory networks of limb skeletal muscle in the embryo and adult

Skeletal muscles of the head, body and limb have distinct embryonic and cellular origins. The trunk and limb muscles are derived from the paraxial mesoderm and the muscles of the head originate from pre-chordal and pharyngeal head mesoderm. The myogenic regulatory factors (MRFs): *myf5*, *myoD*, *myf6*, and *myogenin*, eventually direct the determination and differentiation transcriptional programs of striated skeletal muscle. However, different regulatory circuits control the deployment of the MRFs in these different anatomical regions. The next level of regulatory circuit deployment, following primary determination and differentiation of skeletal muscle (i.e. embryonic myogenesis), involves establishing the future metabolic and contractile and regenerative properties of the muscles of specific anatomical compartments (i.e. fetal

myogenesis). Finally, during post natal development the regulatory circuitry gets further refined by inputs from innervating motor neurons. In essence, the underlying core genomic regulatory apparatus for striated skeletal muscle is partially revealed as we determine the input receiving capacity of each network-associated gene as it encounters varying regulatory states throughout development and into adulthood. We will highlight some of the regulatory circuitry that governs hind limb skeletal muscle development in fast-twitch muscle (i.e. the predominant muscle type of the mouse hind limb) to provide a context from which to better frame the regulatory logic that directs the denervation response in the adult skeletal muscle. Our attention will eventually focus specifically on *myogenin* and its down-regulation after the first week of skeletal muscle denervation.

1.2.1 Regulatory circuitry during delamination and migration of progenitor cells to the limb bud

The limb musculature develops from progenitor cells that delaminate from the hypaxial region of the dermomyotome and and migrate to the limb bud. During the migration to the limb bud these progenitor cells proliferate, but do not yet differentiate. At this time they do not yet express MRFs or muscle specific proteins. The regulatory circuitry active at this point involves expression of *six1/4*, *eya1/2*, *pax3*, *c-Met*, *lbx1*, and *gab1* in the migrating progenitor cells, along with hepatocyte growth factor (HGF), secreted from neighboring tissues. Loss of function experiments confirm that Six1 and Six4 control expression of *pax3* [Grifone et al., 2005] and that Pax3 partially regulates *c-Met* [Bober et al., 1994, Goulding et al., 1994, Tajbakhsh et al., 1997, Relaix et al., 2004, 2005]. The progenitor cells are unable to undergo normal delamination or migration without Six1/4 and Pax3. Pax3 also controls expression of the homeobox-containing transcription factor ladybird,*lbx1*, and when *lbx1* is inactivated, progenitor cells cannot migrate to the limb bud [Schäfer and Braun, 1999, Gross et al., 2000, Brohmann et al., 2000]. c-Met is a tyrosine kinase receptor for

HGF and Gab1 functions in the signaling cascade downstream of c-Met and HGF When *gab1* or *hgf* is mutated, no muscles are formed in the hind limb [Sachs et al., 2000, Dietrich et al., 1999]. The regulatory circuitry at this point, from approximately embryonic day 9.0 (E9.0) to 11.0 (E11.0) in the hind limb, centers around facilitating delamination, proliferation and migration while preventing activation of the myogenic differentiation program (Figure 1.2).

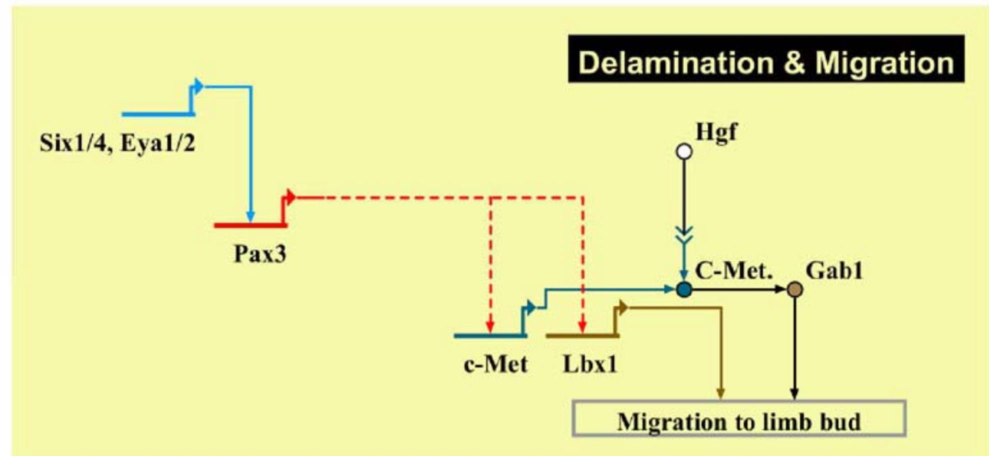

Figure 1.2: Progenitor cell regulatory subcircuit during delamination and migration to hind limb (E9.0 to E11.0). The myogenic differentiation program is not yet initiated until the progenitor cells reach the limb bud. Hgf/C-Met/Gab1 are part of a signaling mechanism that guides migration. Colored circles represent proteins of a signaling pathway and solid arrows connecting the circles represent the signaling hierarchy. Solid lines indicate a direct *cis* interaction and dotted lines indicate an indirect interaction.

1.2.2 Regulatory circuitry of hind limb muscle development in the embryo

Once progenitor cells reach the hind limb, signaling from sonic hedgehog (Shh) and bone morphogenetic protein 4 (BMP4) contribute to activation of the MRFs and the myogenic program [Krüger et al., 2001]. *six1/4*, *meox2*, *pax3*, *myf5*, *myf6*, *myoD* and *myogenin* are expressed in the progenitor cells along with regulatory co-factors (e.g. E-proteins, PBX/Meis, Eya1/2 and Mef2C) which comprise the regulatory circuitry driving primary myogenesis in the hind limb. Mutations in the homeobox gene *meox2* results in partial loss of muscle in the hind limb and down-regulation of *pax3* [Mankoo et al., 1999]. Six1/4 and Pax3 can directly activate *myf5* expression and *myf5* is the first of the MRFs to be expressed in the primary limb musculature [Giordani et al., 2007, Bajard et al., 2006]. Shortly thereafter *myoD* followed by *myogenin* and *myf6* are expressed. Six1/4 along with co-activators Eya1/2 can directly activate *myf5* once in the hind limb, whereas only Pax3 directly activated *myf5* in the dermomyotome. Mutation of *meox2* has no effect on *myoD* expression, while it does directly and indirectly affect *pax3* and *myf5* expression, respectively. In fact, myoD appears to be the primary MRF directing the myogenic program in hind limb muscle as evidenced by normal myogenesis in hind limb in *myf5*:*myf6* double mutants. In contrast, in the somites, activation of the myogenic program is delayed in *myf5*:*myf6* double mutants [Kablar et al., 1997]. Similarly, a distinct regulatory circuitry directs the temporal activation of MRFs in the trunk versus the hind limb. In the trunk, Myf6 is sufficient to direct normal myogenesis in the absence of myoD, since they are expressed at overlapping time points, whereas myogenesis is delayed in the hind limb of mutant *myoD* embryos, from E11.5 to E13.5, until *myf6* is expressed and can rescue the phenotype [Kablar et al., 1997]. Shh is thought to function in maintenance of expression of MRFs and expansion of embryonic myoblasts. Shh mutants display severe deficiencies in hind limb muscles and also result in reduced expression of *bmp4*. *in vitro* studies on explants from Shh mutant embryos showed that treatment with exogenous BMP4 increased the number of primary myotubes [Krüger et al., 2001].

These experiments suggest a regulatory hierarchy where Shh up-regulates *bmp4* and together they support the activation of the myogenic program towards the formation of primary myotubes in the hind limb. The primary muscle fibers express two contractile fiber types: fast embryonic myosin heavy chain (eMyHC) and slow MyHCβ. The regulatory circuitry at this point, from approximately E11.0 to E13.5, is centered around activation of the myogenic program and differentiation of embryonic myoblasts to primary myofibers with specific contractile properties (Figure 1.3).

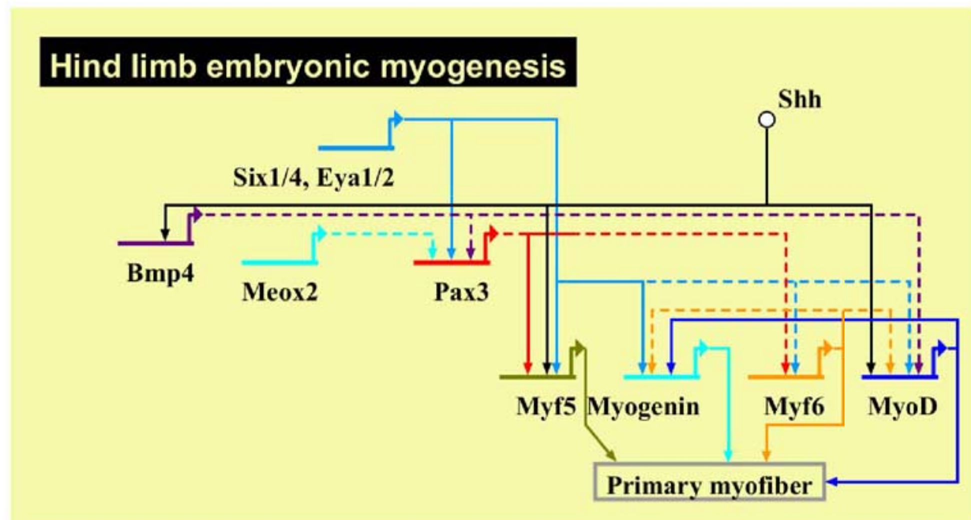

Figure 1.3: Embryonic Myogenesis. Once in the limb bud embryonic myoblasts activate the myogenic program. The white circle represents a signaling protein, Shh, acting to activate its target genes via signaling. The end result of this regulatory subcircuit is the primary myofiber. Solid lines indicate a direct *cis* interaction and dotted lines indicate an indirect interaction.

Fetal myoblasts which differentiate into secondary myotubes or myofibers during fetal development (E14.5-E17.5), follow a similar regulatory circuitry in terms of the deployment of MRFs along with their regulatory co-factors (e.g. E-proteins, PBX/Meis and Mef2C). Together the MRFs and co-factors initiate the myogenic program in fetal

myoblasts. However, *pax7*, as opposed to *pax3*, now helps to drive the differentiation regulatory program towards secondary myofiber formation [Hutcheson et al., 2009b]. *pax7* activates the myogenic program via *myoD* [Relaix et al., 2006]. *myf5* activates the myogenic program independently of *pax7*. This is in contrast to the regulatory dynamic during embryonic myogenesis, where *pax3* regulated *myf5* and *myf6*, but not *myoD* (Figure 1.3). It has been postulated that the regulatory program of fetal myoblasts is held at bay during the embryonic phase of muscle differentiation by signaling molecules, such as transforming growth factor beta (TGFβ) or β-catenin, which inhibit fetal myoblasts from differentiating, but not embryonic myoblasts during embryonic myogenesis [Cusella-De Angelis et al., 1994, Hutcheson et al., 2009a]. Secondary myofibers of the hind limb express eMyHC and perinatal MyHC. As is the case with primary myofibers, the contractile properties of the secondary myofibers at this point, as determined by the expression of different types of MyHC, occurs independently of nerve input. This suggests that the metabolic properties of the developing muscle fibers are encoded in the regulatory genome and, thus, potentially accessible to GRN analysis and discovery (Figure 1.4).

Satellite cells begin to appear midway through the fetal development phase. They take their place between the basal lamina and the developing myofibers and activate their myogenic program at this time. The regulatory circuitry of satellite cells of the hind limb is as Figure 1.4. They are *pax7* positive and activate *myoD* to initiate myogenesis. As with fetal myoblasts, TGFβ is implicated in maintaining satellite cells in the proliferative state at the start of fetal myogenesis. A possible explanation as to why they do not begin to differentiate in parallel with fetal myoblasts is that they possess a sensitivity to a different set of inhibitors of differentiation than fetal myoblasts. Platelet derived growth factor (PDGF) was found to function in this manner during muscle development in chick [Yablonka-Reuveni and Seifert, 1993]. Once differentiated, satellite cells contribute significantly to the future adult myofiber population until the late post natal period when skeletal muscle development is complete[Zhang et al., 1998]. Once in adult muscle, satellite cells are quiescent. Quiescent satellite

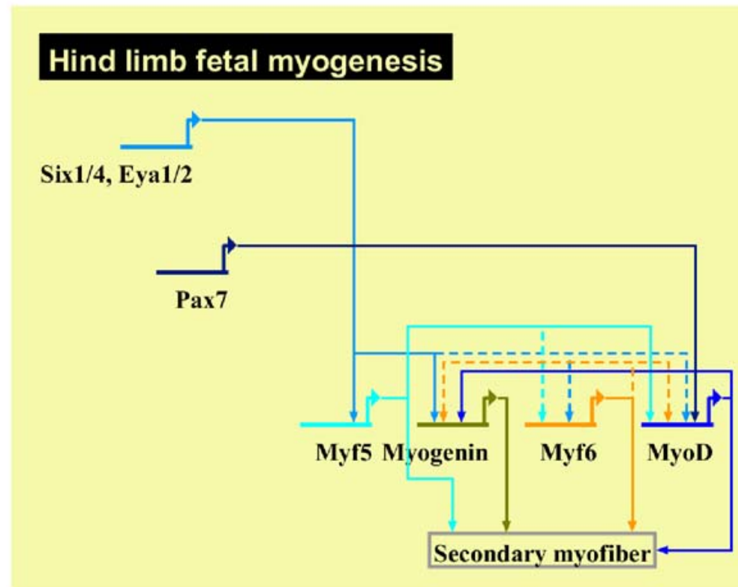

Figure 1.4: Fetal Myogenesis. Pax7 now guides the myogenic program towards secondary myofiber formation via myoD. Myf5 can also activate the myogenic program indepently of Pax7. The Six/Eya transcriptional complex directly regulates Myf5. This is also the regulatory circuitry used by satellite cells once they activate the myogenic program midway through the fetal development stage. Solid lines indicate a direct *cis* interaction and dotted lines indicate an indirect interaction.

cells remain *pax7* positive, but no longer express *myoD*. Interestingly, *pax7* loss of function studies have demonstrated that Pax7 is only required to activate the myogenic program via *myoD* up to post natal day 21.0. After day 21.0, satellite cells are still able to activate the myogenic program in the absence of *pax7* [Lepper et al., 2009]. The regulatory molecules responsible for activating the myogenic program of satellite cells after 21.0 days of post natal development have not been identified, but possible candidates include the Six family and, depending on the stimulus for regeneration of muscle, possibly direct-acting cytokines or hormones that up-regulate *myf5*, *myf6*, or *myoD*. By this time in skeletal muscle development, the regulatory circuitry has elegantly directed waves of proliferation and differentiation programs in embryonic, fetal, and satellite progenitor cells, respectively. Now that the skeletal muscle bundle is structurally and metabolically enabled, and its future regenerative capacity, via satellite cells, secured; innervation of the muscle bundle provides the major regulatory input to further advance the functional capacity of mammalian skeletal muscle.

1.2.3 Regulatory circuitry of adult hind limb muscle

Peri and post natal development is characterized by a down-regulation of the myogenic program and establishment of fiber-type specificity. The latter is driven by innervating slow and fast firing motor neurons. The MRFs, with the exception of *myoD*, are down-regulated, since further differentiation is what is called for at this stage and not further proliferation of progenitor cells. Msy3 and its co-binding factor, Pbx, along with Daschund 2 (Dach2) directly work to repress *myogenin* in innervated skeletal muscle [Berghella et al., 2008, Tang and Goldman, 2006]. *myoD* is expressed at a low level in innervated muscle, regulated in part by the Six/Eya complex [Laclef et al., 2003] (Figure 1.5). The slow and fast firing motor neurons that innervate the muscle fibers during this developmental phase, assist in stimulating the replacement of embryonic and fetal isoforms of the MyHC contractile proteins with the adult isoforms. In rodents, the major MyHC fiber-types of fast-twitch muscle are: IIA, IIX/D, and IIB

[Zhang et al., 1998, Schiaffino and Reggiani, 1994]. TypeIIA is the slowest fiber-type in fast-twitch muscle and IIB is the fastest. They are characterized by oxidative and glycolytic metabolism, respectively. Metabolic enzymes associated with glycolytic metabolism include phospho-fructo kinase (Pfk) and muscle specific enolase (Mse) (Figure 1.5). These enzymes are first expressed in secondary myofibers [Barbieri et al., 1990]. The Six/Eya transcriptional complex positively regulates the fast-type transcritional program [Grifone et al., 2004, Richard et al., 2011]. The neuronal input is conveyed to the transcriptional machinery of the muscle via calcium-dependent signaling following excitation-contraction coupling. Hdac4 is a key molecule connecting nerve activity to muscle transcription, as will be discussed in the next section [Cohen et al., 2007]. In an innervated muscle, calcium influx activates calcium-dependent kinases, which phosphorylate Hdac4 and send it out of the nucleus [McKinsey et al., 2002, Liu et al., 2005] (Figure 1.5). In addition to influencing fiber-type, innervation also maintains muscle mass and plays an important role in forming the neuromuscular junction in adult muscle. The latter we will discuss in further detail below and in Chapter 3.

1.2.4 Regulatory circuitry of denervation response in adult fast-twitch hind limb muscle

Upon denervation of fast-twitch muscles (i.e. MyHCIIB dominant) of the hind limb, calcium influx into the muscle fibers decreases and Hdac4 gets dephosphorylated and shuttled to the nucleus. Once in the nucleus, Hdac4 and its co-factor, Mef2, helps to set in motion a series of regulatory events that characterize the first 7.0 days following denervation, which I will refer to henceforth as the *early denervation* response (Figure 1.6). Hdac4/Mef2 functions to repress *pfk*, *mse*, and *myhcIIB* transcripts, which in turn represses the glycolytic and fast-twitch characteristics of the muscle bundle, respectively. This results in the muscle assuming the metabolic characteristic of the remaining fiber-types, such as MyHCIIA in Figure 1.6. This is indeed indicative of

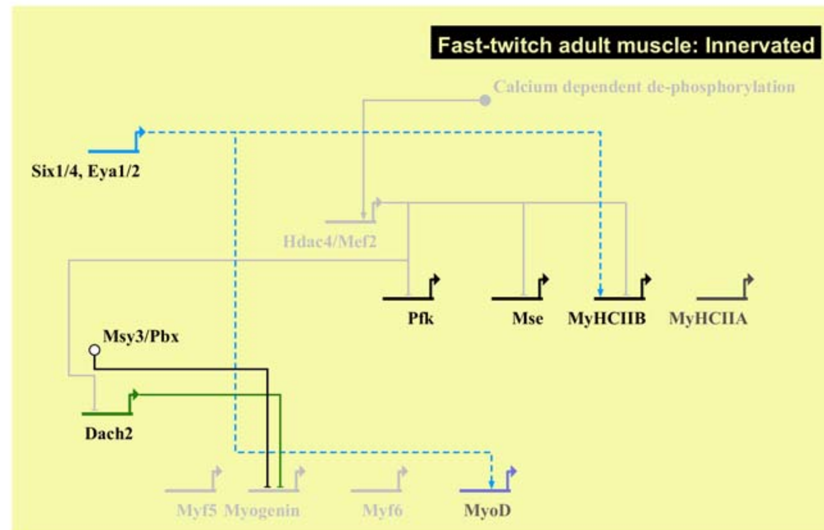

Figure 1.5: Adult fast-twitch innervated muscle is characterized by inactivity of the myogenic program and robust expression of *myhcIIB* (i.e. fast-twitch fiber-type) and *pfk* and *mse* (i.e. glycolytic enzymes). This regulatory state is influenced strongly by nerve input. All gene names in bold text represent actively expressed genes. The Six1/4,Eya1/2 transcriptional complex contributes to the robust expression level of *myhcIIB* and also to low-level *myoD* expression. Dach2 and Msy3/Pbx contribute to down-regulation of *myogenin* in innervated muscle. *myhcIIA* and *myoD* are colored in an intermediate shade to indicate low to intermediate levels of expression for these two genes in innervated muscle. The calcium influx in innervated muscle fibers maintains Hdac4 phosphorylated and in the cytoplasm (i.e. not able to regulate its denervation-associated target genes). Solid lines indicate a direct *cis* interaction and dotted lines indicate an indirect interaction. Satellite cells are quiescent during this state and only become activated following a trauma or other degenerative stimulus that requires muscle regeneration.

a fast-twitch transcriptional program driven by fast-firing motor neurons. *runx1* is activated following denervation and functions to sustain transcription of *mychIIA*, an aforementioned slow-type contractile protein; keratin-type1 cytoskeletal 18 (*krt1-18*), which links sarcomeric Z-lines with M-lines; and acetylcholine receptor subunit gamma (*chrng*), which is a subunit previously only expressed in pre-innervated late fetal and post natal muscle [Wang et al., 2005, Ursitti et al., 2004]. By acting in this manner, *runx1* functions to maintain muscle fiber structural integrity and possible re-innervation capacity in the fibers that remain. Also, as can be appreciated in Figure 1.6, by repressing *dach2*, Hdac4/Mef2 de-represses *myogenin* and indirectly activates all of the target genes of *myogenin* [Tang et al., 2009]. These target genes include *myoD*, which inturn activates itself and *myogenin* [Berkes et al., 2004]; *fbxo32* and *trim63*, which are both E3-ubiquitin ligases that are responsible for the protein degradation events that characterize neurogenic atrophy [Bodine et al., 2001, Moresi et al., 2010]; the acetylcholine subunit receptor genes alpha (*chrna*), gamma (*chrng*), and delta (*chrnd*) [Burden, 1977b]; and muscle specific kinase (*musk*). These four genes at the bottom right corner of Figure 1.6 represent components of the regulatory subcircuit of the developing neuromuscular junction, last activated during peri and post natal development [Burden, 1977b, Mazhar and Herbst, 2012].

Comparison of the different regulatory networks governing myogenesis, not only at different anatomical regions, but during the different phases of skeletal muscle development in cells of the same region, indicates an elegant regulatory design that to date is not well understood. A dedicated approach to functional testing of *cis*-acting regulatory sequences for key regulatory molecules is needed to interweave the current knowledge pertaining to the many GRNs governing skeletal muscle biology during development and during physiological response in the adult. In the present study we will introduce an stable, efficient, cost effective, short time-to-assay and biologically relevant gene transfer technique for functional testing of *cis*-acting regulatory sequences. We will focus on *myogenin* and its transcriptional down-regulation during the late phase of the neurogenic atrophy response.

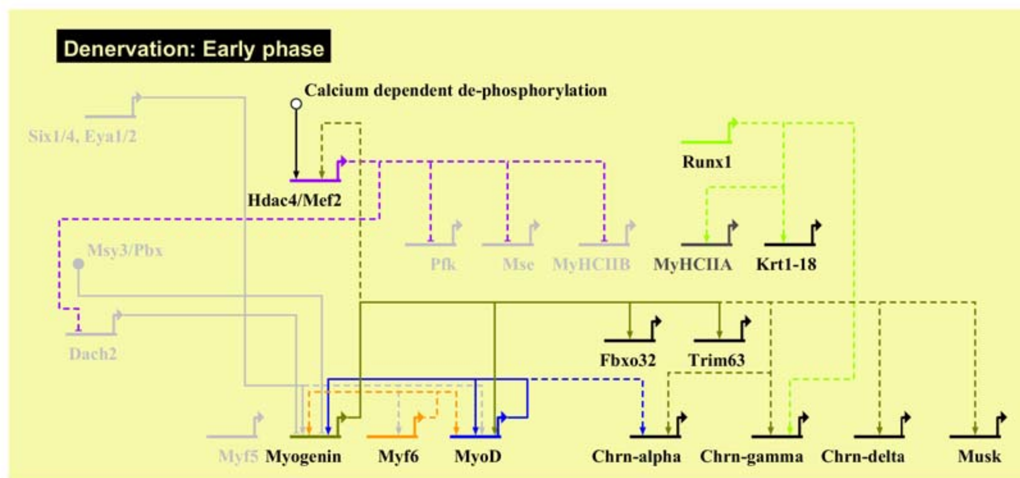

Figure 1.6: Denervation response of adult skeletal muscle out to 3.0 days after denervation (i.e. early phase). A decrease in intracellular calcium results in the dephosphorylation of Hdac4 and it gets shuttled into the nucleus to repress and activate many regulatory subcircuits including metabolic, neuromuscular junction, and neurogenic atrophy programs (see text). Solid lines indicate a direct *cis* interaction and dotted lines indicate an indirect interaction.