Protein Profile and Directed Gene Expression of Developing C2C12 cells

By

Susan Rashid

Submitted in Partial Fulfillment of the Requirements

For the Degree of

Master of Science

In the

Biological Sciences

Program

YOUNGSTOWN STATE UNIVERSITY

August 3, 2015

Protein Profile and Directed Gene Expression of Developing C2C12 cells

Susan Rashid

I hereby release this thesis to the public. I understand that this will be made available from the OhioLINK ETD Center and the Maag Library Circulation Desk for public access. I also authorize the University or other individuals to make copies of this thesis as needed for scholarly research.

Signature:

_____

Susan Rashid, Student Date

Approvals:

_____

Dr. Gary Walker, Thesis Advisor Date

_____

Dr. Jonathan Caguiat, Committee Member Date

_____

Dr. David Asch, Committee Member Date

_____

Dr. Sal Sanders, Associate Dean, Graduate Studies Date

ABSTRACT

Myogenesis is a tightly regulated process resulting in unique structures called myotubes or myofibers, which compose skeletal muscle. Myotubes are multi-nucleated fibers containing a functional unit composed of cytoskeletal proteins called the sarcomere. The specific arrangement of these proteins in the sarcomere works to contract and relax muscles. During embryonic and post-embryonic development, fluctuations in expression of growth factors throughout the program account for the dramatic structural changes from cell to mature muscle fiber. *In vivo*, these growth factors are strictly spatiotemporally regulated according to a 'myogenic program.' In order to assess the dynamics of protein expression throughout this program, we conducted a time course study using the mouse myoblast cell line C2C12, in which cells were allowed to differentiate and insoluble protein fractions were collected at seven time points. The insoluble protein fraction accounts for cytoskeletal proteins, involved in the dramatic remodeling of cells. In order to discover influence of growth factors on protein expression, a similar time course was conducted with the addition of growth factors (serum) at specific time points after induction of myogenesis. The cells were allowed to differentiate through the time course and the detergent- insoluble protein fractions were collected. Proteins were separated according to their eletrophoretic mobility using SDS-PAGE and were identified using LC-MS/MS. Proteins were then functionally annotated using Gene Ontology (GO) terms. Analysis of protein expression during *in vitro* myogenesis shows progression of the myogenic program. Furthermore, a targeted gene expression study was conducted using reverse-transcriptase quantitative PCR. Expression of the sarcomeric proteins actin, myosin, and titin also fluctuated throughout the time course, indicative of a developmental program.

ACKNOWLEDGEMENTS

First and fore-most, I would like to thank my adviser Dr. Gary Walker for his unbridled support throughout my time at Youngstown State University (YSU). Not only was he willing to afford me any opportunity he could academically and in the lab, but always simplified a perceived problem and remained optimistic. More importantly, his patience remained unwavering. As the chair of the biology department, an instructor of classes, and the adviser of graduate students and undergraduate students in a lab, this is no easy feat.

I also have a debt of gratitude to my committee members Dr. Jon Caguiat and Dr. David Asch who were always around to lend a helping hand. They were the instructors of my first classes at YSU and proved continuously constructive both in class and in the lab. In fact, all members of the Proteomics/Genomics Research group were generous with their time and together created an ideal environment for friendly lab work. A special thanks extended to Ed Budde, our resident molecular biologist, for his patience and extra efforts in aiding in molecular technique, which is no simple task.

Without funding, there cannot be science; therefore I would specifically like to acknowledge the Department of Graduate Studies for affording graduate programs, like our, funds for materials, travel, and student support. On the same note, I would like to thank Dr. Chester Cooper, director of the Center for Applied Chemical Biology, for awarding me two consecutive research assistantships in a time when focus solely on lab work and class was critical to me.

Lastly, but not least I would like to honorably mention my peers in this graduate program. They have been helpful, understanding, and encouraging. Their camaraderie was the little-something-extra I needed to get through the program.

Table of contents

Table of Figures

List of Tables

1.INTRODUCTION

1.1 Introduction

Movement is essential for all life; it allows us to obtain necessary resources and sustain vital biological functions, such as breathing. At the basis of the ability to move we find skeletal muscles, which are composed of a series of complex structures working in concert to provide both voluntary and involuntary contraction. The intricacy of muscle tissue indicates an extensive developmental program.

Muscle cells, also known as myocytes, are highly specialized and multinucleated due to the fusion of precursor cells known as myoblasts. These cells appear striated due to an arrangement of intracellular structures, known as myofibrils. Myofibrils are composed of thick and thin filaments, with each filament type containing different proteins serving their respective functions during muscle contraction. These fibers are arranged in a highly unique structure known as the sarcomere, the smallest unit of contraction.The thin filaments, which compose the Z-line of the sarcomere, are predominately actin protein. The thick filament, which forms the M-line of the sarcomere, is composed of the motor protein myosin. Through a cascade of molecular binding and conformational changes, the thousands of tandem sarcomeres found in muscle, shift together to contract the tissue to produce movement. Though muscle function has been well studied, the molecular mechanisms behind developing this elaborate and diverse tissue have not all been elucidated.

1.2. Myogenesis

Myogenesis, the process of muscle development, as most developmental pathways, is complex, tightly regulated, and continues throughout an organism’s lifetime. Myogenesis occurs in multiple and overlapping phases initiated by spatiotemporal cues from cascading transcription factors. The major events occurring

throughout the process include specification of cell lineage, proliferation, migration, and differentiation. Even differentiation is further organized into a specified time line depending on the type of progenitor cell entering this phase (Rossi,Messina 2014). As is well referenced, skeletal muscle is a heterogenous tissue with several different fiber types to serve its wide array of functions. Terminal differentiation and protein composition of these fibers differ, accounting for increased complexity in the signaling transduction pathways (Meressi et al. 2007)

1.2.1 Molecular Embryonic Myogenesis

Myogenesis is a conserved process with studies comparing several vertebrate model organisms: zebrafish, mice, rats and chickens. In fact, the basic helix-loop-helix (bHLH) motif was first characterized in muscle specific regulatory factors (MRF's). The first protein in the family to be discovered was the MyoD transcription factor. In 1987 Davis et al. transfected cDNA from mouse myoblasts into fibroblasts, and they proceeded to fuse into myotubes (Davis et al. 1987). It was hypothesized this cDNA activated loci in the genomic DNA that could establish a myogenic lineage regardless of original cell program (Choi et al.990). They were able to identify a characteristic DNA binding domain that would go on to identify other MRF's, such as Myf5, Myogenin (Myog), and Mrf4. MyoD and Myf5 are both critical in skeletal myoblast lineage specification and are activated by the morphagens sonic hedgehog (Shh) and wingless1 and 3 (wnt1 and wnt3), and the paired-homeobox 3 transcription factor (Pax 3). Another paired-homeobox transcription factor, Pax7 is found upregulated with myogenin in terminal differentiation, although the exact mechanism of Pax7 function is unknown (Parker et al. 2007). MRFs and the Pax transcription family work with a multitude of other proteins in well-known signaling pathways to push embryonic myoblasts out of the

cell-cycle and into the proper localization and mature myocyte. (Braun, Gautel, 2011)

1.2.2. Myogenesis in C2C12 cells

C2C12 cells are a mouse embryonic myoblast cell line that have been used extensively by researchers to investigate properties of muscle development. Although the myogenic lineage has been established in these cells studies have shown that the major transcription factors from early myogenesis can be traced to some extent (Diel et al 2008, Rios et al 2001). Researchers can then take a closer look at secondary myogenesis and terminal differentiation.

Traditionally cells are obtained commercially and proliferated in culture flasks using a nutritious medium supplemented with a high percentage of a growth serum, such as fetal bovine serum (FBS). Myoblasts are then triggered to differentiate by reducing the percentage of growth serum, thereby reducing growth factors and mitogens, after which they begin to fuse and elongate. By approximately 6 days post-differentiation, mature myofibers are formed in culture and by day 12, fully function muscle cells are present (Kislinger et al. 2005). This simple, yet effective methodology has allowed researchers to externally alter the myogenic program by various technique to investigate specific aspects of the program (Blau, Epstein, 1979).

1.3 Sarcomeric Proteins

Myogenic cell populations, *in vivo* and *in vitro*, undergo extensive structural changes throughout their development to mature muscle fibers. Progenitor cells fuse together and subsequently go through cytoskeletal remodeling. Next cells adhere to the extracellular matrix and elongate to form the striated pattern characteristic of muscle tissue (Rui et al. 2010). Although several proteins are involved in making the final structure, we are most interested in the main proteins involved in the basic

contractile unit, the sarcomere. Interestingly, much research has been done on the anatomical and molecular processes in myogenesis; however the assembly of the sarcomere is still poorly understood (Rui et al. 2010). The mechanism of function of the sarcomere is well known; however the foundation of some myopathies still elude us and are continually being investigated. Perhaps, a foundation in sarcomereogenesis may provide crucial details.

1.3.1 Titin

The largest protein known to man, so far, is titin, also known as connectin, and is the third main protein involved in muscle contraction after actin and myosin. Titin is coded for by the single gene TTN; however, this gene contains a total of 363 exons and is known to undergo alternative splicing to produce several isoforms of titin expressed variably in tissues (Guo et al 2009). One isoform known as N2A was found to always be expressed in skeletal muscle. This multidomain protein is the third most abundant in the sarcomere and serves several functions.

Titin spans the half the entire sarcomere the N-terminus ending at the Z-line and the C-terminus at the M-line. The length of the protein is included in two prominent sarcomeric regions known as the I- band, describing thin filaments not overlapped by thick filaments, and the A band, which spans the thick filaments (Tskhovrebova, Trinick, 2004). It is well documented that the structure of titin changes drastically at the A/I band junction. The predominant domains found in the I band regions are tandem immunoglobulin (Ig) domains connect by a domain dominated by proline, glutamate, valine, lysine (PEVK) motif. This region is thought to give elasticity to the sarcomere, as the bound thin filaments contract. For example, the stiff PEVK motif is modulated by calcium ions, crucial for muscle contraction. Continuing through the A-band, fibronectin-3 and Ig domains repeated corresponding to the thick

filaments ending with a kinase domain at the M-line (Tskhovrebova, Trinick, 2004). The corresponding domains allow titin to bind myosin in the thick filament stabilizing the structure. However, these domains are conserved and allow for the interaction of titin with other proteins (Eilersten et al 1994).

Titin has been found to play a role in myopathies and dystrophies. For example, Walker et al initially worked on mysethia gravis (MG), a nueromuscular autoimmune disorder resulting in weakening muscles. They found a subset of patients presented with Autoimmune Rippling Muscle disease (ARMD), exibiting involuntary muscle spasms. Walker et al eventually found that these patients possessed antibodies against titin N2A (Watkins et al. 2006). These autoantibodies were found against the region in which titin and myosin overlap and near the PEVK motif, possibly accounting for altered sarcomere contraction and phenotypic symptoms.

1.3.2 Myosin

Although the actin, the most abundant of the sarcomeric proteins, gives rise to the structure, myosin is the main component in driving contraction. Known as myosin heavy chains (MYH), myosin is a motor protein involved in several cellular processes such as cytokinesis, vesicular transport, and motility (Weiss et al 1999). Those are most concerned with are Myosin II or “conventional” mysoins forming the thick filaments of the sarcomere. These proteins have an ATPase region from which they utilize the energy generated from the hydrolysis of ATP to contract the sarcomere.

There are five isoforms of myosin that are differentially expressed based on developmental time and muscle fiber type. These include MYH 3, an embyronic isoform, MYH8, a perinatal isoform, MYH 2, MYH 1, MYH4, adult isoforms, and lastly MYH7, a slow isoform. Genes encoding these isoforms are found in clusters

along chromosomes 11 in mice and conserved along chromosome 17 in humans (Weiss et al 1999). The isoforms expressed post-natally give rise to the variability of contractile functions and diversity in muscle fibers. Each isoform is expressed according to different external stimuli, such as physical activity (Pandorf et. al 2006, Biressi et al 2007). A shift can occur as muscles adapt to either slower or faster contractile properties. For example a shift in expression from MYH2 to MYH1 to MYH4, fast-twitch myosin heavy chain, can occur with increased muscle use. The transcriptional regulation of this shift has not been elucidated.

1.4 Proteomics

Proteomics was defined by Wu et al. as a "collective study of all measured proteins of cells in a given condition" (Wu et al. 2002). The study of "-omics" and other high through-put data acquisition techniques in recent years has been particularly valuable to developmental studies, such as the exploration of muscle development. As illustrated previously, the process is a concert of genes, proteins and small molecules with most serving several functions. Proteomics and other high-throughput gene expression data can give us a snapshot in time of processes taking place within a living animal or cell culture (Wu et al. 2002). Pertaining to proteomics specifically, the presence and abundance of an identified protein is indicative of gene expression patterns taking place within the cells. By layering proteomic data from several time points, a clear picture of events is illustrated throughout the development program.

Furthermore with the rise of big data sets comes the development of databases and tools for structural, evolutionary, and functional annotation and visualization of genes and gene products. In one example, Kislinger et al uses the Gene Ontology Consortium classification system to associate their identified protein products

collected from their C2C12 cell lysate with biologically relevant terms (Ashburner et al. 2000). The functional clustering of hundreds of identified proteins illustrated the enrichment of mitotic, structural, and metabolic proteins at expected time points throughout myofiber development (Kislinger et al 2011)\.

1.5 Aims

In this study we aim to track the fundamental events of myogenesis using C2C12 myoblasts over time with a focus on cytoskeletal proteins and their role in remodeling to form these unique cells. We take both a global approach in using proteomics to define particular events and a targeted approach in using reverse-transcriptase quantitative PCR (RT-q-PCR) against the prominent sarcomeric proteins actin, titin, and myosin. Perhaps this may provide insight to sarcomeric assembly.

These approaches will give us a comprehensive look inside myogenic cellular development, as other studies have previously. However, we also attempt to define the role of growth stimulating factors from growth serum, such as FBS, used during *in vitro* studies. We have added back high growth serum media at specific time points after cells have been "pushed out" of proliferation using low growth serum, as is traditional. We will look at this effect through proteomics and RT-q-PCR and perhaps see a resetting of the myogenic program.

1.6 Hypothesis

Although myogenesis is tightly regulated, growth serum contains mitogens which can activate several signaling pathways. Our approach to "add-back" these mitogens, along with other growth factors which can affect development, has proved novel. Collective literature review has defined time points at which prominent stages occur during the developmental program. Utilizing this knowledge, we plan to

investigate the role of these external signaling factors by making them available to the cells once more. We hypothesize that the addition of the media supplemented with a high percentage of FBS will reset the myogenic program, and cause at least a subset of cells in culture to revert to a proliferative state.

2. METHODS

2.1 Cell Culture

C2C12 mouse myoblast cell line bought from ATCC (Manassas, VA) was used in all proceeding experiments. Cells were cultured using traditional adhesion culture methodology. 1 mL vial of cells was defrosted in a 37ºC water bath, and gradually mixed with Dulbeco's Modified Eagle Media (DMEM) supplemented with 20% Fetal Bovine Serum (FBS) and 1% Ampicillin (High Serum Growth Media). Cells were grown in an incubator at 37ºC with 5% CO_2 until an observed confluency of 80% to 90%. Cells were treated with 1x 0.05% Trypsin-0.53 mM EDTA (Corning Life Sciences, Manassas, VA) pelleted in a 15 mL conical tube and reseeded at 1mL in two separate culture flasks. The cells were once again grown up to observed confluency of 80-90%.

Cells from each flask were pelleted again and reseeded into triplicates of 25cm^3 flasks representing time points designated: T= 0 Days, T= 2 days, T= 4 days, T= 6 days, T= 8 days and T= 12 days. Cells were proliferated in each flask to 80-90% observed confluency with high growth serum then triggered into differentiation with DMEM supplemented with only 1% FBS and 1% ampicillin denoted as low growth serum media. Cells from triplicate flasks designated T= 0 were collected, pelleted and stored in a cryogenic solution. Remaining cells were fed low growth serum media until the time points designated on each set of flasks after which cells were fed high growth serum media. These cells were all grown and harvested after 13 days, from the day differentiation was triggered. This time course was carried out twice; but during the second time course, T= 0 days was replaced with triplicate control flasks. These control flasks were grown to and harvested at 13 days, but no high growth serum was added back (See Appendix 3 for time course calendars).

2.2 Protein Collection

Cell lysate was collected according to protein solubility in detergent to produce soluble and insoluble fractions. All flasks containing cells from the first time course were collected separately at Day 13, trypsinzed then treated with 1 mL isotonic wash buffer followed by 1 mL buffer containing triton-x-100. Buffer now containing detergent- soluble protein fraction was poured out of the culture flask into microfuge tubes and stored at -80ºC. Remaining lysate, in each flask was treated with 1 mL of 1X SDS buffer and rocked gently on ice for 5 min or until DNA degraded completely. This detergent- insoluble protein fraction was collected in microfuge tubes and stored at -80ºC.

2.3 SDS-PAGE and Gel Analysis

Our interest centered around cytoskeletal protein investigation, therefore only proteins in the detergent-insoluble fraction were used in downstream applications Samples of extracted protein treated with 1x SDS were thawed and 10µl from all triplicates in each time point, as well as bovine serum albumin (BSA) standard, were loaded into two precast 12% polyacrylamide gels, as (Bio-Rad, Hercules, CA) and ran with 1X tris-glycinde-SDS (TGS) running buffer. The powers source was set a constant 40 A/gel and ran for an average of 30 minutes. This process was repeated several times with varying loading volumes to find optimum volume.

Gels were then visualized using Coomassie Brilliant Blue staining. Gels were shaken over night with filtered Coomassie stain (Fisher Scientific, Pittsburgh, PA) and a protein fixing buffer. Gels were then treated with buffers denoted as high and low destain. These were used alternately until gel became translucent and proteins bands were visible. Gel image was then captured and scanned onto the computer using the Pharos FX system (Bio-Rad, Hercules, CA), and analyzed using the software

Quantity One (Bio-Rad, Hercules, CA) according to the quick start guide included in the program.

2.4. LC- MS/MS and Proteomic Analysis

Protein identification was necessary to describe specific events in the *in vitro* myogenic program, and to specify the presence of cytoskeletal proteins of interest. Triplicate samples from each time point were pooled together and 12 µl of pooled sample from each time point were loaded in a precast 12% polyacrylamide gel (Bio-Rad, Hercules, CA) and run as described previously. Triplicates were pooled in order to increase abundance and diversity of proteins at each time point for higher resolution detection. The gel was stained with Coomassie Brilliant Blue and proteins were fixed in the gel as described above.

Eight gel bands were selected (Fig. 2) to be cut out to be sent to Ohio State University Mass Spectrometry and Proteomics Facility for identification. Four band regions at molecular weights of: >212 kDa., ~212 kDa, ~45 kDa and ~18 kDa at days 0 and 6 were chosen and sent. Proteins were partially sequenced using capillary-liquid chromatography tandem mass spectroscopy (LC-MS/MS). The sofware MASCOT (Matrix Science, London, UK) was used for sequence identification and protein abundance. UniProtKB/Swiss-Prot protein database ID's, emPAI abundance values, scores and partial alignments were sent back. Further proteomic analysis was done using various online programs such as Database for Annotation, Visualization and Integration Discover (DAVID), Gene Ontology Consortium, and MetaboAnalyst to functionally annotate and visualize data.

2.5. RNA extraction, RT-qPCR, and Statistical Analysis

Cells from the second time course were collected at day 13; each flask was separately trypzinized, pelleted and stored at -80ºC. At a later time, total RNA was

extracted from each cell pellet using RNeasy Mini Kit (Qaigen, Valencia, CA) according to kit protocol. RNA concentrationand quality were determined using a Nanodrop Spectrophotometer (ThermoScientific, Waltham, MA).

Reverse transcriptase quantitative PCR (RT-qPCR) was then carried out using the iTaq$_{TM}$ Universal SYBR® Green One-Step kit (Bio-Rad, Hercules, CA). Each reaction in a 96-well plate contained: 0.2 µl reverse transcriptase, 2 µl F primer (10 µM), 2 µl R primer (10 µM), 3.75 µl nuclease-free water, 10 µl Sybr green, 2 µl RNA (150 µg). Primers were designed based on previous experimentation in this lab; but MYH4 and MYH8 were based on Zhou et al.'s work (See Table 1 below) (Zhou et al 2010). Reactions took place on the iCylcer iQ$_{tm}$ Real-Time PCR detection system (Bio-Rad, Hercules, CA) and accompanying software. Finally, statistical analysis on amplification data was carried out using DART-PCR excel-base analysis (Peirson et al. 2003).

Table 1. Primer sequences

Gene	Forward	Reverse
Alpha, Actin-1	5' AGGACCTGTATGCCAACAACGTCA 3'	5' TCAACATCTGCTGGAAGGTGGACA 3'
Titin N2A	5 'CCTGAGGACCTGGAAGTCAA 3'	5' AGTACCAATGAGGCGGCTTT 3'
MYH4	5' CACCTGGACGATGCTCTCGAG 3'	5' GCTCTTGCTCGGCCACTCT 3'
MYH8	5' CAGGGAGCAGGAATGATGCTCTGAG 3'	5' AGTTCCTCAAACTTTCAGCAGCCAA 3'

3. RESULTS

3.1.Introduction

Firstly, the cytoskeletal proteome was analyzed using SDS-polyacrylamide gel eletrophoresis (SDS-PAGE) then LS/MS-MS identification. Both time point and molecular weight were criteria to find proteins of interest; however, a complete profile of the dynamics of the proteome was also illustrated using functional annotation. These results were compared to both the work of Kislinger et al. who used similar approaches. The addition of a high growth serum media, such used in proliferation, post initiation of terminal differentiation provided a novel aspect to our proceedings.

Inspired both by previous work in this lab and our proteomic findings, a directed gene study was conducted using reverse transcriptase-quantitative polymerase chain reaction (RT-qPCR). Most importantly, genes targeted are essential to structure and function of skeletal muscle, and can be used as developmental temporal cues.

3.2. Separation of Proteins through SDS-PAGE

Upon collection of cell lysate from each of the twenty-one cultures, two separate types of samples were collected: a detergent-insoluble protein sample and detergent-soluble protein sample. Each sample fraction denotes the solubility of proteins in the detergent, triton-x-100. Our interest is primarily in cytoskeletal proteins which are predominantly found in the detergent-insoluble fraction; therefore our analyses only concerned the samples collected from this fraction.

Proteins in the detergent-insoluble lysate for each culture were separated through SDS-PAGE. Firstly, biological triplicates at each time point were run next to one another to observe consistent banding patterns for each time point. Upon confirming consistency of samples in a single time point, a single SDS-PAGE was conducted

containing all time points for a reliable time course comparison. Banding patterns between lanes, indicating time points after high serum was added, show various shifts in protein density. Prominent bands appear within each lane and at the same electrophoretic distance between lanes. For example, a prominent band can be observed between 66.2- 45.0 kDa in all lanes. However the pronounced prominence of the band at Day 6 is a result of contamination of a BSA quantification standard. The bands also appear in the well region of each lane indicating the presence of proteins too large to migrate through 12% polyacrylamide. Gel was analyzed using QuantOne software (Bio-Rad, Hercules, CA) which showed overall varied protein abundance at all molecular weights over time.

A.

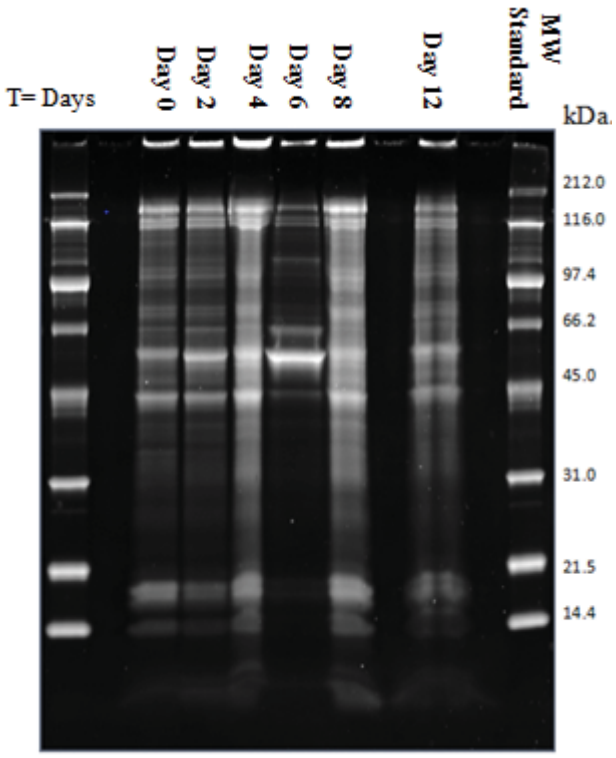
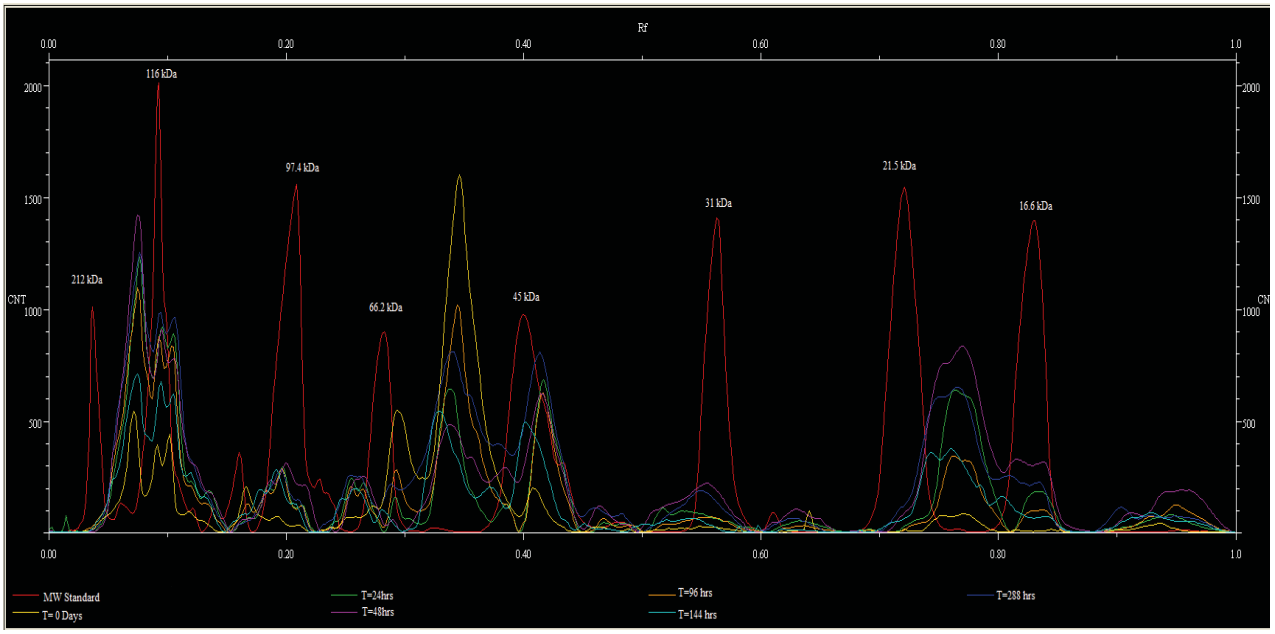


Fig 1. **Protein profile of detergent insoluble cell lysate**.
(A) Gel image from 12% SDS-PAGE of detergent- insouble protein fraction. Lane represent time points after add-back of high-serum; molecular weight standards present on both sides. Band densities vary across time points. (B) Protein abundance across migration distance. X-axis represent protein counts, y-axis represent Rf values. Each color represents a lane, representing a time point after add-back of high serum. High abundance found at >116 kDa and at 45 kDa.

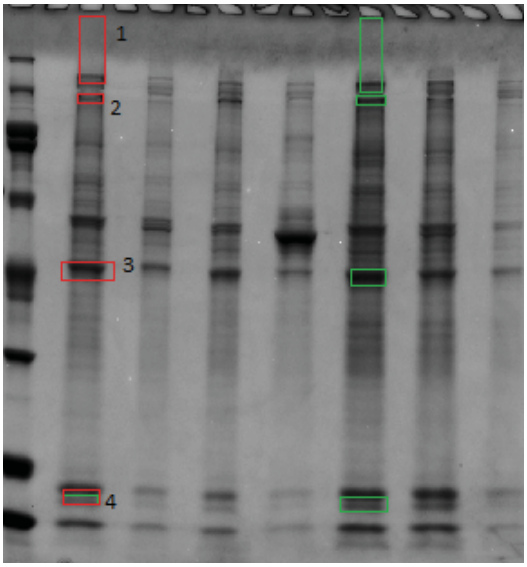
B.



3.3 Proteomic analysis

Further investigation of the proteome required identification of proteins present in the detergent-insoluble fraction of the first set of cultures. Band regions were selected from T= 0 days and T= 6 days across molecular weights (Table 2). Band regions at >212 kDa and ~212 kDa were selected with the intent of identifying titin and myosin. Band regions at ~45 kDa and ~18 kDa were chosen with the intent of identifying actin isoforms. Bands were cut from an electrophoretic gel and sent to Ohio State University Mass Spectrometry and Proteomics Facility for LC/MS-MS processing, and preliminary bioinformatic analysis. T= 0 days represents high growth serum add-back 24 hours after differentiation was triggered; and T= 6 days represents high growth serum add-back six days after differentiation was triggered. Although cells were grown to thirteen days, the most dramatic observed morphological changes occur through ~ Day 6 when elongated myotubes are present in the culture (Kislinger et al. 2005). Although, high serum was added back, we still observed these distinct morphological changes of cells in culture.

Using the MASCOT protein search engine (Matrix Science, London, UK), Ohio State University Mass Spectrometry and Proteomics Facility identified a total of 80 different proteins from Day 0 and 130 different proteins from Day 6 (A5). All protein band regions from both time points were selected from the same SDS-PAGE procedure (Fig 2) to account for consistency in protein separation.



Sample	Time	Molecular Weight (KDa)
1	0, 6 Days	>212
2	0, 6 Days	~212
3	0, 6 Days	~45
4	0, 6 Days	~18

Table 1. Molecular weights of bands excised and sequenced using LC-MS-MS

Fig 2/Table 2. **Gel Band Region Excision for LC-MS/MS analysis, high growth serum conditions**. SDS-PAGE was performed on insoluble protein fraction from high serum treatment. Portions (indicated in red for Day= 0 and green for Day= 6) of the gel containing proteins of interest were sent for sequencing using LC-MS-MS. SDS-PAGE gel sampling regions are indicated by boxes (Band regions 1-4 for both day 0 and day 6).

3.4. Functional Annotation

High-throughput data sets can be challenging to parse through in a meaningful way. In order to identify specific events in the myogenic program and monitor major changes during these events, identified proteins from each time point were functionally annotated using the online program Database for Annotation, Visualization and Integration Discovery (DAVID). The analysis included both KEGG biological pathway identification and Gene Ontology (GO) (Ashburner et al

2002) term enrichment ($p>0.05$), allowing us to elucidate the global dynamics of the cell culture.

Concerning Day 0, 76 of the 80 identified proteins were mapped onto separate GO annotations specific to mouse with significance ($p>0.05$); in Day 6, 119 of 130 proteins were mapped onto separate GO annotations specific to mouse with significance ($p>0.05$). Each protein was mapped onto at least one or more of the three GO categories: biological process, cellular component, molecular function. DAVID yielded hundreds of GO terms corresponding to events reflecting myogenic development such as cytoskeleton, proteosome, motor activity, and skeletal muscle contraction.

3.4.1. Myogenic Program

Upon induction of differentiation by growth factor withdrawal, a dramatic shift in cellular processes occurs accompanied with an increase in intracellular signaling. At this stage, embryonic myoblasts exit the cell cycle and begin drastic structural remodelling. Among the enriched GO terms included histones, DNA binding, RNA binding, and chromatin indicating an upregulation of transcriptional regulation. For example, core nuclesosome components such as histones 2 and 3 and chromatin remodeling factors such as chromodomain-helicase-DNA binding protein 5 were found throughout the time course. Accompanying these basic transcription regulatory factors is the upregulation of downstream signaling proteins, such as Ras GTPase-activating-like protein and A-kinase anchor protein 2 further indicative of a changing landscape within the cells.

However of the 210 identified proteins, cytoskeletal factors dominated the proteome. GO terms enriched include basement membrane, cytoskeleton, adherens binding all under the umbrella GO term striated muscle tissue development.

Subterms even included sarcomere, actin binding, and microfilament motor activity. Specific proteins include organizational factors laminins, nidogen 2, filamin-A,-B,-C, fibronectin, and various isoforms of actin. In fact, the most abundant proteins were in the data were Actin, cytoplasmic 1 and Actin, alpha skeletal muscle. Adhesion proteins such fibronectin and talin-1, -2. Finally, crucial to cell remodelling are proteases; proteosome subunits and proteases such as 26S protease regulatory subunit and proteasome-associated protein ECM29 homolog are also enriched throughout the terminally differentiating proteome. The comprehensive function of these proteins was mapped onto the KEGG biological pathways denoted as “skeletal muscle contraction” and “tight junctions.”

Table 3. **Transcription regulatory proteins. (**A) Proteins identified in transcriptional regulation. "+" represents presence of protein at Day 0 and Day 6 , and "-" represents absence of protein at Day 0 and Day 6. (B) GO terms associated with these particular proteins. P-values and enrichment scores signify significance of terms. Analysis generated by DAVID.
A

Uniprot ID	Day 0	Day 6	Protein
H2A2B_MOUSE	+	+	Histone H2A type 2-B
H2A1F_MOUSE	+	+	Histone H2A type 1-F
H3C_MOUSE	+	+	Histone H3.3C
H32_MOUSE	-	+	Histone H3.2
H31_MOUSE	-	+	Histone H3.1
HNRPD_MOUSE	+	+	Heterogeneous nuclear ribonucleoprotein
TADBP_MOUSE	+	+	TAR DNA-binding protein 43
DNJA2_MOUSE	-	+	DnaJ homolog subfamily A member 2
CHD5_MOUSE	-	+	Chromodomain-helicase-DNA-binding protein 5
RBP2_MOUSE	-	+	E3 SUMO-protein ligase RanBP2
PRP8_MOUSE	-	+	Pre-mRNA-processing-splicing factor 8
TPR_MOUSE	+	+	Nucleoprotein TPR

B.

Category	Term	RT	Genes	Count	%	P-Value	Fold Enrichment
GOTERM_BP_FAT	chromatin assembly or disassembly	RT		6	54.5	7.4E-9	68.0
GOTERM_BP_FAT	nucleosome assembly	RT		5	45.5	1.6E-7	84.6
GOTERM_BP_FAT	chromatin assembly	RT		5	45.5	1.8E-7	82.4
GOTERM_BP_FAT	protein-DNA complex assembly	RT		5	45.5	1.8E-7	81.3
GOTERM_BP_FAT	nucleosome organization	RT		5	45.5	1.8E-7	81.3
GOTERM_BP_FAT	DNA packaging	RT		5	45.5	5.8E-7	61.2
GOTERM_BP_FAT	chromatin organization	RT		6	54.5	1.5E-6	23.5
GOTERM_BP_FAT	chromosome organization	RT		6	54.5	5.1E-6	18.3
GOTERM_BP_FAT	cellular macromolecular complex assembly	RT		5	45.5	1.2E-5	28.5
GOTERM_BP_FAT	cellular macromolecular complex subunit organization	RT		5	45.5	2.0E-5	25.2
GOTERM_BP_FAT	macromolecular complex assembly	RT		5	45.5	7.0E-5	18.3
GOTERM_BP_FAT	macromolecular complex subunit organization	RT		5	45.5	9.7E-5	16.8
GOTERM_BP_FAT	mRNA metabolic process	RT		3	27.3	2.0E-2	12.3
GOTERM_BP_FAT	protein folding	RT		2	18.2	9.0E-2	19.5

Most proteins were found to overlap at both time points, as expected due to the multi-layered process of myogenesis. However, for myocytes to become functional, an alteration in metabolic processes needs to occur to provide energy for muscle

contraction. There were 73 proteins found to be exclusive to day 6, many of which are associated with mitochondrial metabolism and development of biochemical processes. Oxidative phosphorylation, ATP synthesis coupled proton transport and generation of precursor metabolites and energy were highly enriched according to GO annotations. Key metabolic enzymes such as alpha-enolase, creatine kinase, citrate synthase, and cytochromes composing the electron transport chain are expressed at Day 6. Coupled with this data is the presence of components of calcium ion channels, such as voltage-dependent calcium channel subunit alpha2/d; calcium is a crucial regulator of muscle contraction.

Table 4. **Metabolic Process Proteins.** **(**A) Proteins identified in metabolic processes. “+” represents presence of protein at Day 0 and Day 6 , and “-” represents absence of protein at Day and Day 6. (B) GO terms associated with these particular proteins. P-values and enrichment scores signify significance of terms. Analysis generated by DAVID

A.

Uniprot ID	Day 0	Day 6	Proteins
PPIA_MOUSE	+	+	Peptidyl-prolyl cis-trans isomerase A
QCR2_MOUSE	+	+	Cytochrome b-c1 complex subunit 2, mitochondrial
ATPB_MOUSE	-	+	ATP synthase subunit beta, mitochondrial
QCR1_MOUSE	-	+	Cytochrome b-c1 complex subunit 1, mitochondrial
TFR1_MOUSE	-	+	Transferrin receptor protein 1
NSF1C_MOUSE	-	+	NSFL1 cofactor p47
TXND5_MOUSE	-	+	Thioredoxin domain-containing protein 5
ASSY_MOUSE	+	+	Argininosuccinate synthase
VATL_MOUSE	+	+	V-type proton ATPase 16 kDa proteolipid subunit
ENOA_MOUSE	+	+	Alpha-enolase
ACOT2_MOUSE	-	+	Acyl-coenzyme A thioesterase 2, mitochondrial
NDUAD_MOUSE	-	+	NADH dehydrogenase [ubiquinone] 1 alpha subcomplex subunit 13
SCMC1_MOUSE	-	+	Calcium-binding mitochondrial carrier protein SCaMC-1
OSTC_MOUSE	+	+	Oligosaccharyltransferase complex subunit OSTC
ATPD_MOUSE	-	+	ATP synthase subunit delta, mitochondrial
OLA1_MOUSE	+	+	Obg-like ATPase 1
OST48_MOUSE	-	+	Dolichyl-diphosphooligosaccharide--protein glycosyltransferase
KCRB_MOUSE	+	+	Creatine kinase B-type
SUCB1_MOUSE	-	+	Succinyl-CoA ligase [ADP-forming] subunit beta, mitochondrial
GTR1_MOUSE	-	+	Solute carrier family 2, facilitated glucose transporter member 1
SOAT1_MOUSE	-	+	Sterol O-acyltransferase 1
CA2D1_MOUSE	+	+	Voltage-dependent calcium channel subunit alpha-2/delta-1

B.

Category	Term	RT	Genes	Count	%	P-Value	Fold Enrichment
GOTERM_BP_FAT	generation of precursor metabolites and energy	RT		8	36.4	2.4E-8	21.9
GOTERM_BP_FAT	oxidative phosphorylation	RT		4	18.2	5.2E-5	51.1
GOTERM_BP_FAT	ATP synthesis coupled proton transport	RT		3	13.6	1.1E-3	58.0
GOTERM_BP_FAT	energy coupled proton transport, down electrochemical gradient	RT		3	13.6	1.1E-3	58.0
GOTERM_BP_FAT	ion transmembrane transport	RT		3	13.6	1.4E-3	51.1
GOTERM_BP_FAT	proton transport	RT		3	13.6	2.1E-3	41.3
GOTERM_BP_FAT	hydrogen transport	RT		3	13.6	2.2E-3	40.5
GOTERM_BP_FAT	transmembrane transport	RT		5	22.7	2.7E-3	7.8
GOTERM_BP_FAT	ATP biosynthetic process	RT		3	13.6	5.2E-3	26.2
GOTERM_BP_FAT	ATP metabolic process	RT		3	13.6	6.2E-3	23.8
GOTERM_BP_FAT	ribonucleoside triphosphate biosynthetic process	RT		3	13.6	6.6E-3	23.1
GOTERM_BP_FAT	purine ribonucleoside triphosphate biosynthetic process	RT		3	13.6	6.6E-3	23.1
GOTERM_BP_FAT	purine nucleoside triphosphate biosynthetic process	RT		3	13.6	6.7E-3	22.8
GOTERM_BP_FAT	nucleoside triphosphate biosynthetic process	RT		3	13.6	6.9E-3	22.6
GOTERM_BP_FAT	nitrogen compound biosynthetic process	RT		4	18.2	6.9E-3	9.5
GOTERM_BP_FAT	purine ribonucleoside triphosphate metabolic process	RT		3	13.6	7.7E-3	21.2
GOTERM_BP_FAT	ribonucleoside triphosphate metabolic process	RT		3	13.6	7.9E-3	21.0
GOTERM_BP_FAT	purine nucleoside triphosphate metabolic process	RT		3	13.6	8.5E-3	20.2
GOTERM_BP_FAT	purine ribonucleotide biosynthetic process	RT		3	13.6	8.7E-3	20.1
GOTERM_BP_FAT	ribonucleotide biosynthetic process	RT		3	13.6	9.3E-3	19.3
GOTERM_BP_FAT	electron transport chain	RT		3	13.6	9.5E-3	19.2

3.4.2 Molecular Weight Classification

The nature of our sample collection also included the property of molecular weight as a parameter to be investigated. Additionally, proteins of each band region were analyzed based on abundance. Protein abundance was quantitated using the emPAI protocol yielding values that reflect the number of observed peptides for a protein over the number of observable peptides (Ishihama et al 2005). MetaboAnalyst online program was used to visualize a direct comparison of protein abundance over time.

In band regions 1 and 2 at both 0 days and 6 days, a large number of cytoskeletal proteins were identified. Most of these proteins resided in band regions >212 kDa and ~212 kDa; identification of large, structural proteins was expected. However

there is enrichment of the smaller (~45 kDa) cytoskeletal and extracellular membrane factors in Day 6 relative to Day 0. These include desmin, reticulon-2, troponin, and tubulin. Also, high abundances of various isoforms of actin are enriched at both days. However, there is, in general, a higher abundance of these cytoskeletal factors at Day 6 as the more complex myotubes are arranged. Titin, a protein of particular interest to us, was identified in day 0, but in low abundance and could not be considered in further analyses.

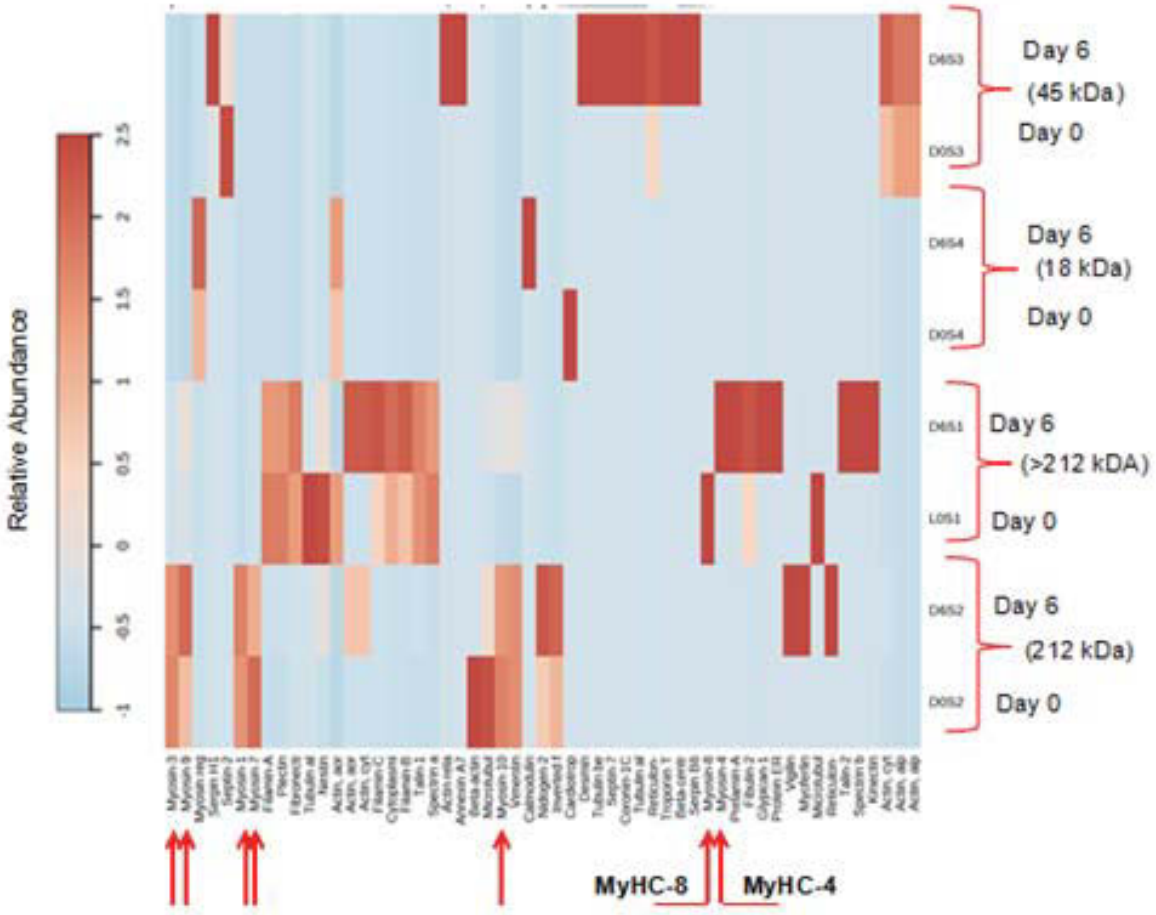


Fig 3. **Relative abundance of cytoskeletal proteins during development under high growth serum conditions**. Arrows indicate myosin heavy chain isoforms, brackets indicate time and molecular weights. Relative abundance based on emPAI values.

One of the more striking results is the temporal protein shift illustrated by Fig. 3

deals with myosin heavy chain isoforms. These isoforms are well characterized and each play a separate role during muscle development and post-natally. Myosin-9 and Myosin-10 are not traditional motor proteins, but aid in cytoskeletal remodeling thus are found in high abundance on both days. Myosin-3, an embryonic isoform is also found enriched on both days, as Day 6 is not yet considered post-natal (Kislinger et al 2011). Myosin-7, the slow isoform, is present at both days, as well, although it should be noted Myosin-7 is also found in cardiac muscle. The drastic shifts can be observed in Myosin-8 and myosin-4. Myosin-8, perinatal isoform, shows high relative abundance at day 0, but seems to disappear at day 6; while myosin-4, a fast-twitch adult isoform, is highly expressed later in development. Also striking, is the presence Myosin-1, adult isoform, at Day 0; although its relative abundance increases at Day comparatively.

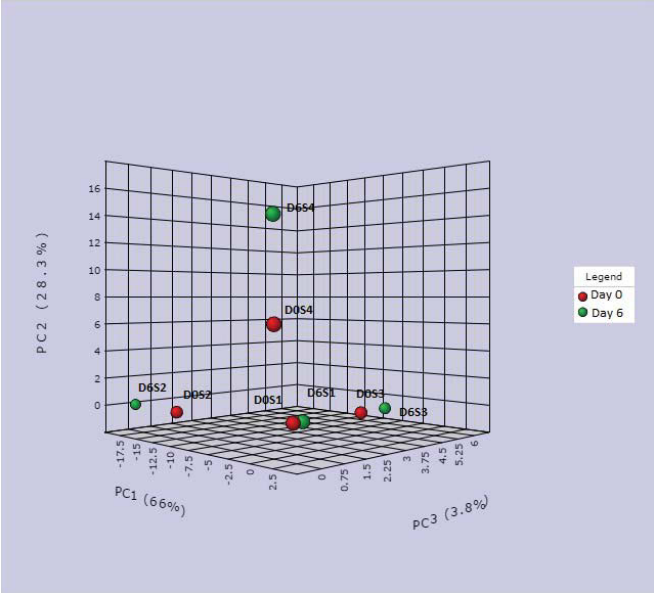


Fig 4. **Principal Coordinate Analysis (PCA) of Proteomic Data.** PC's 1, 2 and 3 describe over 90% variance between samples. Samples mainly cluster by molecular weight, as expected. As molecular weight of the samples decrease, there is more variance between samples.

Although most proteins are present across the time points, an increase in cellular

function resulting in a higher diversity of proteins at day 6. At lower molecular weights, there is more variation between T= 0 and T= 6, especially in regulatory proteins, as the functional annotation above illustrates. These finding are further reflected in Fig. 3 There is less diversity among the larger proteins between time points, but much more diversity as molecular weight decreases. It is speculated the "average protein size" is 53 kDa; although with the diversity of proteins and gene expression, an average molecular weight seemd difficult to pinpoint, proteome shows more dynamics over time, as molecular weight decreases.

3.4.3 Muscle Fiber Heterogeneity

Lastly variations in proteins also indicates the presence of heterogenous muscle fibers, such as is the case with the presence of myocardial associated proteins such as, actin, alpha cardiac muscle, actin, aortic smooth muscle, and cardiotrophin like cytokine factor 1. Although our particular is on skeletal muscle fibers, this further indicates the complex nature of *in vitro* myogenesis and the lack of direct control over terminal differentiation program of the myoblasts.

3.5 Gene expression

As an extra step to confirm findings in our proteomic data, and asses mRNA expression over the course of development, RT-q-PCR was carried out on select genes. Our focus was on striated filament proteins in an attempt to discover more about the poorly understood sarcomere assembly. Genes for skeletal actin (alpha, actin-1), composing the thin filament of the sarcomere, myosin isoforms, composing the thick filaments, and skeletal isoform Titin N2A, spanning both thick and thin filaments for stability and elasticity, were amplified from total RNA of cells after a 13 day time course with addition of high growth serum after initial withdrawal from the cell cycle (Appendix 3).

mRNA transcript levels were measured over a 12 day period at five distinct time points: control, 2 days, 6 days, 8 days, and 12 days. Control mRNA refers to 13 days of cell differentiation without addition of high growth serum. However, high growth factors post-differentiation were added back to at the respective time points to each cell culture, as per the previous time course. All cultures were collected after 13 days. Therefore, "2 days" can be interpreted as: 11 days with high growth serum treatment. "12 days" can be interpreted as: 1 day with high growth serum treatment. Because of the dynamic nature the myogenic program, relative gene expression was measured rather than absolute transcript levels. DART-PCR, an Excel based program, was used to measure relative variation among sample amplification. (Peirson et al. 2003)

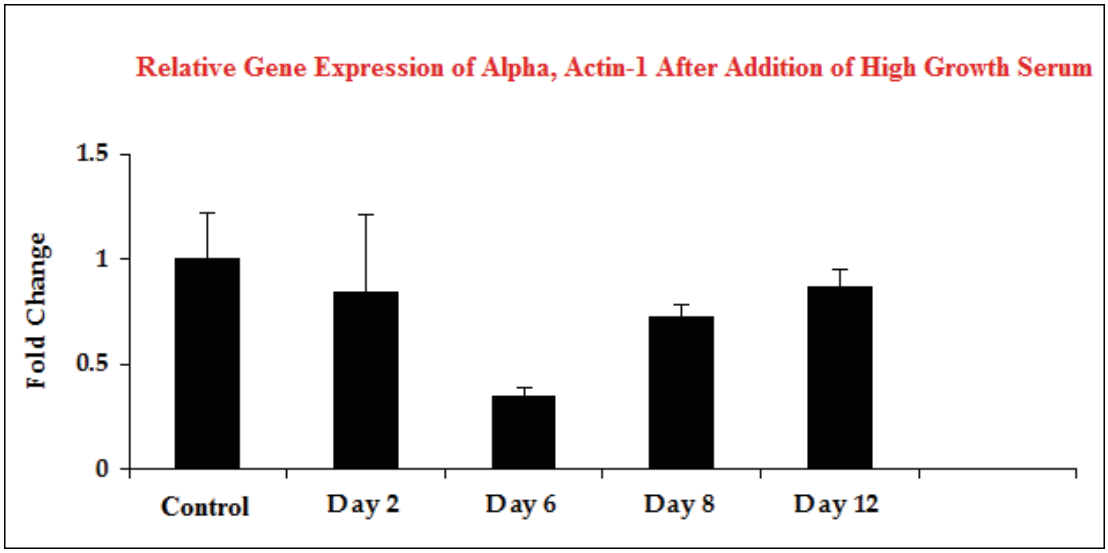


Fig 5. **Relative fold change of Alpha Actin-1 expression.** Measured mRNA transcript levels over time course. Time points indicate day high growth serum was added back

Expression of skeletal actin stayed mostly consistent over time, despite addition of high growth serum. The proteome revealed skeletal actin to be highly abundant throughout terminal differentiation. These gene expression results are congruent with our previous findings. Titin N2A mRNA levels are elevated in

the control time course, sans addition of high growth serum, compared to experimental treatments. Addition of medium containing 20% FBS seems to reduce mRNA transcript levels of skeletal titin, even in day 12 in which culture contains high growth factors for only 1 day before collection. Although not significantly identified in our proteomic data, GO terms annotated to titin include sarcomere, actin-binding, and striated muscle contraction; these same terms are significant across both day 0 and day 6.

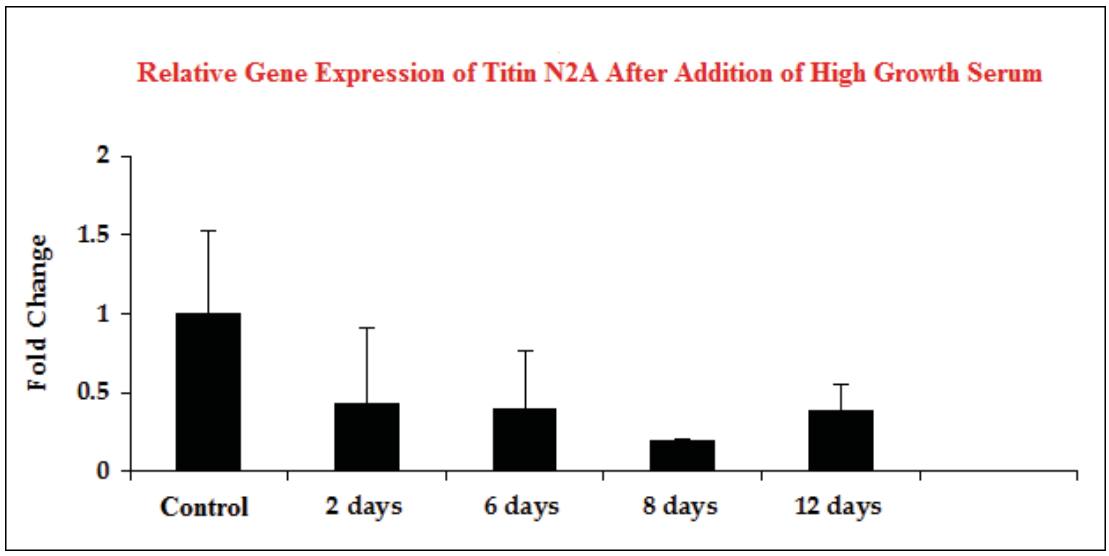
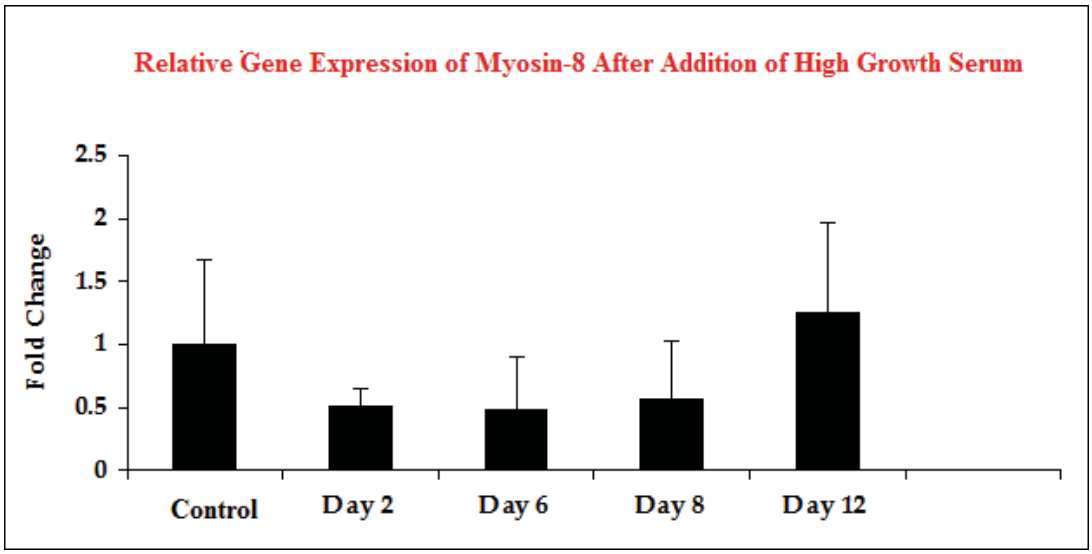


Fig 6. **Relative fold change in Titin N2A expression.** Measured mRNA transcript levels over time course. Time points indicate day high growth serum was added back

The proteome also revealed a striking shift in myosin isoforms throughout development; myosin is also an integral part of the sarcomere, fulling the functional requirements of muscle fibers. Myosin-8 was chosen due to its dramatic upregulation and decline over day 0 and day 6. Conversely, myosin-4 was chosen to for its dramatic upregulation at day 6. Fig. 6 shows relative mRNA transcript levels for each myosin isoform. Like titin, myosin-4 and myosin-8 transcript levels are elevated during control time course relative to culture treated with high growth serum. However in the case of myosin-8, day 12

cells (1 day with high growth factor) show increase of mRNA transcripts; as well as, in myosin-4, day 12 cells show slight elevation of mRNA. Cells in the day 12 cultures have spent the least time in the presence of high growth factor during terminal differentiation. Yet, these results do not reflect our findings from our proteomic data analysis.

A



B

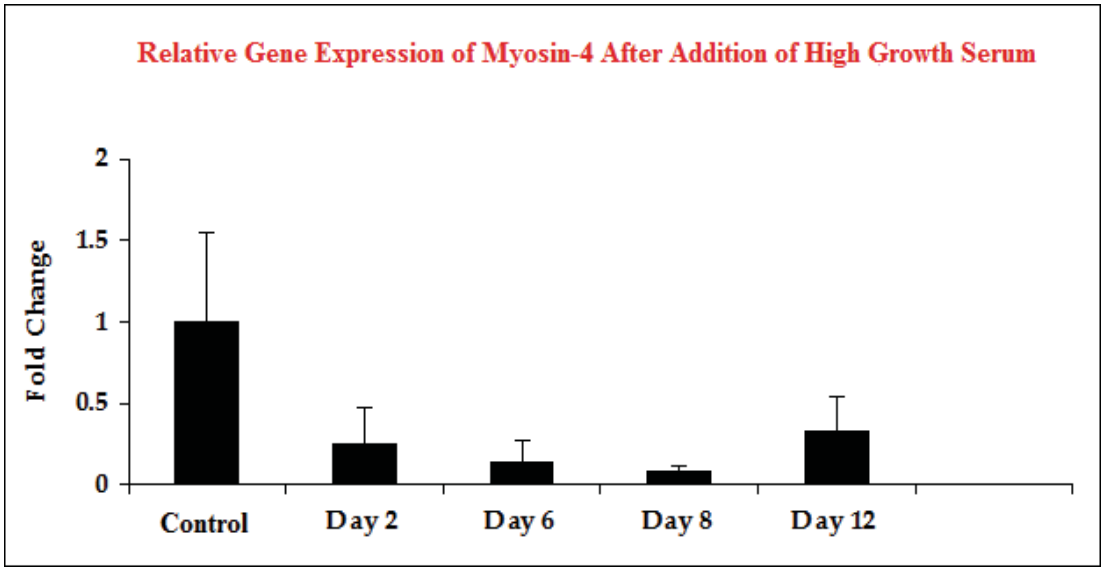


Fig 7. **Relative fold change in Myosin isoform expression.** (A) Myosin-8 (B) Myosin-4. Measured mRNA transcript levels over time course. Time points indicate day high growth serum was added back.

4 DISCUSSION

4.1. Introduction

Although embryonic myogenesis has been well defined, the *in vitro* myogenic program must also be mapped out in such detail. Traditionally, myogenic regulation has been investigated by knock-out/down/in genetic experiments, trailing only few transcription factors at a time. Muscle tissue is also often extracted directly from mice, rats, chickens, and other organisms; and although the tissue is cultivated in a living system, extraction methods can be difficult and the quality and quantity of protein, DNA, and RNA can be very limited. Cell culture technique offers a quick and relatively simple solution to investigating a living system. With the emergence of high-throughput technologies and biocomputational tools, systems biology will help further define developmental processes *in vitro*..

4.2 *In vitro* Myogenesis

Characterization of our proteome revealed *in vitro* late myogensis to be congruent with myogenic progression *in vivo*. With the use of C2C12 cells we forgo early phases of myogenesis such as lineage establishment and begin at differentiation. There is no true migration of cells in culture, but the embryonic myoblasts do maintain their proliferative properties so long as they are fed growth factors. Differentiation is artificially triggered with withdrawal of growth serum. C2C12 cells exit the cell cycle and begin to fuse to form mature myocytes. This switch in global programs was of most interest to us.

4.2.1 High Growth Serum Add-Back

In addition to allowing cells to grow for 13 days, we took a novel approach by adding back mitogens post-differentiation by adding high levels of growth serum to media. Our original hypothesis indicating high levels of mitogens would affect the

transcriptional cascade of the cells in culture. However, our proteomic data indicated addition of mitogens and growth factors post-induction of differentiation did not affect the myogenic program. It seems once these signaling cascades are triggered, they cannot be reset. To further support this hypothesis we see an absence of proteins from both the Notch signaling pathway and the SMAD signaling pathway; both negative regulators of myogensis. Their role *in vivo* is to keep progenitor cells in a proliferative state accounting for the presence of satellite cells for post-natal muscle development (Braun,Gautel et al 2011). Although our sampling relied on molecular weight, both NOTCH3 and SMAD4 at 205 kDa and 48 kDa, respectively, would have conceivably been within selected band regions if present.

More evidence for the contrary to our original hypothesis is the elevated expression of alpha-enolase. Initially thought of solely in its glycolytic role, alpha-enolase converts the precursor 2-phosphoglycerate to phosphoenolpyruvate in the ninth step in glycolysis (Subramanian, Miller, 2000). However, alpha-enolase is elevated at day 0, as well; seeming to be early expression compared to other advanced metabolic machinery. ENO1, the gene coding for the enzyme also codes for myc-binding protein-1 (MBP1) which binds and downregulates the activity of c-myc. C-myc is a well-studied transcription factor directly involved in DNA replication and cell proliferation. C-myc is activated by mitogenic factors, such as those found in fetal bovine serum, through the mitogen activated protein kinase pathway (MAPK pathway) which is well documented to be upregulated during early myogenesis. The artificial addition of mitogens were not enough to activate the, most-likely, bound c-myc within the developing myoblasts (Subramanian, Miller 200). Our revised hypothesis is the myogenic program i*n vitro* is irreversible once terminal differentiation has been induced.

4.2.2. Proteomic Evidence of Myogenic Progression

A hierarchical expression of proteins reflecting significant events revealed itself upon functional annotation. Firstly, the presence of basic transcriptional regulators such as histones and DNA topoisomerase 2-beta which affect DNA accessibility and methylation which functions in alternatively splicing. These events most likely lead to the translation of new proteins necessary for further development of the myotube. Alternatively, the presence of signaling proteins such as Ras GTPase activating like protein and A anchor protein kinase 2 indicate a regulatory function, as well as creation of new cell machinery. Ras GTPase activating like protein is a late downstream target of Ras, which activates other pathways during myogenesis. This protein is thought to participate in actin reorganization during remodeling (Kislinger et al. 2005). An anchor protein kinase 2 binds the subunit protein kinase A (PKA) which acts through cAMP for downstream signaling and may play a role in functional regulation of ion channels (Shwartz, 2011).

The most obvious class of proteins to participate in myogensis is cytoskeletal elements. Although we only identified 210 total proteins compared to Kislinger et al who found ~2000 proteins; both our proteome and their proteome is predominantly cell remodeling and organizational agents such as proteases, cytoskeletal proteins, adhesion proteins, and extracellular membrane proteins. Not only are these classes of proteins functionally enriched according to their GO terms annotations, but they as Fig. 2 shows, their spectral counts were most abundant according to their emPAI's.

The other notable class of proteins that are upregulated upon mature myofiber formation are involved in the biochemical pathways to create energy to allow muscles to contract. ATP binds the myosin II motor protein in the sarcomere; these motor proteins have ATPase domains. Upon the hydrolysis of ATP, myosin move

along actin filaments to produce contraction. Proteins such as cytochrome b-c1 complex subunits indicate the assembly of the electron transport chain crucial for abundant ATP synthesis. Other major metabolic enzymes enriched, include citrate synthase which participates in the first step of the citric acid cycle for the total breakdown of glucose and creatine kinase which functions to form ATP from creatine.

Unexpectedly we also found a significant number of ribosomal proteins, such as 60s ribosomal protein L31, 40s ribosomal protein s19, etc. (Table 5). These proteins appeared at both time points, and most had relatively high emPAI values. Whilst functionally annotating the proteome, common GO terms associated with these proteins were RNA-binding, translation, and ribosome. In fact, the highest GO enrichment score, according to DAVID analysis, was for translation. In recent years, more research has been done in post-transcriptional regulation of gene levels by microRNA (miRNA's) (Braun, Gautel, 2011). These bind to transcription factors, modulating further downstream application. More research is being done on their exact mechanism and function Klerk et al 2015).

Table 5. **Ribosomal proteins.** Proteins identified as ribosomal compartments or RNA-binding proteins. “+” represents presence of protein at Day 0 and Day 6 , and “-” represents absence of protein at Day and Day 6.

Uniprot ID	Day 0	Day 6	Protein names
RLA2_MOUSE	+	+	60S acidic ribosomal protein P2
RL22_MOUSE	+	+	60S ribosomal protein L22 (Heparin-binding protein HBp15)
RS16_MOUSE	+	+	40S ribosomal protein S16
RS19_MOUSE	+	+	40S ribosomal protein S19
RL3_MOUSE	+	+	60S ribosomal protein L3 (J1 protein)
RL30_MOUSE	+	+	60S ribosomal protein L30
IF5A1_MOUSE	+	+	Eukaryotic translation initiation factor 5A-1 (eIF-5A-1)
RS27A_MOUSE	+	+	Ubiquitin-40S ribosomal protein S27a
RS26_MOUSE	+	+	40S ribosomal protein S26
RL34_MOUSE	+	-	60S ribosomal protein L34
RL23_MOUSE	+	+	60S ribosomal protein L23
RL4_MOUSE	+	+	60S ribosomal protein L4
EIF3A_MOUSE	+	+	Eukaryotic translation initiation factor 3 subunit A (eIF3a)
U520_MOUSE	+	+	U5 small nuclear ribonucleoprotein 200 kDa helicase
RL31_MOUSE	+	+	60S ribosomal protein L31
RS20_MOUSE	-	+	40S ribosomal protein S20
IF4A1_MOUSE	+	+	Eukaryotic initiation factor 4A-I (eIF-4A-I)
RRBP1_MOUSE	+	+	Ribosome-binding protein 1 (Ribosome receptor protein)
SMCA4_MOUSE	+	-	Transcription activator BRG1

4.3 Sarcomeric Gene expression

Transcription during any development process is highly diverse, complex, and dynamic; an attempt to measure the transcription of all genes would prove a lofty goal. However, we focused on the process of sarcomerogenesis, and therefore chose to measure mRNA transcript levels for skeletal titin, alpha actin, skeletal, and myosin.

A model for sarcomere formation exists describing titin as a scaffold that binds transiently formed Z-lines, as a first step. The elevated titin expression level during

the control time course (no addition of high growth factors), is congruent with this model as several new sarcomeres form in the cell culture (Rui et al 2010). Addition of high growth factors, even for 24 hours (Day 12) seems to downregulate transcription of the gene. TTN is a large gene with several transcription factor binding sites; any one of these diverse transcription factors are elements in additional signaling pathways which may be directly affected by the addition of mitogens and indirectly affect TTN transcription (Zhang et al 2003)

Contrarily, skeletal actin relative expression remained constant consistent with our proteomic findings. At both Days 0 and Days 6, cytoskeletal proteins served as evidence of drastic actin cytoskeletal remodeling which accounts for the consistently elevated levels of actin. Alternatively, alpha-actin gene expression is under regulation of the serum response growth factor (SRF) which in turn is downtream element in the MAPK pathway which is involved in both cell proliferation and differentiation. For this reason, the addition of high growth serum may not have negatively impacted expression. Further alpha-actin is shown to recruit myogenic elements once upregulated, accounting for the presence proper myogensis-associated proteins in our proteomic data. (Gunning et al 2000)

Lastly, our results for myosin seemed more ambigous, as they were not consistent with our observations in the proteome. However, myosin is a highly conserved protein with several paralogs; therefore mRNA levels in the results may have reflected accidental amplification of other isoforms. Myosin isoforms are found in a gene cluster on chromosome 11 in mice; the regulation of transcription of each isoform according to its respective environmental or temporal stiumli has not been elucidated. (Weiss et al 1999) Added external growth factors may have directly or indirectly affected this transcription.

Myogensis is a prolific process including countless signaling pathways with a series of checks and balances. It is possible the add back of mitogens may have altered the transcription of very specific genes, however, post-transcriptional and post-translation regulation may have prevented an altered proteomic result. Regulation, after all, is present to correct for abnormalities in any complex program (Puente et al 2004)

4.4 Limitations

There are unfortunately, some limitations in our methodologies due mainly to cost. For example, only two time points were chosen for proteomic analysis; ideally, the inclusion of more time points in our investigation would allow for the discovery of more specific processes along the myogenic program. Additionally, gel electrophoresis, although not discriminatory to protein type, depends on the proteins electrophoretic mobility which is hindered by giant proteins or conversely very small proteins may have run off the gel. Selecting proteins for identification based on four molecular weights and protein solubility may have limited our total protein diversity.

Many vital myogenic factors were not detected such as members from the MRF transcription factor family, or a member from the paired homeobox transcription factors. For example Myogenin at 25 kDa and Pax7 at 54 kDa, both important for terminal differentiation, may have been present in the cell lysate but fell outside selected band regions, or were not included in the detergent-insoluble fraction of the total cell lysate. A protein profile from a proliferative myoblast would have been helpful, as well, to directly reflect on the global change immediately prior to differentiation.

4.5 Future work

We will continue looking into the dynamics of the myogenic proteome and

sarcomeric gene expression. The detergent-insoluble fraction of the cell lysate has been filtered to increase likelihood of titin detection. Titin was identified through MASCOT in our first set of proteomic data, however it received an emPAI value of 0.00 and could not be considered in our analysis. As titin is a giant protein, its electrophoretic mobility is limited. Bands at >212 kDa from all time points have been selected and sent to the Ohio State University Mass Spectroscopy facility. Data will be functionally annotated in the same fashion as our data discussed here.

Efforts are also being put forth to more accurately measure mRNA transcript levels of isoforms of myosin. Our previous approach included a one-step RT-qPCR methodology; however, we are generating cDNA from total RNA extracted from developing cells. qPCR will be carried out on the cDNA, and data will be analyzed through DART-PCR. Also crucial for accurate measurement of relative gene expression is the inclusion of a housekeeping gene; we are developing a reliable housekeeping gene for our system. With the tumultuous intracellular signaling processes during development, finding a gene with a constant level of expression proves difficult (Jeong et al 2014). So far this study has provided a road map of essential events in the myogenic program, using an *in vitro* approach. This will allow for more targeted studies in the future as events in cell culture will be easier to pinpoint.

REFERENCES

Ashburner, Michael, Catherine A. Ball, Judith A. Blake, David Botstein, Heather Butler, J. Michael Cherry, Allan P. Davis, Kara Dolinski, Selina S. Dwight, Janan T. Eppig, Midori A. Harris, David P. Hill, Laurie Issel-Tarver, Andrew Kasarskis, Suzanna Lewis, John C. Matese, Joel E. Richardson, Martin Ringwald, Gerald M. Rubin, and Gavin Sherlock. "Gene Ontology: Tool for the Unification of Biology." *Nat Genet Nature Genetics* 25.1 (2000): 25-29. Web.

Biressi, Stefano, Mario Molinaro, and Giulio Cossu. "Cellular Heterogeneity during Vertebrate Skeletal Muscle Development." *Developmental Biology* 308.2 (2007): 281-93. Web.

Braun, Thomas, and Mathias Gautel. "Transcriptional Mechanisms Regulating Skeletal Muscle Differentiation, Growth and Homeostasis." *Nature Reviews Molecular Cell Biology Nat Rev Mol Cell Biol* 12.6 (2011): 349-61. Web.

Diel, P., D. Baadners, K. Schlupmann, M. Velders, and J. P. Schwarz. "C2C12 Myoblastoma Cell Differentiation and Proliferation Is Stimulated by Androgens and Associated with a Modulation of Myostatin and Pax7 Expression." *Journal of Molecular Endocrinology* 40.5 (2008): 231-41. Web.

Eilertsen, K. J. "Cellular Titin Localization in Stress Fibers and Interaction with Myosin II Filaments *in Vitro*." *The Journal of Cell Biology* 126.5 (1994): 1201-210. Web.

Gunning P, Ponte P, Okayama H, Engel J, Blau H, Kedes L. "Isolation and characterization of full-length cDNA clones for human alpha-, beta-, and gamma-actin mRNAs: skeletal but not cytoplasmic actins have an

amino-terminal cysteine that is subsequently removed". *Mol Cell Biol* **3** 5 (1983): 787–95. Web.

Ishihama, Y. "Exponentially Modified Protein Abundance Index (emPAI) for Estimation of Absolute Protein Amount in Proteomics by the Number of Sequenced Peptides per Protein." *Molecular & Cellular Proteomics* 4.9 (2005): 1265-272. Web.

Jeong, Jae-Kyo, Min-Hee Kang, Sangiliyandi Gurunathan, Ssang-Goo Cho, Chankyu Park, Han Seo, and Jin-Hoi Kim. "Evaluation of Reference Genes in Mouse Preimplantation Embryos for Gene Expression Studies Using Real-time Quantitative RT-PCR (RT-qPCR)." *BMC Research Notes BMC Res Notes* 7.1 (2014): 675. Web.

Kislinger, T. "Proteome Dynamics during C2C12 Myoblast Differentiation." *Molecular & Cellular Proteomics* 4.7 (2005): 887-901. Web.

Klerk, E. De, I. F. A. C. Fokkema, K. A. M. H. Thiadens, J. J. Goeman, M. Palmblad, J. T. Den Dunnen, M. Von Lindern, and P. A. C. 't Hoen. "Assessing the Translational Landscape of Myogenic Differentiation by Ribosome Profiling." *Nucleic Acids Research* 43.9 (2015): 4408-428. Web.

Pandorf, C. E., F. Haddad, R. R. Roy, A. X. Qin, V. R. Edgerton, and K. M. Baldwin. "Dynamics of Myosin Heavy Chain Gene Regulation in Slow Skeletal Muscle: ROLE OF NATURAL ANTISENSE RNA." *Journal of Biological Chemistry* 281.50 (2006): 38330-8342. Web.

Parker, Maura H., Patrick Seale, and Michael A. Rudnicki. "Looking Back to the Embryo: Defining Transcriptional Networks in Adult Myogenesis." *Nat Rev Genet Nature Reviews Genetics* 4.7 (2003): 497-507. Web.

Peirson, S. N. "Experimental Validation of Novel and Conventional Approaches to

Quantitative Real-time PCR Data Analysis." *Nucleic Acids Research* 31.14 (2003): n. pag. Web.

Puente, Lawrence G., Jean-François Carrière, John F. Kelly, and Lynn A. Megeney. "Comparative Analysis of Phosphoprotein-enriched Myocyte Proteomes Reveals Widespread Alterations during Differentiation." *FEBS Letters* 574.1-3 (2004): 138-44. Web.

Ríos, Ramón, Isabel Carneiro, Víctor M. Arce, and Jesús Devesa. "Myostatin Regulates Cell Survival during C2C12 Myogenesis." *Biochemical and Biophysical Research Communications* 280.2 (2001): 561-66. Web.

Rossi, Giuliana, and Graziella Messina. "Comparative Myogenesis in Teleosts and Mammals." *Cellular and Molecular Life Sciences Cell. Mol. Life Sci.* 71.16 (2014): 3081-099. Web.

Rui, Yanning, Jianwu Bai, and Norbert Perrimon. "Sarcomere Formation Occurs by the Assembly of Multiple Latent Protein Complexes." *PLoS Genetics PLoS Genet* 6.11 (2010): n. pag. Web.

Schwartz, J. H. "The Many Dimensions of CAMP Signaling." *Proceedings of the National Academy of Sciences* 98.24 (2001): 13482-3484. Web.

Subramanian, A. "Structural Analysis of Alpha -Enolase. MAPPING THE FUNCTIONAL DOMAINS INVOLVED IN DOWN-REGULATION OF THE C-myc PROTOONCOGENE." *Journal of Biological Chemistry* 275.8 (2000): 5958-965. Web.

Supek, Fran, Matko Bošnjak, Nives Škunca, and Tomislav Šmuc. "REVIGO Summarizes and Visualizes Long Lists of Gene Ontology Terms." *PLoS ONE* 6.7 (2011): n. pag. Web.

Tannu, N. S. "Comparative Proteomes of the Proliferating C2C12 Myoblasts and

Fully Differentiated Myotubes Reveal the Complexity of the Skeletal Muscle Differentiation Program." *Molecular & Cellular Proteomics* 3.11 (2004): 1065-082. Web.

Tskhovrebova, L., and J. Trinick. "Properties of Titin Immunoglobulin and Fibronectin-3 Domains." *Journal of Biological Chemistry* 279.45 (2004): 46351-6354. Web.

Watkins, Thomas C., Lisa M. Zelinka, Matt Kesic, Carl F. Ansevin, and Gary R. Walker. "Identification of Skeletal Muscle Autoantigens by Expression Library Screening Using Sera from Autoimmune Rippling Muscle Disease (ARMD) Patients." *J. Cell. Biochem. Journal of Cellular Biochemistry* 99.1 (2006): 79-87. Web.

Weiss, A., D. Mcdonough, B. Wertman, L. Acakpo-Satchivi, K. Montgomery, R. Kucherlapati, L. Leinwand, and K. Krauter. "Organization of Human and Mouse Skeletal Myosin Heavy Chain Gene Clusters Is Highly Conserved." *Proceedings of the National Academy of Sciences* 96.6 (1999): 2958-963. Web.

Wu, Xiaogang, Mohammad Al Hasan, and Jake Yue Chen. "Pathway and Network Analysis in Proteomics." Journal of Theoretical Biology 362 (2014): 44-52. Web.

Zhang, Xiao-Lian, Xue-Ju Qu, and Inder K. Vijay. "STAT5a Regulates the GlcNAc-1-phosphate Transferase Gene Transcription and Expression." *Cellular Physiology and Biochemistry Cell Physiol Biochem* 13.2 (2003): 85-92. Web.

Zhou, Yuefang, Dan Liu, and Henry J. Kaminski. "Myosin Heavy Chain Expression in Mouse Extraocular Muscle: More Complex Than Expected." *Invest. Ophthalmol. Vis. Sci. Investigative Opthalmology & Visual Science* 51.12 (2010): 6355-363. Web.

APPENDICES

A1. Materials

Mammalian Cell Culture	
Product	Manufacturer
C2C12 (ATCCR CRL-1772™)	American Type Culture Collection (ATCC) Manassas, VA
Dublecco's Modified Eagle's Media (DMEM)	American Type Culture Collection (ATCC) Manassas, VA
Medium 199	BioWhittaker Walkersville, MD
Fetal Bovine Serum (FBS)	American Type Culture Collection (ATCC) Manassas, VA
0.05%Trypsin- 0.53mM EDTA	Corning Life Sciences, Manassas, VA

Gel Electrophoresis	
Product	Manufacturer
Criterion TGX 12% 1mm Gel	Bio-Rad Laboratories Hercules, CA
Wide Range Protein Molecular Weight Standard	Amresco, Corporation, Solon OH
Equipment	Manufacturer
Power PAC 3000	Bio-Rad Laboratories Hercules, CA

Gel Analysis	
Equiment/Software	Manufacturer
Pharos FX Plus Molecular Imager	Bio-Rad Laboratories Hercules, CA
Quantity One 1-D Analysis Sofware	Bio-Rad Laboratories Hercules, CA

RNA extraction and PCR	
Products	Manufacturer
RNeasy Mini Kit	Qiagen, Inc. Valencia, CA
OneStep RT-PCR Kit	Qiagen, Inc. Valencia, CA
iTaq$_{tm}$ Universal SYBR® Green One-Step Kit	Bio-Rad Laboratories Hercules, CA
Equipment/Software	Manufacturer
IQ$_{tm}$5 Multicolor Real-Time PCR Detection System	Bio-Rad Laboratories Hercules, CA
iCylcer	Bio-Rad Laboratories Hercules, CA
iQ$_{tm}$ 96 well PCR plates, semi-skirted, high-profile	Bio-Rad Laboratories Hercules, CA
IQ$_{tm}$5 Optical System Software	Bio-Rad Laboratories Hercules, CA

Miscellaneous Chemicals and Equipment	
Product	Manufacturer
Coomaissee Brilliant Blue R-250	Fisher Scientific Pittsburgh, PA
Glacial acetic acid	
Protease inhibitor tablet	Roche, USA San Francisco, CA
Sodium dodecyl sulfate (SDS)	Amresco Corporation, Solon, OH
Triton-X-100	Sigma-Aldrich Corporation, St. Louis, MO
Equipment	Manufacturer
Olympus LH 50A Inverted Phase Contrast Microscope	Olympus America, Center Valley, PA

A2. Solutions and Media

Mammalian Cell Culture	
Media	Recipe
High Serum Growth Media	790mL DMEM 200mL FBS 10mL antibiotic
Low Serum Growth Media	980mL DMEM 10mL FBS 10mL antibiotic

Cell Lysate Extraction	
Solution	Recipe
4X SDS Sample Buffer	12.5 g glycerol 0.76 g tris base 2.30 g SDS 12.3 ml 0.5 M hydrochloric acid 17.75 ml deionized water
Isotonic Wash Buffer	1.211 g tris base 0.406 g magnesium chloride heptahydrate 51.345 g sucrose 500 ml deionized water Adjust pH to 7.4
Triton-x-100 Extraction Buffer	100 ml Isotonic Wash Buffer 0.5 ml Triton X-100 protease inhibitor tablet

Gel Eletrophoresis and Analysis	
Solution	Recipe
Tris-Glycine-SDS Running Buffer (TGS)	800 ml deionized water 30.2 g tris base 10.0 g SDS 188 glycine
Coomassie Stain	2.5 g Coomassie Brilliant Blue R-250 100 ml Glacial Acetic Acid 450 Methanol 450 ml H20
High Destain	100 ml glacial acetic acid 400 ml methanol 500 ml deionized water
Low Destain	100 ml glacial acetic acid 150 ml methanol 750 ml deionized water

A3. Time Course Calendars

Proteomics Time Course Calendar (13 Days)		
Sample Time Point	No. Of Days Treated with High Serum Add-back	Day of Harvest
Day 0	None	Day 1
Day 2	11 Days	Day 13
Day 4	9 Days	Day 13
Day 6	7 Days	Day 13
Day 8	5 Days	Day 13
Day 12	1 Day	Day 13

Gene Expression Time Course Calendar (13 Days)		
Sample Time Point	No. Of Days Treated with High Serum Add-back	Day of Harvest
Control	None	Day 13
Day 2	11 Days	Day 13
Day 4	9 Days	Day 13
Day 6	7 Days	Day 13
Day 8	5 Days	Day 13
Day 12	1 Day	Day 13

A4. Bioinformatic Online Resources

Databse/Software	Website
Database for Annotation, Visualization and Integration Discovery (DAVID)	http://david.niaid.nih.gov
Gene Ontology Consortium	http://geneontology.org/
MetaboAnalyst 3.0	http://www.metaboanalyst.ca/faces/ModuleView.xhtml
Uniprot	http://www.uniprot.org/

A5. Proteomic Data

Day 0

Uniprot ID	Protein	emPAI
RS27A_MOUSE	Ubiquitin-40S ribosomal protein S27a	0.4
ACTB_MOUSE	Actin, cytoplasmic 1	0.35
NEST_MOUSE	Nestin	0.32
TBA1A_MOUSE	Tubulin alpha-1A chain	0.29
H2B1B_MOUSE	Histone H2B type 1-B	0.29
CLH1_MOUSE	Clathrin heavy chain 1	0.29
HNRPD_MOUSE	Heterogeneous nuclear ribonucleoprotein	0.28
RS26_MOUSE	40S ribosomal protein S26	0.26
ACTA_MOUSE	Actin, aortic smooth muscle	0.25
RL34_MOUSE	60S ribosomal protein L34	0.25
ACTBL_MOUSE	Beta-actin-like protein 2	0.25
MAP4_MOUSE	Microtubule-associated protein 4	0.24
RL23_MOUSE	60S ribosomal protein L23	0.23
HNRPF_MOUSE	Heterogeneous nuclear ribonucleoprotein F	0.23
VATL_MOUSE	V-type proton ATPase 16 kDa proteolipid subunit	0.22
RL4_MOUSE	60S ribosomal protein L4	0.22
H3C_MOUSE	Histone H3.3C	0.22
TOM20_MOUSE	Mitochondrial import receptor subunit TOM20 homolog	0.21
OSTC_MOUSE	Oligosaccharyltransferase complex subunit OSTC	0.2
IF1AX_MOUSE	Eukaryotic translation initiation factor 1A	0.2
CISY_MOUSE	Citrate synthase, mitochondrial	0.2
VIME_MOUSE	Vimentin	0.2
RS27A_MOUSE	Ubiquitin-40S ribosomal protein S27a	0.18
MYH10_MOUSE	Myosin-10	0.17
TBB5_MOUSE	Tubulin beta-5 chain	0.14
SPTN1_MOUSE	Spectrin alpha chain, non-erythrocytic 1	0.13
CLCF1_MOUSE	Cardiotrophin-like cytokine factor 1	0.13
MYH8_MOUSE	Myosin-8	0.12
TOP2B_MOUSE	DNA topoisomerase 2-beta	0.11
TLN1_MOUSE	Talin-1	0.1
FLNC_MOUSE	Filamin-C	0.1
RRBP1_MOUSE	Ribosome-binding protein 1	0.08
ILF2_MOUSE	Interleukin enhancer-binding factor 2	0.08
RBMX_MOUSE	RNA-binding motif protein, X chromosome	0.08
KCRB_MOUSE	Creatine kinase B-type	0.08
ASSY_MOUSE	Argininosuccinate synthase	0.07
TADBP_MOUSE	TAR DNA-binding protein 43	0.07
SERPH_MOUSE	Serpin H1	0.07
PSMD6_MOUSE	26S proteasome non-ATPase regulatory subunit 6	0.07

INF2_MOUSE	Inverted formin-2	0.07
FLNB_MOUSE	Filamin-B	0.07
PGBM_MOUSE	Basement membrane-specific heparan sulfate proteoglycan core protein	0.06
SPTB2_MOUSE	Spectrin beta chain, non-erythrocytic 1	0.06
DYHC1_MOUSE	Cytoplasmic dynein 1 heavy chain 1	0.06
CA2D1_MOUSE	Voltage-dependent calcium channel subunit alpha-2/delta-1	0.06
NID2_MOUSE	Nidogen-2	0.04
EIF3A_MOUSE	Eukaryotic translation initiation factor 3 subunit A	0.04
RTN4_MOUSE	Reticulon-4	0.03
U520_MOUSE	U5 small nuclear ribonucleoprotein 200 kDa helicase	0.03
IQGA1_MOUSE	Ras GTPase-activating-like protein	0.03
AKAP2_MOUSE	A-kinase anchor protein 2	0.03
PRP8_MOUSE	Pre-mRNA-processing-splicing factor 8	0.02
MACF1_MOUSE	Microtubule-actin cross-linking factor 1	0.02
SMCA4_MOUSE	Transcription activator BRG1	0.02
FBLN2_MOUSE	Fibulin-2	0.02
TPR_MOUSE	Nucleoprotein TPR	0.01

DAY 6

Uniprot ID	Protein	emPAI
ACTB_MOUSE	Actin, cytoplasmic 1	14.22
PPIA_MOUSE	Peptidyl-prolyl cis-trans isomerase A	9.96
ACTC_MOUSE	Actin, alpha cardiac muscle 1	8.56
ACTS_MOUSE	Actin, alpha skeletal muscle	8.5
RLA2_MOUSE	60S acidic ribosomal protein P2	5.11
H2A2B_MOUSE	Histone H2A type 2-B	4.73
MYH9_MOUSE	Myosin-9	3.4
H2A1F_MOUSE	Histone H2A type 1-F	2.63
MLRS_MOUSE	Myosin regulatory light chain 2, skeletal muscle isoform	2.12
MYH3_MOUSE	Myosin-3	2.08
RL22_MOUSE	60S ribosomal protein L22	1.82
EF1A1_MOUSE	Elongation factor 1-alpha 1	1.76
RS16_MOUSE	40S ribosomal protein S16	1.54
H2B1B_MOUSE	Histone H2B type 1-B	1.39
RL31_MOUSE	60S ribosomal protein L31	1.33
RS19_MOUSE	40S ribosomal protein S19	1.15
SERPH_MOUSE	Serpin H1	0.98
NEST_MOUSE	Nestin	0.88
H3C_MOUSE	Histone H3.3C	0.82
H32_MOUSE	Histone H3.2	0.82

IQGA1_MOUSE	Ras GTPase-activating-like protein IQGAP1	0.79
RL3_MOUSE	60S ribosomal protein L3	0.73
QCR2_MOUSE	Cytochrome b-c1 complex subunit 2, mitochondrial	0.7
MYH7_MOUSE	Myosin-7	0.61
MYH1_MOUSE	Myosin-1	0.61
FLNA_MOUSE	Filamin-A	0.58
CISY_MOUSE	Citrate synthase, mitochondrial	0.54
FINC_MOUSE	Fibronectin	0.51
H31_MOUSE	Histone H3.1	0.48
TOM20_MOUSE	Mitochondrial import receptor subunit TOM20 homolog	0.45
NDKA_MOUSE	Nucleoside diphosphate kinase A	0.43
PLEC_MOUSE	Plectin	0.42
HNRPF_MOUSE	Heterogeneous nuclear ribonucleoprotein F	0.41
ATPB_MOUSE	ATP synthase subunit beta, mitochondrial	0.41
TNNC1_MOUSE	Troponin C, slow skeletal and cardiac muscles	0.4
RS27A_MOUSE	Ubiquitin-40S ribosomal protein S27a	0.4
QCR1_MOUSE	Cytochrome b-c1 complex subunit 1, mitochondrial	0.35
ACTA_MOUSE	Actin, aortic smooth muscle	0.35
CLH1_MOUSE	Clathrin heavy chain 1	0.33
DNJA2_MOUSE	DnaJ homolog subfamily A member 2	0.32
ANXA7_MOUSE	Annexin A7	0.29
HNRPD_MOUSE	Heterogeneous nuclear ribonucleoprotein	0.28
FLNC_MOUSE	Filamin-C	0.27
RS26_MOUSE	40S ribosomal protein	0.26
RL30_MOUSE	60S ribosomal protein L30	0.26
SEPT2_MOUSE	Septin-2	0.26
NSF1C_MOUSE	NSFL1 cofactor p47	0.26
RS20_MOUSE	40S ribosomal protein S20	0.25
TFR1_MOUSE	Transferrin receptor protein 1	0.25
TXND5_MOUSE	Thioredoxin domain-containing protein 5	0.23
RL23_MOUSE	60S ribosomal protein L23	0.23
ASSY_MOUSE	Argininosuccinate synthase	0.23
PRS8_MOUSE	26S protease regulatory subunit	0.23
PSD11_MOUSE	26S proteasome non-ATPase regulatory subunit 11	0.22
MYH3_MOUSE	Myosin-3	0.22
VATL_MOUSE	V-type proton ATPase 16 kDa proteolipid subunit	0.22
PRS7_MOUSE	26S protease regulatory subunit 7	0.22
ENOA_MOUSE	Alpha-enolase	0.22
ARP3_MOUSE	Actin-related protein 3	0.22
RS14_MOUSE	40S ribosomal protein S14	0.21
ACOT2_MOUSE	Acyl-coenzyme A thioesterase 2, mitochondrial	0.21
VIME_MOUSE	Vimentin	0.2
NDUAD_MOUSE	NADH dehydrogenase [ubiquinone] 1 alpha	0.2

	subcomplex subunit 13	
IF5A1_MOUSE	Eukaryotic translation initiation factor 5A-1	0.2
CALM_MOUSE	Calmodulin	0.2
VIME_MOUSE	Vimentin	0.2
SCMC1_MOUSE	Calcium-binding mitochondrial carrier protein SCaMC-1	0.2
OSTC_MOUSE	Oligosaccharyltransferase complex subunit OSTC	0.2
ATPD_MOUSE	ATP synthase subunit delta, mitochondrial	0.19
IF4G1_MOUSE	Eukaryotic translation initiation factor 4 gamma 1	0.18
ACTA_MOUSE	Actin, aortic smooth muscle	0.16
TADBP_MOUSE	TAR DNA-binding protein 43	0.15
OLA1_MOUSE	Obg-like ATPase 1	0.15
MYH10_MOUSE	Myosin-10	0.15
RL4_MOUSE	60S ribosomal protein L4	0.14
FLNB_MOUSE	Filamin-B	0.14
EF1A1_MOUSE	Elongation factor 1-alpha 1	0.14
OST48_MOUSE	Dolichyl-diphosphooligosaccharide--protein glycosyltransferase	0.14
BZW1_MOUSE	Basic leucine zipper and W2 domain-containing protein 1	0.14
DESM_MOUSE	Desmin	0.13
INF2_MOUSE	Inverted formin-2	0.12
TLN1_MOUSE	Talin-1	0.11
PGBM_MOUSE	Basement membrane-specific heparan sulfate proteoglycan core protein	0.11
SPTN1_MOUSE	Spectrin alpha chain, non-erythrocytic 1	0.11
TNNT2_MOUSE	Troponin T, cardiac muscle	0.09
NID2_MOUSE	Nidogen-2	0.09
MAK16_MOUSE	Protein MAK16 homolog	0.09
DYHC1_MOUSE	Cytoplasmic dynein 1 heavy chain 1	0.09
SPB6_MOUSE	Serpin B6	0.08
RTN4_MOUSE	Reticulon-4	0.08
MYH4_MOUSE	Myosin-4	0.08
DAZP1_MOUSE	DAZ-associated protein 1	0.08
APJ_MOUSE	Apelin receptor	0.08
KCRB_MOUSE	Creatine kinase B-type	0.08
ACTY_MOUSE	Beta-centractin	0.08
TOP2B_MOUSE	DNA topoisomerase 2-beta	0.07
TBL2_MOUSE	Transducin beta-like protein 2	0.07
SNF5_MOUSE	SWI/SNF-related matrix-associated actin-dependent regulator of chromatin subf	0.07
ENOA_MOUSE	Alpha-enolase	0.07
EF2_MOUSE	Elongation factor 2	0.07
U520_MOUSE	U5 small nuclear ribonucleoprotein 200 kDa helicase	0.07

TBB5_MOUSE	Tubulin beta-5 chain	0.07
SUCB1_MOUSE	Succinyl-CoA ligase [ADP-forming] subunit beta, mitochondrial	0.07
IF4A1_MOUSE	Eukaryotic initiation factor 4A-I	0.07
CTBP1_MOUSE	C-terminal-binding protein 1	0.07
TBA1B_MOUSE	Tubulin alpha-1B chain	0.06
PPGB_MOUSE	Lysosomal protective protein	0.06
MAP4_MOUSE	Microtubule-associated protein 4	0.06
LMAN1_MOUSE	Protein ERGIC-53	0.06
GTR1_MOUSE	Solute carrier family 2, facilitated glucose transporter member 1	0.06
COR1C_MOUSE	Coronin-1C	0.06
SEPT7_MOUSE	Septin-7	0.06
NEST_MOUSE	Nestin	0.06
MYOF_MOUSE	Myoferlin	0.06
SOAT1_MOUSE	Sterol O-acyltransferase 1	0.05
GPC1_MOUSE	Glypican-1	0.05
VIGLN_MOUSE	Vigilin	0.05
FBLN2_MOUSE	Fibulin-2	0.05
CLH1_MOUSE	Clathrin heavy chain 1	0.05
PRP8_MOUSE	Pre-mRNA-processing-splicing factor 8	0.04
MYH7_MOUSE	Myosin-7	0.04
HNRPU_MOUSE	Heterogeneous nuclear ribonucleoprotein U	0.04
MYH10_MOUSE	Myosin-10	0.04
LMNA_MOUSE	Prelamin-A/C	0.04
TLN2_MOUSE	Talin-2	0.03
MYOF_MOUSE	Myoferlin	0.03
CHD5_MOUSE	Chromodomain-helicase-DNA-binding protein 5	0.03
RTN4_MOUSE	Reticulon-4	0.03
MAP4_MOUSE	Microtubule-associated protein 4	0.03
CA2D1_MOUSE	Voltage-dependent calcium channel subunit alpha-2/delta-1	0.03
LAMB1_MOUSE	Laminin subunit beta-1	0.02
KTN1_MOUSE	Kinectin	0.02
SPTB2_MOUSE	Spectrin beta chain, non-erythrocytic 1	0.02
RRBP1_MOUSE	Ribosome-binding protein 1	0.02
PLEC_MOUSE	Plectin	0.02
MRC2_MOUSE	C-type mannose receptor 2	0.02
FLNA_MOUSE	Filamin-A	0.02
EIF3A_MOUSE	Eukaryotic translation initiation factor 3 subunit A	0.02
ECM29_MOUSE	Proteasome-associated protein ECM29 homolog	0.02
U520_MOUSE	U5 small nuclear ribonucleoprotein 200 kDa helicase	0.01
TPR_MOUSE	Nucleoprotein TPR	0.01
RBP2_MOUSE	E3 SUMO-protein ligase RanBP2	0.01

MAP1B_MOUSE	Microtubule-associated protein 1B	0.01
TLN1_MOUSE	Talin-1	0.01
MYH9_MOUSE	Myosin-9	0.01
MACF1_MOUSE	Microtubule-actin cross-linking factor 1	0.01