A Functional Genomics Approach for Characterizing the Role of Six Transcription Factors in Muscle Development

Alphonse Chu

Thesis submitted to the
Faculty of Graduate and Postdoctoral Studies
In partial fulfillment of the requirements
For the MSc degree in Biochemistry

Department of Biochemistry, Microbiology and Immunology
Faculty of Medicine
University of Ottawa

Abstract

Proper development of skeletal muscle occurs through a highly complex process where activation and repression of genes are essential. Control of this process is regulated by timely and spatial expression of specific transcription factors (TFs). Six1 and Six4 are homeodomain TFs known to be essential for skeletal muscle development in mice. Using the C2C12 cell line, a model for skeletal muscle differentiation, I used a functional genomics approach, employing siRNA specific to both these TFs, to characterize their role in skeletal myogenesis. To identify the genes that are regulated by both these TFs, gene expression profiling by microarray of cells treated with siRNA against Six1 and/or Six4 was performed. The knock-down of these TFs caused lower expression of markers of terminal differentiation genes in addition to an impairment of myoblast fusion and differentiation. Interestingly, transcript profiling of cells treated with siRNA against myogenin revealed that several of the Six1 and Six4 target genes are also regulated by myogenin. Through a combination of bioinformatic analyses it was also found that specific knock-down of Six4 causes an up-regulation of genes involved in mitosis and the cell cycle. In summary, these results show that Six1 and Six4 can both independently regulate different genes, but can also cooperate together with other TFs where they play an important role in the proper regulation of skeletal myogenesis.

Acknowledgements

I would first like to thank Dr. Alexandre Blais for providing me the opportunity to work in his lab and teaching me many of the laboratory techniques that I have mastered during the course of my graduate degree. I would like to acknowledge also the members of my lab, Yubing Liu and Imane Chakroun, who have provided critical comments and suggestions to improve the results of my project and during the editing of this thesis. Additionally, without their help, I would have not been able to obtain and work with primary myoblasts.

I also would like to thank members of my thesis advisory committee, Dr. Ilona Skerjanc, Dr. Mary-Ellen Harper and Dr. Mads Kaern that have provided important suggestions and constructive comments on my project. In addition, I would like to thank the members of the Skerjanc lab for providing good constructive criticisms of my results and research project.

I would especially like to thank the students and researchers from the 4th floor of RGN hall, particularly the labs of the OISB in addition to the BMIGSA student council, for creating a fun working environment. A special thanks to the COOP Education Program and staff for providing me with the opportunity to gain valuable laboratory experience during work-terms.

Finally, I would like to thank my parents and sister for their support and encouragements throughout my education. Of course, this work would not have been possible without the financial support provided by the Muscular Dystrophy Association (MDA), the Natural Sciences and Engineering Research Council of Canada (NSERC), the Canadian Fund for Innovation (CFI) and the University of Ottawa awarded to Dr. Alexandre Blais. Special thanks to the Fonds de Recherche en Santé du Québec (FRSQ) for awarding me the Master's Training Scholarship.

Table of Contents

List of Abbreviations

AD	Transcriptional Activation Domain
ALP	Alkaline Phosphotase
ANOVA	Analysis of Variance
ARE	Atp1a1 Regulatory Element
ATCC	American Type Culture Collection
ATP	Adenosine Triphosphate
bFGF	Basic Fibroblast Growth Factor
bHGH	Basic Hepatocyte Growth Factor
bHLH	Basic Helix-Loop-Helix
bp	Base Pairs
BrdU	Bromodeoxyuridine
BRE	TFIIB Recognition Element
CBP	CREB Binding Protein
cDNA	Complementary Deoxyribonucleic Acid
ChIP	Chromatin Immunoprecipitation
ChIP-on-Chip	Chromatin Immunoprecipitation followed by Microarray Hybridization
ChIP-re-ChIP	Chromatin Immunoprecipitation followed by another ChIP with a different Antibody
ChIP-Seq	Chromatin Immunoprecipitation followed by Sequencing
CKI	Cyclin Dependent Kinase Inhibitor
CLICK	CLuster Identification via Connectivity Kernels
cRNA	Complementary Ribonucleic Acid
CTD	C-Terminal Domain of RNA Polymerase II
Cy3	Cyanine 3 Dye
Cy5	Cyanine 5 Dye
dach	Dachshund
DAPI	4',6-diamidino-2-phenylindol
DM	Differentiation Medium
DMEM	Dulbecco's Modified Eagle Medium
DMSO	Dimethyl Sulfoxide
DNA	Deoxyribonucleic Acid
dNTP	Deoxyribonucleotide triphosphate
DPE	Downstream Promoter Element
DSHB	The Developmental Studies Hybridoma Bank
dsRNA	Double Stranded RNA
ey	Eyless
FACS	Fluorescence-Activated Cell Sorting
FBS	Fetal Bovine Serum
FDR	False Discovery Rate
FRET	Fluorescence Resonance Energy Transfer or Förster resonance energy transfer
FWER	Family-Wise Error Rate
Gln	Glutamine
GM	Growth Medium
GO	Gene Ontology Analysis
GSEA	Gene Set Enrichment Analysis
H3K27me	Histone H3 Lysine 27 Mono-methylation
H3K27me2	Histone H3 Lysine 27 Di-methylation
H3K27me3	Histone H3 Lysine 27 Tri-methylation
H3K4me2	Histone H3 Lysine 4 Di-methylation

H3K4me3	Histone H3 Lysine 4 Tri-methylation
H3K9me	Histone H3 Lysine 9 Mono-methylation
HAT	Histone Acetyltransferase
HCS	Highly Connected Subgraphs
HD	Homeodomain
HRP	Horse Raddish Peroxidase
kb	Kilobase Pairs
MADS	MCM1, Agamous, Deficiens, Serum Response Factor
MB	Myoblast
MD	Muscular Dytrophies
Mef2	Myogenic Enhancer Factor 2 Family of Transcription Factors
MHC	Myosin Heavy Chain
miRNA	MicroRNA
MMLV-RT	Moloney Murine Leukemia Virus Reverse Transcriptase
MRF	Myogenic Regulatory Factors
MT	Myotube
NS	Non Specific or Non Targeting siRNA Sequence
nt	Nucleotide
P/S	Penicillin and Streptomycin
PBS	Phosphate Buffered Saline
PBST	Phosphate Buffered Saline with 0.1% Triton X-100
PBST-BSA	Phosphate Buffered Saline with 0.1% Triton X-100 and 3% BSA Fraction V
PCR	Polymerase Chain Reaction
PIC	Preinitiation Complex
PTM	Post-Translational Modification
PVDF	Polyvinylidene fluoride
PWM	Position Weight Matrix
qRT-PCR	Quantitative Reverse-Transcriptase PCR
RISC	RNA Induced Silencing Complex
RNA	Ribonucleic Acid
RNAi	RNA Interference
RNA-Seq	RNA Sequencing
RT-PCR	Reverse-Transcriptase Polymerase Chain Reaction
SD	Six Domain
SEM	Standard Error over the Mean
SFM	Serum-Free Medium (DMEM)
shh	Sonic Headgehog
shRNA	Short Hairpin RNA
siRNA	Small Interfering RNA
so	Drosophila Sine Oculis
SOM	Self-Organizing Maps
ssRNA	Single Stranded RNA
SYBR	SYBR Green I Dye
TAF	TBP Associated Factor
TBP	TATA Binding Protein
TF	Transcription Factors
TIFF	Tagged Image File Format
TSS	Transcriptional Start Site
UTP	Uridine Triphosphate
XDR	Extended Dynamic Range

List of Figures and Illustrations

List of Tables

List of Supplemental Tables

1. Chapter 1 – Introduction

Muscular dystrophies (MD) are a group of hereditary muscle disorders with an incidence rate of 1 in 3500 and a prevalence rate in the male population of 1 in 25000 (Emery, 1991). These disorders are characterized by progressive weakening and wasting of skeletal muscle. MDs can be categorized in many different subgroups depending on the distribution of the muscle weakness, which include Duchenne and Becker; Emery-Dreifuss; distal; facioscapulohumeral; oculopharyngeal; and limb-girdle (Reviewed in Emery, 2002). The severity of the symptoms affecting the patients varies between the types of MDs. However, patients affected by these diseases have considerably reduced mobility and quality of life. Affected individuals may also develop a number of health complications as they age, of note cardiomyopathy and lung failure (Gilroy et al., 1963; Voit et al., 1988) which consequently leads to decreased life expectancy of affected individuals. Currently, treatments are available to treat the symptoms of patients affected by MDs, but no cures are available for patients with MD, therefore it is important to continue research in this field.

1.1. Development of Therapies for Muscular Dystrophy Patients

Many therapeutic approaches have been explored such as, cell based and gene therapies. Cell based therapies involves transplantation or injection of muscle precursor cells into degenerating muscle tissues in the hope of regenerating it. Many obstacles still remain before cell-based therapies can be used to treat patients with MD. For instance, prevention of immune rejection and improving the survival of injected cells are still challenges to overcome. Additionally, obtaining an adequate number of progenitor cells for the technique is still difficult (Reviewed in Cossu and Sampaolesi, 2004). Gene therapy on the other hand

aims at replacing defective genes, for example the dystrophin gene for Duchenne's MD, or introducing a specific gene that would promote muscle regeneration (Reviewed in Chamberlain, 2002). However, the latter approach requires a very detailed knowledge of the regulatory mechanisms that govern myogenesis. In summary, development of potential therapies for MD patients requires an understanding of muscle development, adult muscle regeneration and the gene regulatory network that governs both processes.

1.2. Embryonic Muscle Development

Embryonic muscle development has been studied extensively (Reviewed in Bryson-Richardson and Currie, 2008; Buckingham, 2001; Buckingham et al., 2003). The different tissues forming the organism originate from three germ layers during embryogenesis: the endoderm, the ectoderm and the mesoderm. The paraxial mesoderm, the tissue structure located on either side of the neural tube, subsequently segments and forms structures named somites. The somites then give rise to the dermomyotome and the myotome (Brent and Tabin, 2002). Both the dermomyotome and the myotome have been found to be the origin of muscle precursor cells (Ordahl and Le Douarin, 1992). The myogenic progenitor cells then migrate to different locations in the embryo to form other anatomical structures, such as the limbs (Williams and Ordahl, 1994). The migration process is dependent on c-Met, a tyrosine receptor, and myogenic identity of the cells is conferred by a number of signalling pathways such as Sonic hedgehog (shh); Wnt and Fibroblast growth factor (FGF) (Bladt et al., 1995; Dietrich et al., 1999; Schmidt et al., 1995). Additionally, the Pax family of paired box TFs, particularly Pax3 and Pax7, have been shown to be essential for proper migration of skeletal muscle progenitor cells (Bladt et al., 1995; Bober et al., 1994; Franz et al., 1993; Goulding et al., 1994) and myogenesis (Lagha et al., 2008; Relaix et al., 2006; Relaix et al., 2003; Relaix

et al., 2005; Ridgeway and Skerjanc, 2001; Tajbakhsh et al., 1997). Of particular note, Pax3 has been shown to induce expression of Myf5 and MyoD, both are key TFs in committing the progenitor cells into a myogenic lineage (Bajard et al., 2006; Lagha et al., 2008; Tajbakhsh et al., 1997). After the muscle precursor cells have migrated to their locations, these cells proliferate to increase the muscle mass.

Proliferation of the myogenic cells, or myoblast cells, is essential to populate areas requiring muscle cells which then allows for proper differentiation. Myoblast cells are characterized by their high proliferation rate and expression of genes regulating the cell cycle (Andres and Walsh, 1996; Tomczak et al., 2004). This process ensures the survival of the organism and plays an essential role in muscle tissue regeneration and repair. The cell cycle is a highly complex and regulated process requiring a number of events to occur in an orderly fashion (Reviewed in Schafer, 1998; Vermeulen et al., 2003). Interaction between different classes of proteins, cyclins and cyclin dependent kinases (CDK), allow cells to proceed through the different phases of the cell cycle. Entry into any of the different phases is a highly regulated process and must occur in a precise order at the correct time (Nigg, 1995). Different groups of proteins must interact with each other to allow transition from one phase to the other; this includes cyclin dependent kinases, cyclins and cyclin dependent kinase inhibitors (CKI) (Elledge and Harper, 1994). However, it is necessary to terminate the cell division process to allow expression of myoblast genes subsequently leading to muscle cell differentiation and fusion. Therefore, genes participating in the regulation of the cell cycle are down-regulated, whereas cell cycle inhibitors are up-regulated at the onset of differentiation (Andres and Walsh, 1996; Tomczak et al., 2004). Differentiation proceeds with expression of the myogenic regulatory factors resulting in the formation of mature muscle fibers.

1.3. Skeletal Muscle Regeneration and Adult Myogenesis

Skeletal muscle is known to have the ability to regenerate after injury or degeneration (Carlson, 1973). This is possible through the presence of satellite cells that can fuse and form new myofibers (Mintz and Baker, 1967). The satellite cells are found between the basal lamina surrounding the muscle fibers (Mauro, 1961). Satellite cells remain quiescent in the basal lamina and express Pax3 and Pax7, where Pax7 is essential for cell survival (Relaix et al., 2006). Cells expressing both Pax3 and Pax7 also express Myf5, although the protein is undetectable (Beauchamp et al., 2000). Upon skeletal muscle damage or injury, satellite cells become activated, where they proliferate, increase the expression levels of Myf5, which in turn induces the expression of Myod. After induction of Myod, Pax7 expression is reduced and myoblast differentiation proceeds normally (Relaix et al., 2006; Zammit et al., 2002; Zammit et al., 2006). However, during the proliferative stage where Pax7, Pax3 and Myod are expressed, a number of cells are hypothesized to undergo asymmetric cell division (Conboy and Rando, 2002; Shinin et al., 2006; Zammit et al., 2004). This creates a population of activated satellite cells returning to the quiescent state where only Pax7 is expressed whereas, the second population of cells undergo myogenic differentiation (Conboy and Rando, 2002; Shinin et al., 2006; Zammit et al., 2004). Understanding the detailed molecular mechanisms regulating skeletal muscle regeneration is an essential step in developing therapies for MD patients.

1.4. Model Systems of Muscle Development and Regeneration

To better understand muscle regeneration and development, a number of models have been developed. The first involves studying muscle regeneration in mice. To accomplish this, it is necessary to cause injury to the muscle. Several methods are used to cause injury in mouse muscle models (Reviewed in Schiaffino and Partridge, 2008). One method is to mechanically damage the muscle through crushing or causing a contusion (Jarvinen, 1976; Minamoto et al., 1999; Stauber et al., 1990). The disadvantage of this technique is the reproducibility of the crush injury from one animal to the other, in addition to unequal regeneration for different muscle tissue types (Bassaglia and Gautron, 1995; Fink et al., 2003). Another method to cause muscle injury in mice is the use of venom myotoxins (Gutierrez and Ownby, 2003; Harris, 2003). Cardiotoxin is the most widely used method to injure mice muscles. When injected into muscle, the cardiotoxin causes membrane disruption and activation of phospholipases. This subsequently leads to the destruction of muscle cells and activates regeneration (Fletcher and Jiang, 1993; Lin Shiau et al., 1976). An important consideration when using myotoxins is the selectivity of the certain agent to damage a particular muscle fiber type (Harris et al., 1975) and the infiltration of immune cells into the damaged muscle (Tidball and Villalta, 2010). An alternative method to cause muscle injury in mice is to use barium chloride instead of cardiotoxin (Caldwell et al., 1990). This latter method results in the death of muscle fibers without damaging the basal membrane and does not cause immune cells infiltration. Therefore, barium cholride can be used as an alternate model for muscle regeneration after injury (Caldwell et al., 1990). Mouse models have the advantage of studying muscle regeneration in the context of an organism as a whole. It allows assessment of the contribution of other cell types, the basal membrane, local cytokine

concentrations, signalling molecules and other factors. Although mouse models are a useful tool in studying muscle regeneration, there are some disadvantages to these models. One of the difficulties with mouse models is the high level of variability between experiments due to the complexity of the organism. The complexity of the *in vivo* model also makes it difficult to study an individual molecular event or pathway. Additionally, genetic manipulations in mice are lengthy and more difficult to accomplish. For the purpose of systems biology studies, it is important to use simple model systems, as they reduce the complexity level during interpretation of the data.

Alternatively, one can study muscle differentiation and regeneration in isolation (*ex-vivo*). One widely used model system is the C2C12 cell line (Yaffe and Saxel, 1977). This cell line was originally established by isolating myogenic precursor cells from C3H mice where their muscles were crushed to induce regeneration and repair. Purification of the myogenic cells was obtained through serial passaging where a single clone was purified. Over time, serial passaging of the C2C12 cells caused them to spontaneously become immortal. Primary myoblasts on the other hand are myogenic precursor cells that have been freshly isolated from mice muscle and put into culture (Rando and Blau, 1994). Primary myoblasts are a pool of cells obtained by collagenase digestion that are grown before senescence and immortalization. Both of these models are very useful because they recapitulate the sequence of events necessary for muscle development and repair. However, both models do not take into account the contribution of the entire organism in the muscle differentiation and regeneration process as outlined in the *in-vivo* model. The *ex-vivo* models are a compromise between variability in mouse models and studying development of muscle in a complex organism.

In our experience, primary myoblasts have a higher tendency to differentiate spontaneously as soon as cell-cell contacts are made, even under growth conditions with high concentrations of growth factors and serum. This is in line with what others have observed (Asp et al., 2011). For these reasons and others outlined by Asp and colleagues (Asp et al., 2011), C2C12 were used in this study due to the ability to obtain homogeneous cell populations, which aids in reducing variability in genomics studies.

1.5. Regulation of Gene Expression

The process by which muscle cells differentiate into mature muscle fibers, termed skeletal myogenesis, is a highly regulated and complex process. The genes responsible for muscle differentiation must both be expressed at the correct location and time during this process. The regulation of these genes is controlled by a number of DNA sequence specific transcription factors (TF) which orchestrate the order in which genes are induced during embryonic muscle differentiation (Tajbakhsh et al., 1998; Tajbakhsh and Buckingham, 2000; Tajbakhsh et al., 1996; Tajbakhsh et al., 1997). Although Pax3 and Pax7 play a role in the specification of cells into the myogenic lineage, a different regulatory network governs the final steps of skeletal muscle differentiation. Even if muscle specific TF regulate the process, other factors play an important role.

Gene expression can be regulated at many levels. One layer of regulation occurs at the gene transcription step. For transcription to occur, RNA polymerase II is recruited to the transcriptional start site (TSS) of genes. Recruitment of the polymerase requires specialized proteins known as general transcription factors. These proteins, named TFIIA to TFIIH

(Function of each subunit reviewed in Hahn, 2004), recognize specific DNA sequences located up-stream of the TSS of genes (Cosma, 2002; Orphanides et al., 1996).

1.5.1. Basal Promoter and Basic Transcription Machinery

Three types of DNA elements direct and recruit RNA polymerase II: (1) the core or basal promoter (Nakajima et al., 1988); (2) proximal promoters and (3) promoter distal elements or often referred to as enhancers (Banerji et al., 1981; Gillies et al., 1983). The basal promoter contains several sequence elements necessary to recruit general transcription factors to the TSS. The first is the TATA box, found ~35 bp up-stream of the TSS, where a subunit of TFIID, the TATA binding protein (TBP) binds to this sequence (Dynlacht et al., 1991). Other elements of the core promoter include the TFIIB recognition element (BRE) which is found directly up-stream of the TATA and aids in recruiting TBP to the TATA box (Lagrange et al., 1998; Qureshi and Jackson, 1998); the Initiator element (Inr), located within the TSS (Smale and Baltimore, 1989) and the Downstream Promoter Element (DPE) (Kutach and Kadonaga, 2000). Both the Inr and the DPE work together and are targeted by the TBP associated factors (TAF), TAF1/2 (Chalkley and Verrijzer, 1999; Oelgeschlager et al., 1996) and TAF6/9 (Burke and Kadonaga, 1997) respectively. Both elements aid in recruitment of TFIID to the correct location on the gene promoter. All of these components of the basal promoter contribute to the recruitment, or cooperative binding, of general transcription factors and RNA polymerase II to the gene to be transcribed.

Once TFIID is bound to the basal promoter of the gene, other members of the general family of TF and RNA Polymerase II are recruited to the TSS forming the preinitiation complex (PIC) (Forget et al., 1997; Kim et al., 1997). Assembly of the PIC does not initiate

transcription. Phosphorylation on key residues of the c-terminal domain (CTD), unique to RNA Polymerase II, is needed to initiate transcription (Brickey and Greenleaf, 1995). Other RNA processes, such as RNA splicing, capping and polyadenylation have also been associated with the CTD (Hirose and Manley, 1998; McCracken et al., 1997a; McCracken et al., 1997b).

1.5.2. Role of Enhancers

Activation of genes does not only depend on the basal promoter: other DNA elements located further away from the TSS were discovered to enhance gene transcription (Banerji et al., 1981; Gillies et al., 1983). Enhancers can be located several kilobases up- or down-stream of gene promoters. Studies have shown that enhancers are involved in the recruitment of histone acetyl transferases (Shang et al., 2002) such as p300 and CREB Binding Protein (CBP) (Louie et al., 2003), known to alter chromatin structure (discussed in the next section) which subsequently aids in activation of gene expression (Brown et al., 2000).

1.5.3. Chromatin Structure and Post-Translational Modifications of Histones

The most basic unit of DNA compaction in cells involves DNA being wrapped around a nucleosome. Nucleosomes are formed by an assembly of an octamer of histones comprised of 2 of each subunit: H2A, H2B, H3 and H4 (Luger et al., 1997). Indeed, the presence of the nucleosomes impedes DNA transcription (Knezetic and Luse, 1986; Lorch et al., 1987) and adds an additional layer of regulation to gene expression. Histones, particularly the basic tail, are known to play an important role in gene expression regulation (Han and Grunstein, 1988; Kayne et al., 1988). Post-translational modifications (PTM), particularly on histone tails, such as methylation of arginine residues; methylation, acetylation, ubiquitination and ADP-

ribosylation of lysines; and phosphorylation of serine or threonine residues can impact gene transcription (Reviewed in Kouzarides, 2007; Li et al., 2007). For example, acetylation of H3, or di- or tri-methylation of H3 on lysine 4 (H3K4me2/3) are associated with active genes (euchromatin) (Bernstein et al., 2006; Kim et al., 2005), while methylation of lysine 9 or 27 of H3 (H3K9me and H3K27me) are associated with inactive or repressed genes (heterochromatin) (Barski et al., 2007; Boyer et al., 2006; Lee et al., 2006; Roh et al., 2006). Acetylation of histone causes an effect on chromatin structure. This particular PTM on histone tails neutralizes the positive charges from lysines of histones which weakens the strength of the DNA-histone interactions (Lee et al., 1993a). Therefore, accessibility of the DNA is increased allowing for gene transcription. Different histone methylation marks, on the other hand, would work as a signal for recruitment of different remodelling complexes, either activator or repressor complexes (Cheung and Lau, 2005; Fischle et al., 2003; Santos-Rosa et al., 2003; Zegerman et al., 2002). Overall, histone PTMs play an important role in gene expression and adds an additional level of complexity to gene regulation.

These PTMs are added or removed by various enzymes and complexes (Brown et al., 2000). More importantly, chromatin remodelling complexes have been found at enhancer regions of genes and location of active transcription (Heintzman et al., 2007; Visel et al., 2009). The interaction between protein complexes located at enhancers and the basal promoter have been shown to occur through looping of DNA over extended distances (Carter et al., 2002; de Laat and Grosveld, 2003). Additionally, the Mediator complex, a multi-subunit protein complex, also serve as scaffolds to link chromatin remodelers with the basic transcriptional machinery and RNA Polymerase II at the TSS (Conaway and Conaway, 2011; Kagey et al., 2010). A generally accepted model of the transcriptional complex is illustrated in Figure 1.

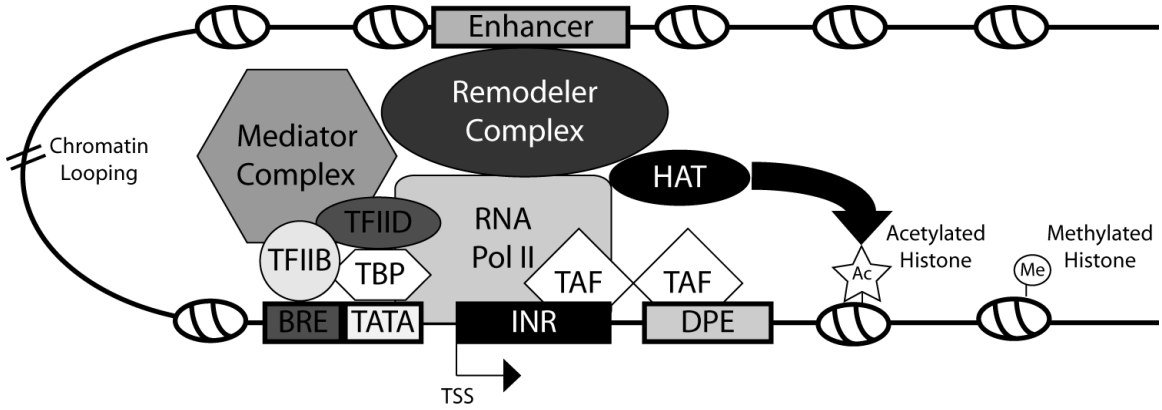

Figure 1 – The Transcriptional Complex
Schematic illustration of the major components found basal promoter of a gene near the transcriptional start site (TSS). This complex includes general transcription factors (TFIIB, TFIID and TBP) bound to the core promoter elements (BRE, TATA box, INR and DPE) with the aid of TBP associated factors (TAF). Chromatin remodelers and histone acetyltransferases (HAT) possibly bound to enhancers located further away from the TSS. Also, the mediator complex is known to help in the recruitment of general transcription factors and RNA polymerase II to the TSS. The chromatin remodelers and HAT subsequently modify the chromatin to allow transcription of the gene by RNA polymerase II. Other proteins involved in the complex have been left out to simplify the figure. (Figure inspired by Smale and Kadonaga, 2003; Szutorisz et al., 2005).

It is to note that other levels of gene expression regulation still exist, such as translational control, mRNA degradation and DNA methylation which also impact gene expression. However, they are outside the scope of this work. Up to now, the transcriptional regulatory network governing the myogenic process, specifically the key TFs involved in this process, is well characterized (discussed in the next section). The overall framework of skeletal myogenesis has been established, but the details and our understanding of the network is still incomplete.

1.6. Myogenic Regulatory Factors (MRF)

Myogenic regulatory factors (MRF) are a family of basic helix-loop-helix (bHLH) transcription factors that have been identified to play a key role in skeletal myogenesis (Reviewed in Tapscott, 2005). These TFs bind to DNA sequences named E-boxes (consensus sequence CANNTG) (Murre et al., 1989a; Murre et al., 1989b). Myogenic differentiation 1 (Myod1 or Myod) was the first gene to be identified by subtractive cloning as being able to convert fibroblasts (Davis et al., 1987) and other cell types into myoblasts (Weintraub et al., 1989). Soon after, three other family members with similar sequences and structures were identified, also with the ability to convert non-muscle cells into myogenic cells: myogenic factor 5 (Myf5), Myf6 (also known as Mrf4) and myogenin (Myog or Myf4) (Braun et al., 1989a; Braun et al., 1989b; Edmondson and Olson, 1989; Miner and Wold, 1990; Weintraub et al., 1989).

Knock-out mice were then generated to study how each member of the MRF family contributed to the regulation of skeletal myogenesis. Individual knock-out of Myod (Rudnicki et al., 1992) and Myf5 (Braun et al., 1992) surprisingly resulted in normal muscle

development with minor defects in trunk skeletal muscle in the Myf5$^{-/-}$ and delayed development of branchial arch muscle in Myod$^{-/-}$. This suggested that other members of the MRFs can compensate for the loss of an individual TF. In the compound Myf5/Myod null mice, no muscle cells are formed, indicating that both were required for myogenic differentiation (Rudnicki et al., 1993). However, concerns were raised about the contribution of Mrf4, a gene located about 8.5 kb from the Myf5 (Yoon et al., 1997). The proximity of both these genes could potentially affect the outcome of knock-out studies, since there is a possibility that both TF share a common enhancer and/or regulatory region (Carvajal et al., 2001; Hadchouel et al., 2000). Recently, a study with new Myf5 mutants, where the expression of Mrf4 was verified to be intact, showed that, in addition to being required for terminal differentiation (Zhu and Miller, 1997), Mrf4 is also necessary for early differentiation and can compensate for the combined loss of Myod and Myf5 (Kassar-Duchossoy et al., 2004). Myogenesis in the Myf5/Myod double mutants occurs normally if Mrf4 is expressed correctly. Studies performed on the Myog$^{-/-}$ mice revealed that myoblasts are still formed, but the differentiation of these cells is severely impaired and disorganized (Hasty et al., 1993; Nabeshima et al., 1993) and supports the hypothesis that Myog is required for the terminal differentiation of myoblasts. Additional work showed that Myog by itself is insufficient to induce myogenic differentiation efficiently in mice further supporting the idea that Myog is required for the terminal, rather than the initial steps of differentiation (Wang and Jaenisch, 1997). The role of Myog during postnatal life has also been characterized. Using conditional Myog null mice, researchers found that loss of Myog leads to normal skeletal muscle in adult mice, but the mice exhibit smaller body size than their wild-type counterparts (Knapp et al., 2006). Additionally, Myog has been shown to be an essential player in muscle atrophy upon muscle denervation during adult mice life (Moresi et

al., 2010). Put together, these results suggest that expression of the MRFs is crucial for proper skeletal myogenesis to occur.

The MRF family has been found to cooperate with other TFs in order to perform their function as master regulators of skeletal myogenesis. The search for other TFs revealed that MRFs cooperate with E-proteins and myocyte enhancer factor 2 (Mef2) family of TF. E-proteins are a different class of bHLH factors that can bind to E-boxes. Both MRFs and E-proteins factors can form homodimers, but only MRF and not E-proteins homodimers can convert non myogenic cells into muscle (Lassar et al., 1991; Murre et al., 1989a; Murre et al., 1989b; Zhang et al., 1999). In fact, both bHLH factors prefer to heterodimerize with each other on the E-box to activate transcription of their target genes (Blackwell and Weintraub, 1990). Mef2 TFs contain four family members, named Mef2A to Mef2D (Reviewed in Black and Olson, 1998). These proteins belong to the MCM1, Agamous, Deficiens, Serum response factor (MADS) box-containing family of TFs. Mef2 factors bind to a DNA consensus sequence ($C/TTA(A/T)_4TAG/A$), which is often found in promoters of muscle function genes (Gossett et al., 1989). Mef2 proteins have been shown to cooperate with members of the MRF family, but expression of Mef2 TFs alone is insufficient to induce muscle differentiation (Molkentin et al., 1995). Promoter analysis of muscle specific genes reveal that Mef2 sites are found in close vicinity of E-boxes further supporting the notion that Mef2 and MRFs cooperate with each other to regulate skeletal muscle differentiation (Andres et al., 1995; Blackwell and Weintraub, 1990; Blais et al., 2005; Fickett, 1996a, b; Li and Capetanaki, 1994; Molkentin et al., 1995, 1996).

Additionally, in line with the earlier discussion of the requirement of chromatin remodelers at active sites of gene transcription, these protein complexes have been shown to play an

important role in the expression of terminal muscle genes. Myod is known to recruit PCAF a histone acetyltransferase through interaction with p300 (Puri et al., 1997a; Puri et al., 1997b), and Myog has been shown to recruit RBP3, a core unit from RNA polymerase II complex (Corbi et al., 2002).

Identification of MRF targets across the genome have revealed a number of DNA TF binding sites that are specifically enriched: E-boxes; Mef2 and Mef3 sites (Blais et al., 2005; Cao et al., 2006; Cao et al., 2010). E-boxes and Mef2 sites as previously discussed allow binding of MRF and Mef2 factors respectively. Mef3 DNA elements (consensus TCAGGTTTC) on the other hand are known to be recognized by members of the Six family of TF (Spitz et al., 1998). Six1 and Six4 TFs have been shown to be required for proper skeletal myogenesis (Giordani et al., 2007; Grifone et al., 2005; Ohto et al., 1999; Ozaki et al., 2001). However, little is known about how they function and what genes they control during skeletal muscle development. Of particular interest, do the Six factors cooperate with the MRFs? And if so, what are the mechanisms which orchestrate this cooperation?

1.7. Six Family of Transcription Factors

The Six family of TFs in mice was first identified by sequence homology to the *Drosophila sine oculis* (*so*) gene (Kawakami et al., 1996a). *so* proteins are necessary for compound eye formation (Cheyette et al., 1994). Genetic analysis in *Drosophila* revealed that *so* proteins participated in a gene regulatory network involving *eyeless* (*ey*), *eyes absent* and *dachshund* (*dach*) proteins (Chen et al., 1997; Pignoni et al., 1997). Independent studies performed by Kawakami and Oliver found, through nucleotide sequence analysis, that homologues of *so* genes were also expressed in mice (Kawakami et al., 1996a; Kawakami et al., 1996b; Oliver

et al., 1995a; Oliver et al., 1995b). Further work was performed to understand the roles of these proteins in mice.

1.7.1. Characteristics of the Six Family of Transcription Factors

Identification of other *so* homologues in mice revealed a total of six genes with high sequence homology, named Six1 to Six6 (Figure 2A) (Kawakami et al., 1996b; Oliver et al., 1995a; Oliver et al., 1995b) (Reviewed in Kawakami et al., 2000; Kumar, 2009). Further characterization of these TFs found the presence of a homeodomain (HD) of about 60 amino acids which functions as the DNA binding domain (Kawakami et al., 1996a). The Six domain (SD) comprised of 110-115 amino acids has been shown to be essential in mediating protein-protein interactions with transcriptional repressors such as Groucho (Kenyon et al., 2005; Kobayashi et al., 2001; Lopez-Rios et al., 2003; Zhu et al., 2002) and Dach (Li et al., 2003; Li et al., 2002) or activators such as Eya (Ikeda et al., 2002; Ohto et al., 1999). Moreover, Six4 and Six5 contain an additional transcriptional activator domain (AD) at the C-terminus (Kawakami et al., 1996a) (Figure 2A). Six proteins have been shown to locate to the nucleus further supporting the idea that Six factors function as a TF (Kawakami et al., 1996a; Spitz et al., 1998). Careful analysis of the Six-type HD reveals that it lacks two highly conserved amino acids which are typical of all other HD. Arginine5 and Glutamine12 of helix 1 are absent from the Six HD and are replaced with Threonine, Serine or Valine (Figure 2B).

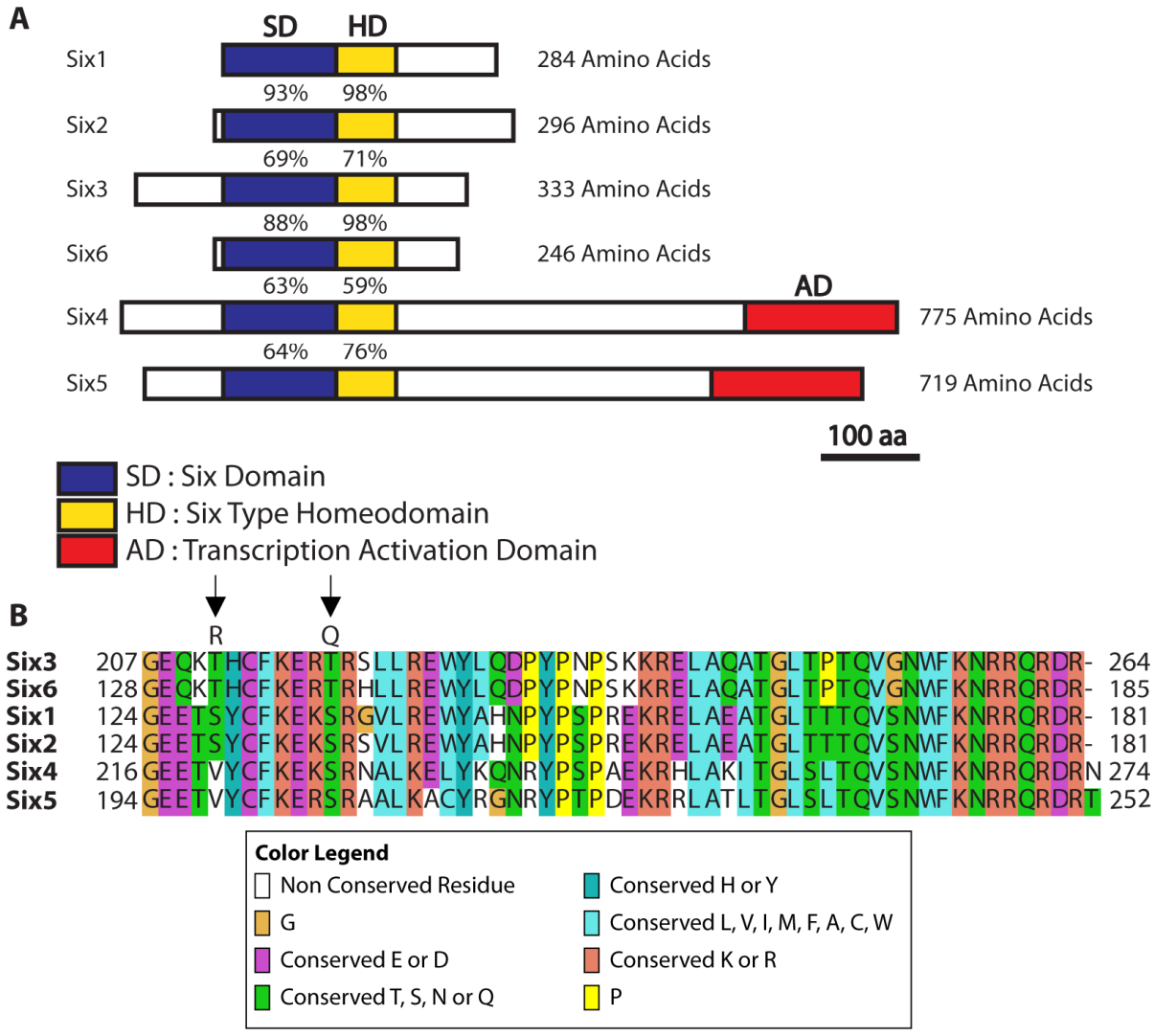

Figure 2 – Six Family of Transcription Factors.
A) The Six family of TFs contains 6 family members, named Six1 to Six6. Six factors contain a Six Domain (SD) which is unique to this family and a homeodomain (HD) needed for DNA binding. Six4 and Six5 also have an additional transcriptional activation domain (AD) at the C-terminus of the protein. Percentages represent amino acid sequence conservation of the Six family members compared to Six1. Amino acid length of proteins are also indicated to the right of the figure (Figure adapted from Kawakami et al., 2000). **B)** Six Family protein sequence alignment of the homeodomain. Six family of TF have a Threonine, Valine or Serine at position 5 and 12 instead of an Arginine and Glutamine present in the homeodomain. Corresponding amino acid residue to the respective Six protein are indicated (Figure adapted from Kawakami et al., 2000; Kumar, 2009). Sequence alignment made with ClustalX2 (Larkin et al., 2007).

This finding suggests that Six TFs may adopt a novel structure which affects its binding to the TAAT DNA consensus sequence of canonical HD proteins (Reviewed in Gehring et al., 1994; Kissinger et al., 1990). Even after the discovery of other family members and characterization of their protein domains, very little was known about their roles and functions as TFs in mice.

1.7.2. Expression of Six Factors in Mouse Embryos

Following the characterization of the Six TFs, studies were carried out to identify the location and the timing of expression of Six factors in order to better understand their roles in development (Reviewed in Kawakami et al., 2000). Six3 and Six6 are expressed exclusively in the cranial area such as the optic vesicle, ventral forebrain and Rathke's pouch (Jean et al., 1999; Oliver et al., 1995a). Six2 has a broader expression. It is expressed in the head mesoderm, the branchial arches, the fore and midgut, the nephrotomes, the genital eminence and the limb mesenchyme (Oliver et al., 1995b). Six1 and Six4 on the other hand are co-expressed in a wider number of tissues during development, including Otic vesicles, nasal placodes, branchial arches, Rathke's pouch, nephrogenic cords, limb mesenchyme, dorsal root ganglia and somites (Ohto et al., 1998; Oliver et al., 1995b).

1.7.3. Six Factors and Myogenesis

The first indication that Six TFs were involved in skeletal myogenesis was the discovery of Six4 binding to the Atp1a1 Regulatory Element (ARE) (Kawakami et al., 1996a). This gene codes for the Na^+,K^+-ATPase α1 subunit, which is responsible for maintaining cellular homeostasis. The sodium and potassium adenosine triphosphate (ATP) dependent pump is expressed specifically in the brain, lung, heart, kidney and skeletal muscle (Orlowski and

Lingrel, 1988). Kawakami and colleagues therefore performed a northern screening of mouse tissues and discovered that Six4 was expressed at high levels in skeletal muscle and in very minute amounts in heart muscle, lungs and kidneys (Kawakami et al., 1996a). Also, *in situ* hybridization studies of Six1 and Six4 reveal that both transcription factors are expressed in the somites, the source of myogenic precursor cells (Ohto et al., 1998; Oliver et al., 1995b). Additionally, ectopic expression of Six1 with Eya2 in chick embryos induced the expression of terminal muscle differentiation markers, such as myogenin, Myod and myosin heavy chain, in somite explants (Heanue et al., 1999). In summary, the results suggest that Six1 and Six4 play an important role in regulating gene expression in muscle.

1.7.4. Six1 and Six4 Knock-Out Mice

Mice that are homozygous null for Six1 and Six4 were then generated to better understand the role of these transcription factors in muscle development. The generation of $Six1^{-/-}$ knock-out mice has revealed that it plays an important role in skeletal myogenesis. The first observation was that $Six1^{-/-}$ mice died shortly after birth with a reduced muscle mass. Careful examination of the transgenic mice revealed that the diaphragm was devoid of muscle resulting in the death of the mice due to respiratory failure. $Six1^{-/-}$ mice also have rib and sternum malformations, including truncation of the distal ribs and failure of certain ribs to attach to the sternum. The loss of Six1 caused muscle hypoplasia due to impairment of primary myogenesis. This is evidenced by disorganization and reduced number of the primary myofibers, which serves as a scaffold for the formation of secondary myofibers. Examination of the cause of primary fiber disorganization showed that the phenotype was not caused by impaired delamination or migration of muscle precursors from the somites to the limbs. Inspection of early specification markers characterizing myogenic cells was found

to be unaffected. The authors of the study found that only the later steps of myogenesis were impaired by the Six1 knock-out, but Myf5 and Myog were correctly expressed; although their expression was delayed during development (Laclef et al., 2003).

The $Six4^{-/-}$ mice on the other hand are viable and show no morphological abnormalities (Grifone et al., 2005; Ozaki et al., 2001). Terminal differentiated muscle genes, such as Atp1a1, a sodium/potassium-transporting ATPase, are expressed normally (Grifone et al., 2005; Ozaki et al., 2001). This finding, reminiscent of the situation with the Myod- or Myf5-null animals, suggested that other Six family members may compensate for the loss of Six4. A likely candidate could be Six1 since it is expressed in the same location and has a similar DNA binding preference as Six4 (Kawakami et al., 1996b; Spitz et al., 1998).

Indeed, it was found that the $Six1^{-/-}$ $Six4^{-/-}$ compound knock-out mice displayed a more severe phenotype than the single $Six1^{-/-}$. Muscle hypoplasia and rib defects still occur as in the $Six1^{-/-}$ mice, but the defects in the double knock-out are more severe. The distal forelimbs and hindlimbs are completely absent from the $Six1^{-/-}$ $Six4^{-/-}$ knock-out, compared to the $Six1^{-/-}$ mice, where the same muscle groups are still formed, but with a reduced size (Laclef et al., 2003). The expression of key factors for migration such as Pax3, Met and Lbx1 were reduced, in contrast to the $Six1^{-/-}$ where these factors are still expressed and migration is normal (Laclef et al., 2003). Reduced expression of these factors impairs the migration of the myogenic precursors from the dermomyotome to the limbs. The expression of the MRFs family (Myog, Mrf4, Myf5 and Myod) is completely abolished (Giordani et al., 2007; Grifone et al., 2005).

Different studies have been performed in an attempt to identify gene targets of Six1 and Six4. Both Six1 and Six4 have been shown to directly regulate Atp2a1 and Tnnt3, genes found in fast-type muscle fibers (Niro et al., 2010) and Slc12a involved with kidney development (Ando et al., 2005). Gene targets unique to Six1 have also been identified, for example Grem1, gene essential in ureteric bud development (Nie et al., 2011). Additionally, Six1 and Eya1 have been found to regulate the expression Fgf8, required for cardiomyogenesis (Guo et al., 2011), c-myc and Gdnf, both required for organogenesis (Li et al., 2003). These studies show that Six1 and Six4 can directly regulate different genes involved in development. However, being TFs, both Six1 and Six4 can have indirect targets. For example, Six1 and Six4 may directly regulate other transcription factors such as Myod, Myog and Myf5, which in turn regulate other genes. The loss of either or both TFs can cause down-regulation of other genes indirectly.

Despite the requirement of Six TFs for skeletal muscle development, very little is known about the exact roles of Six1 and Six4 TFs in myogenesis. More importantly, how the Six TFs work in conjunction with the MRFs myogenic transcriptional network is widely uncharacterized. This therefore leads to the hypothesis and aims of this thesis.

1.8. Hypothesis and Aims

Given that Six1 and Six4 have a broad expression pattern in the developing mouse embryo, I hypothesize that Six1 and Six4 regulate genes involved in muscle cell functions and differentiation through cooperation with the MRFs, but can also regulate genes implicated in other distinct biological pathways not obviously related to the functions mentioned above. Additionally, little is known about the role of Six1 and Six4 in post-natal muscle and during muscle regeneration. For these reasons, the C2C12 (Yaffe and Saxel, 1977) cell line is used in the course of this work, since these cells recapitulate key events of post-natal muscle myogenesis and regeneration.

Aim 1: The first aim is to identify the genes and gene functions that are regulated by the Six factors. Using clustering analysis of the genes up or down regulated by the knock-down of either Six factors, it will give us clues to the roles of these transcription factors during muscle differentiation.

Aim 2: The second aim of my project is to functionally validate the role of Six factors in signalling pathways or biological roles identified in the previous aim (Aim 1).

Overall, this thesis aims at identifying the genes controlled by Six factors and understanding their function during myogenesis with the MRFs. Additionally, this project aims to provide a better understanding of their role in the myoblast differentiation process.

To accomplish these aims, the use of a number of functional genomics and systems biology techniques such as RNA interference and whole genome transcript profiling will be used. A description of the systems biology approach and these experimental techniques are described in the following sections.

1.9. The Functional Genomics and Systems Biology Approach

Over the past 10 years, systems biology has become a fast evolving field of research (Hubner et al., 2011). This field of research focuses on the understanding of a biological system in its entirety, instead of the dissection of the individual components of the system (Chong and Ray, 2002). Systems biology is therefore a holistic approach to research instead of the established paradigm which employs a reductionist approach to solve a scientific question. In fact, as Sauer explains: "*The reductionist approach has successfully identified most of the components and many of the interactions but, unfortunately, offers no convincing concepts or methods to understand how system properties emerge"* (Sauer et al., 2007). In other words, the classical approach (the reductionist approach) has provided valuable answers to biological questions by focusing on individual interactions, but this is done without considering the impact on the entire biological system. The systems biology approach is well summarized by Denis Noble: *"Systems biology […] is about putting together rather than taking apart, integration rather than reduction"* (Noble, 2006). Another difference between both research approaches is the scientific method. In the classical scientific method, a researcher would begin with (1) making an observation, (2) formulating a hypothesis to explain the observation, (3) designing experiments to test the hypothesis and (4) formulating a model or a theory based on the results. However, in systems biology, the scientific process is reversed. The researcher would start with (1) construction of a model or a theory, (2) performing appropriate experiments, whether *in silico* or wet-lab, (3) formulating a hypothesis and (4) making observations based on the new hypothesis. In summary, systems biology is a new approach to the scientific method using a holistic approach. This requires

the integration of many different experimental data sets and mathematical models which ultimately is aimed at understanding the complexity of a biological system.

The ultimate goal of the systems biology approach is to analyze the system as a whole entity and more importantly, integrate the data collected into a model. The process begins with a hypothesis and designing experiments aimed at testing the hypothesis. The proposed experiments are intended to study the interactions between the genome, the transcriptome, the proteome and the metabolome. Acquiring the experimental data requires the use of high-throughput techniques to generate many different data sets. These include: protein interaction maps or protein interactome, transcriptome analysis in a variety of conditions (i.e. gene knock-down, knock-out, drug treatment, conditional mutant, etc.), protein-DNA interactions, metabolite monitoring, etc. The analysis of these data sets individually is often insufficient to understand how these biological systems behave and function. Integration of the data collected from these studies in a computer model is therefore required to better understand how the system functions (Reviewed in Vidal, 2001). In fact, a key element of systems biology is the use of computational biology to combine the data collected into a model. From the model, predictions can be made about the system and other experiments may be designed to test these new hypotheses. Testing these new hypotheses requires the use of perturbations, or in other words, single changes to the system to confirm or refute the original hypothesis. Examples of these perturbations can include: knock-down of a gene by RNA interference (RNAi), genetic knock-out, drug treatment, over-expression or a mutation of a specific gene. Once again, after performing the experiments, data sets are analyzed and integrated into a new and improved model, where the cycle is repeated (see Figure 1 of Kitano, 2002). Due to the complex nature of the systems biology approach, hypotheses are generally broader than

hypotheses generated for the reductionist approach. In summary, systems biology has changed the philosophy of biochemistry research to consider not only the function of one gene in an isolated model, but to consider its role in the entire biological system.

As mentioned before, the aim of this work is to identify the roles of Six family of TFs during skeletal myogenesis. The hypothesis put forward was: Six1 and Six4 directly regulate genes involved in muscle cell functions and differentiation, but can also regulate genes implicated in other distinct biological pathways not obviously related to the functions mentioned above. To test this hypothesis, RNAi combined with genome wide transcript profiling by microarray and computational clustering analysis (described in the following sections) are used to knock-down the expression of Six1 and/or Six4 and reveal the identity of genes regulated by both factors. This will also lead to a better understanding of the interplay of the Six factors with the MRFs. Overall, the use of both these techniques, a combination of RNAi and genome wide transcript profiling, illustrates the importance of the systems biology approach to resolve biological complexity.

1.9.1. RNA Interference (RNAi)

RNA mediated gene silencing was originally discovered by injecting sense and antisense RNA specific to *Par1*, a gene essential for polarity in *Caenothabditis elegans*, resulting in their death (Guo and Kemphues, 1995). However, the cause was not investigated at that time. Later, other discovered that the gene knock-down was more effective and potent when using double stranded RNA versus single stranded RNA (Fire et al., 1998). Today, this discovery has been widely used as a tool to specifically knock-down expression of a gene. It has

become a rapid, easy and cost effective way to perform loss of function studies compared to the generation of a knock-out animal which is more expensive and time consuming.

Since the discovery of RNA silencing, the mechanism by which double stranded RNA (dsRNA) causes gene silencing has been described (Reviewed in Hannon, 2002; Meister and Tuschl, 2004). First dsRNA longer than 21-22 nucleotides (nt) are initially processed by Dicer, a RNase III ribonuclease containing 2 catalytic domains, a helicase and a PAZ motif (Bernstein et al., 2001), cleaves the dsRNA molecule into 21-22 nt RNA molecules with 5' phosphorylated overhangs (Elbashir et al., 2001; Zamore et al., 2000). These short dsRNA molecules are then incorporated into a nuclease named RNA-induced silencing complex (RISC) (Hammond et al., 2000) which is subsequently guided by the siRNA strand to target complementary mRNA sequence (Martinez et al., 2002; Schwarz et al., 2002). This leads to the degradation or translational inhibition of the target mRNA resulting in the silencing of the selected gene (Kennerdell and Carthew, 1998; Tuschl et al., 1999). In fact, the exact mechanism by which a gene is silenced, whether by RNA degradation (Guo et al., 2010) or translational inhibition (Guang et al., 2010; Selbach et al., 2008), is still debated. An important consideration is whether siRNA, shRNA, endogenous siRNA and microRNA (miRNA) all work with the same mechanism. Some of the findings or paradigms may apply to some but not other types of interference. A schematic summarizing the mechanism of siRNA is found in Figure 3. The identification of the key players in the siRNA pathway has provided scientists with a novel tool to perform loss of function studies.

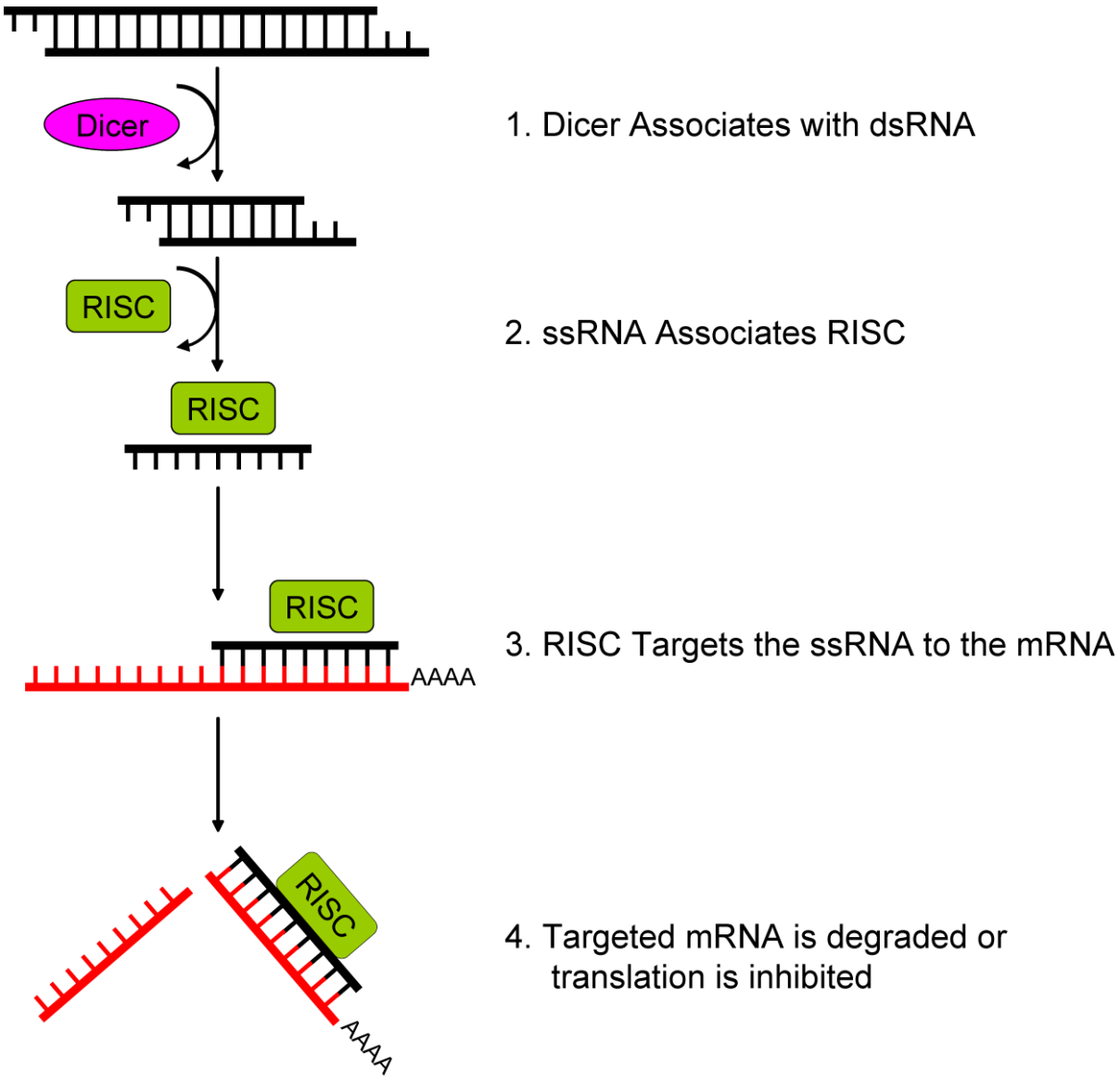

Figure 3 – Mechanism of RNA Interference (RNAi).
This schematic illustrates the mechanism by which small interfering RNA silences genes. dsRNA molecules longer than 21-22 nucleotides are first processed by Dicer, resulting in 21-22 nucleotides RNA molecules. The double stranded RNA is then associated with the RNA Induced Silencing Complex (RISC). Once association is complete, RISC targets the mRNA with the complementary RNA sequence. This leads to mRNA translation inhibition or degradation of the mRNA. This leads to lower levels of the cognate protein and/or of the mRNA itself. (Figure adapted from Meister and Tuschl, 2004)

Different RNA interference (RNAi) molecules have been developed to knock down the expression of a target gene (Reviewed in Amarzguioui et al., 2005; Rao et al., 2009). The most common RNAi technique makes use of chemically synthesized 21-22 nt dsRNA molecules with 5' phosphorylated overhangs often referred to as small interfering RNA (siRNA). The siRNA duplexes can then be transfected into cells to induce knock-down of a target gene. This method is directly processed by RISC without the need of Dicer to cleave the RNA molecules. This approach is relatively simple, siRNA sequences can be synthesized quickly and transfection is straightforward. However, the resulting knock-down is dependent on transfection efficiency and it is only transient. Monitoring of siRNA transfection by Fluorescence Resonance Energy Transfer (FRET also called Förster resonance energy transfer) showed that less than 1% of the siRNA sequences introduced in the cells remained intact after 48h leading to loss of silencing and re-expression of the target gene (Jarve et al., 2007).

Today, RNAi has become an indispensible tool used for loss of function studies. There are, however, limitations to the technology. Despite RNAi being sequence specific, there have been reports of non specific off-target effects from RNAi mediated knock-down (Reviewed in Rao et al., 2009). These off target effects are indirect effects of the exogenous RNAi molecule. For example, the use of short hairpin RNA (shRNA) has been shown to possibly induce an interferon response and immune response due to the activation of the dsRNA-dependent kinase PKR (Bridge et al., 2003; Sledz et al., 2003). Also, introduction of a shRNA sequence can result in oversaturation of the exportin-5, a nuclear export protein, impairing transport of essential miRNAs, from the nucleus to the cytoplasm (Grimm et al.,

2006). Other off target effects may include toxicity of the nucleotide sequence or the transfection method.

The use of RNAi technology has many different applications. For example, this method can be used to knock-down the expression of a key transcription factor and monitor the changes that occur at the transcript levels to identify its putative target genes, such as in this study. This technique is therefore a powerful tool for use in systems biology studies. Although some drawbacks are associated with this genomics technique, such as possible interferon response and non-specificity of siRNA or shRNA sequences, some precautions can be taken when interpreting the results from these types of experiments. These precautions may include, verifying the expression levels of known genes involved in the immune response and confirming that they were not up-regulated with the RNAi treatment. Additionally, using a second siRNA sequence (targeting a different region of the target mRNA) and obtaining similar results rules out the off-target effects from the siRNA. Both of these verifications were used in this study to verify specificity of the knock-down.

1.9.2. Gene Expression Profiling and Microarrays

Genome wide transcript profiling by microarray has now become a common technique used in systems biology to gain insight in changes in gene expression due to the dynamic nature of the system or caused by experimental perturbations to the biological system (Recent advancement in the technology see Hoheisel, 2006).

The first report of oligonucleotide arrays was in 1995, where only 48 genes were investigated for differential expression (Schena et al., 1995). Over 15 years, the technology has been improved and now a single microarray experiment can quantify up to 40000 genes. It has

become an indispensible tool in systems biology. Today, a variety of commercial microarrays are available using different DNA immobilization techniques (Commercial platforms reviewed in Dalma-Weiszhausz et al., 2006; Fan et al., 2006; Wolber et al., 2006). In this study, Agilent arrays were used to quantify mRNA expression in RNAi treated cells.

The overall procedure to perform a genome wide transcript profiling experiment (such as on the Agilent platform) is fairly straightforward (other labelling techniques reviewed in Michael, 2006). RNA is first isolated and purified from cells that have been subjected to treatments, or coming from a time series experiment. Then the RNA is reverse transcribed into complementary DNA (cDNA) using the RNA dependent DNA polymerase: Moloney Murine Leukemia Virus Reverse Transcriptase (MMLV-RT) and a T7-oligo dT primer. Once synthesis of the cDNA is complete, the T7 RNA polymerase is used to synthesize complementary RNA (cRNA). To label the cRNA fluorescently, uridine triphosphate (UTP) coupled to Cy5 or Cy3 dyes are used in the transcription mix. cRNA is used to quantify the levels of RNA from the samples because the signal is amplified from the cDNA template. The fluorescently labelled cRNA is then purified and subsequently hybridized to the glass arrays. After hybridization, the arrays are washed and then scanned to quantify the fluorescence of the mRNA in the samples. A schematic illustrating the overall procedure is shown in Figure 4. This revolutionary technique has enabled researchers to monitor the expression of genes across the entire genome in a single experiment and has allowed them to unravel the complexities of biological systems.

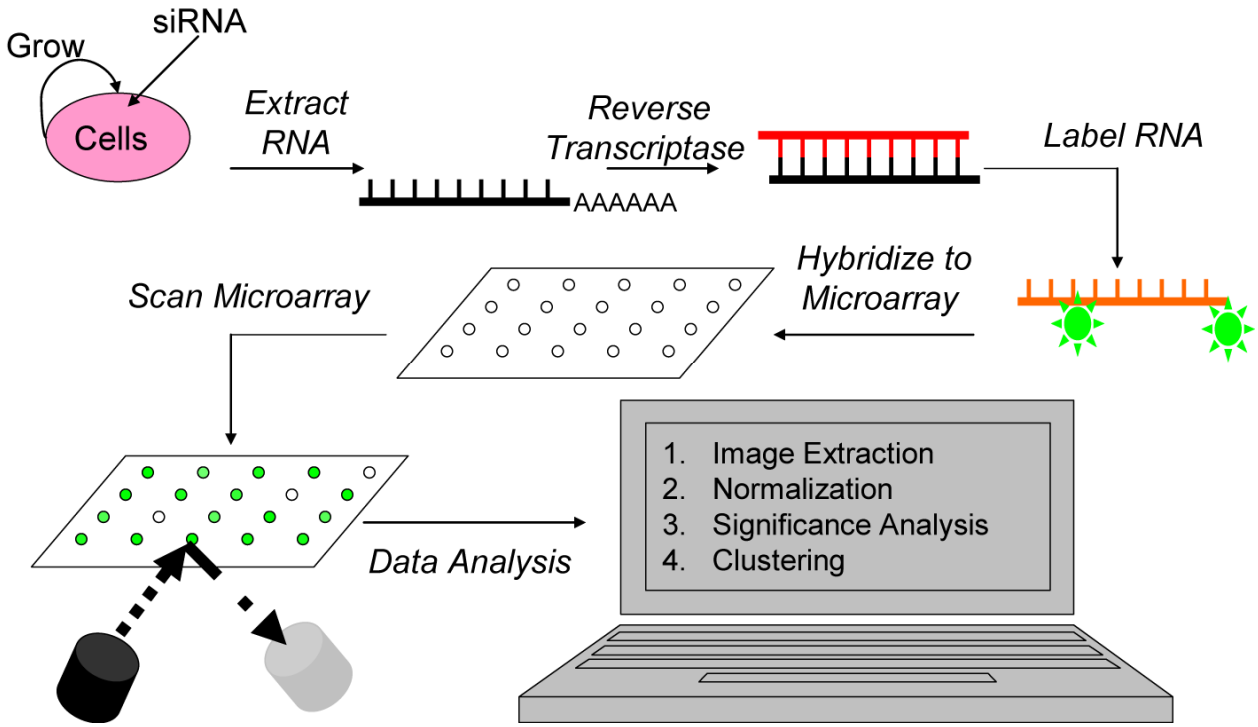

Figure 4 – Schematic of Microarray Transcript Profiling.
This figure shows a flowchart of RNA transcript profiling by microarray. RNA is first extracted from cells or animal tissues which may be treated with chemicals, signaling molecules, transfected siRNA, DNA plasmid expression vectors or infected with a wide variety of viruses. Isolated RNA is then reverse-transcribed and then labeled with a RNA polymerase with a fluorescently labeled nucleotide. The fluorescent single stranded RNA is then hybridized to a microarray chip and then scanned to quantify the amount of a specific RNA in the genome. Once the fluorescence information is collected, the raw data are log base 2 transformed, normalized and analyzed through clustering and the significance of regulated gene is assessed.

Since the development of this technology, some concerns were raised about the accuracy (how close is the measured value to the true value?), sensitivity (how accurate is the measurement at extremely low or high concentrations?), specificity (how does the immobilized DNA probe distinguish between similar mRNA sequences?) and reproducibility (can the results be replicated under the same conditions?) of the technology (Brody et al., 2002; Draghici et al., 2006; Hardiman, 2004; van Bakel and Holstege, 2004). The concerns arose when it was discovered that the technique tends to yield a high rate of false-positives (Gusnanto et al., 2007) due to limitations of statistical methods used (Brody et al., 2002; Dharmadi and Gonzalez, 2004; Nadon and Shoemaker, 2002; Rensink and Hazen, 2006; Wu, 2009). Genes identified by the analysis as changing significantly over different conditions were often false-positives, meaning changes in expression values were in reality minimal. However, the optimization and the standardization of the methods, including data processing and statistical analysis have now improved accuracy, sensitivity, specificity and reproducibility of the technique (Yauk and Berndt, 2007; Yauk et al., 2004).

1.9.3. Identification of Differentially Expressed Genes

Identification of differentially expressed genes in a microarray experiment requires pre-processing of the data beforehand (Reviewed in Geller et al., 2003; Leung and Cavalieri, 2003; Quackenbush, 2002). To explain the mathematics involved with genome wide expression profiling, consider two microarray experiments, where A is the reference sample and B is the sample treated with a chemical compound. Contained on one array, many different DNA oligonucleotide sequences specific to a gene are immobilized onto a glass slide (referred to as “gene” or “probe”) where fluorescently labelled RNA molecules (referred to as “labelled RNA”) can hybridize to. Calculation of the difference in expression

of a specific gene, often referred to as fold change, between condition A and B can be represented by the following equation: *Expression Difference or Fold Change = B/A*.

The major disadvantage of this method of representing the fold change is that the up-regulated genes and down-regulated genes are represented differently. For example, a gene up-regulated by 2 fold over the control would have a fold change of 2 (i.e. If B = 2 and A = 1, thus Fold Change = B/A = 2/1 = 2). However, if a gene was down-regulated 2 fold compared to the control, the fold change value would be 0.5 (i.e. If B = 1 and A = 2, thus Fold Change = B/A = 1/2 = 0.5). Additionally, use of raw (data that has not been log transformed) fold change ratios are often skewed or asymmetric in replicated experiments (Nadon and Shoemaker, 2002). For example, consider a triplicate experiment where the researcher compares a treatment : reference and obtained the following values: 4:3, 3:1.15 and 5:2.20 the average fold change would be 2.07 with a standard deviation of 0.66. If the inverse experiment was considered (reference : treatment) the ratios would be 3:4, 1.15:3 and 2.20:5. The average would be 0.52 with a 0.19 standard deviation. Therefore, the treatment: reference ratio is not equal to the inverse of the reference : treatment ratio ($0.52^{-1} \neq 2.07$), which is expected if the samples was reversed. To resolve this issue, the raw data is log base 2 transformed. Using a base 2 logarithm allows for a wider distribution of fold changes on a histogram than a logarithm base 10. For the example above, log base 2 transform for the treatment : reference gives 0.993 ($2^{0.993}$) and the inverse ratio gives -0.993 ($2^{-0.993}$). This makes it easy to distinguish between up-regulated and down-regulated genes, indicated by positive and negative fold changes respectively.

1.9.4. Gene Expression Data Normalization

Once the raw fluorescence values are base 2 log transformed, the expression values are normalized. This step aims at removing technical biases and systematic error between the different microarrays. These include differences in the hybridization, the labelling efficiencies, amount and purity of starting RNA. There are a number of normalization algorithms (Reviewed in Chiogna et al., 2009; Geller et al., 2003). However, only the algorithms used by the Agilent platform are discussed here. On the Agilent platform, three algorithms are available for use: 75^{th} percentile shift, scale and quantile. 75^{th} percentile normalization ranks all probes in order of intensity, from the most to the least fluorescent probe in one array independently of the other ones in the experiment. The algorithm then subtracts the intensity value at the 75^{th} percentile from all probes. The quantile method normalizes the entire data set so that in all of the arrays, the mean, median and the percentile would be identical to each other. The scale normalization is similar to the 75^{th} percentile shift; the only difference is the value of the shift is user defined.

Once the expression values (fluorescence intensities values) have been normalized, the relative fold changes in expression can be used to identify differentially expressed genes. A drawback from this method is that no statistical significance is associated with the fold change. Also, this data analysis technique is prone to systematic technical and biological variations that are inherent to microarray gene expression profiling experiments. To collect reliable data, it is imperative to obtain multiple biological replicates to increase the reliability of the data sets (Lee et al., 2000b).

1.9.5. Statistical Analysis of Gene Expression Data

Using statistical methods to detect differentially expressed genes is often the preferred approach (Reviewed in Leung and Cavalieri, 2003; Nadon and Shoemaker, 2002). Initially, to identify differentially expressed genes, statistical methods such as Student t-tests, used for pair wise comparisons, and analysis of variation (ANOVA), used for time series or for multiple pair wise comparisons, were applied to data sets. However, these computational methods often yielded high number of false positives (Gusnanto et al., 2007). This problem was attributed to certain limitations of the statistical analysis methods used. Applying statistical analysis to a microarray experiment, which contains thousands of genes, creates a multiple-hypothesis testing problem. For example, if the p_{value} was set at 5% (probability of a randomly selected gene to be not differentially expressed) and applied to a 40000 gene data set, the number of genes possibly being a false positive would be 2000 (40000 x 0.05). This is an unacceptable number of gene candidates potentially not differentially expressed. To resolve this issue, many have implemented multiple testing corrections such as the Familywise Error Rate (FWER) and the Benjamini-Hochberg algorithms (Benjamini and Hochberg, 1995; Dudoit et al., 2003). As statistical models specific for microarrays were tested, statisticians soon preferred the Benjamini-Hochberg solution to the multiple hypothesis testing problem over the FWER method. The FWER method was deemed too stringent resulting in a high false-negative rate (Nadon and Shoemaker, 2002).

Although identification of differentially expressed genes has improved through normalization and statistical analysis techniques adapted for microarrays, it is crucial to validate candidate genes by using alternative and independent techniques (Reviewed in Chuaqui et al., 2002). The first step is to measure the levels of candidate genes at the RNA

and protein level using quantitative Reverse-transcription PCR (qRT-PCR) and western blots respectively. If successful, both of these methods rules out the possibility that the differentially expressed genes were an artefact from the arrays. However, the levels of mRNA or protein may not always corroborate with each other (i.e. levels of mRNA are down-regulated but the protein levels remain constant). This can often be explained by the stability of the protein.

Another drawback to this technology is that novel transcripts cannot be detected by this method, since hybridization of probes to the arrays requires that oligonucleotide sequences be known beforehand. However, newer techniques such as whole transcript sequencing or RNA-Sequencing (RNA-Seq) are currently being used to quantitate novel transcript sequences and have been shown to be highly reproducible with little technical variations (Marioni et al., 2008).

Microarray technology has changed the way of studying gene function; but data sets are very often difficult to interpret, due to the vast amount of information they generate. Computational analysis of the transcriptome data, or clustering, is often used to interpret the data. Ultimately, these data sets are integrated with other types of experimental data sets from systems biology studies.

1.10. Clustering and Pathway Analysis

The use of these new technologies has generated a tremendous number of RNA expression profile data sets. The challenge in interpreting the data is to assign significantly modulated genes and relate them to a molecular pathway or a biological function. This difficulty is due to the large number of genes quantified simultaneously in one experiment and the complexity

of gene regulatory networks that are involved. One solution to this problem is to use computational methods, such as gene clustering, to help in the interpretation of the data.

Gene clustering of transcriptome analysis is a technique to classify genes that share similar gene expression profiles across conditions. The assumption behind gene clustering is genes having similar expression patterns across different conditions are more likely to be involved within the same biological pathway than randomly selected genes, in contrast, to genes having different expression patterns (Figure 5) (Lee et al., 2004; Stuart et al., 2003). In this case, when a TF is knocked-down, it is assumed that a group of genes that are co-regulated upon a given change are enriched and they are under the direct control of a TF. Of course, there it is also possible that certain genes within a given cluster may be under indirect control of the TF studied (e.g. knock-down of Six causes an up-regulation of a gene which normally induces a number of downstream genes). However, the list of direct and indirect targets still reveals the processes regulated by the TF under study. Other useful inferences can also be made in clustering analysis. Genes sharing the same expression profiles will likely be under the same regulatory control and therefore more likely to be regulated by a common transcription factor binding to a unique DNA sequence.

The most common clustering algorithms calculate the metric distance, which represents how close the expression of each individual gene is to each other. Then genes within a certain distance limit according to this metric are considered to form a cluster of co-expressed genes. Once the genes with similar expression patterns (i.e. shortest distance between them) are identified, they are partitioned into what are known as clusters. These groups of genes are then further analyzed by other bioinformatics tools to further uncover their function.

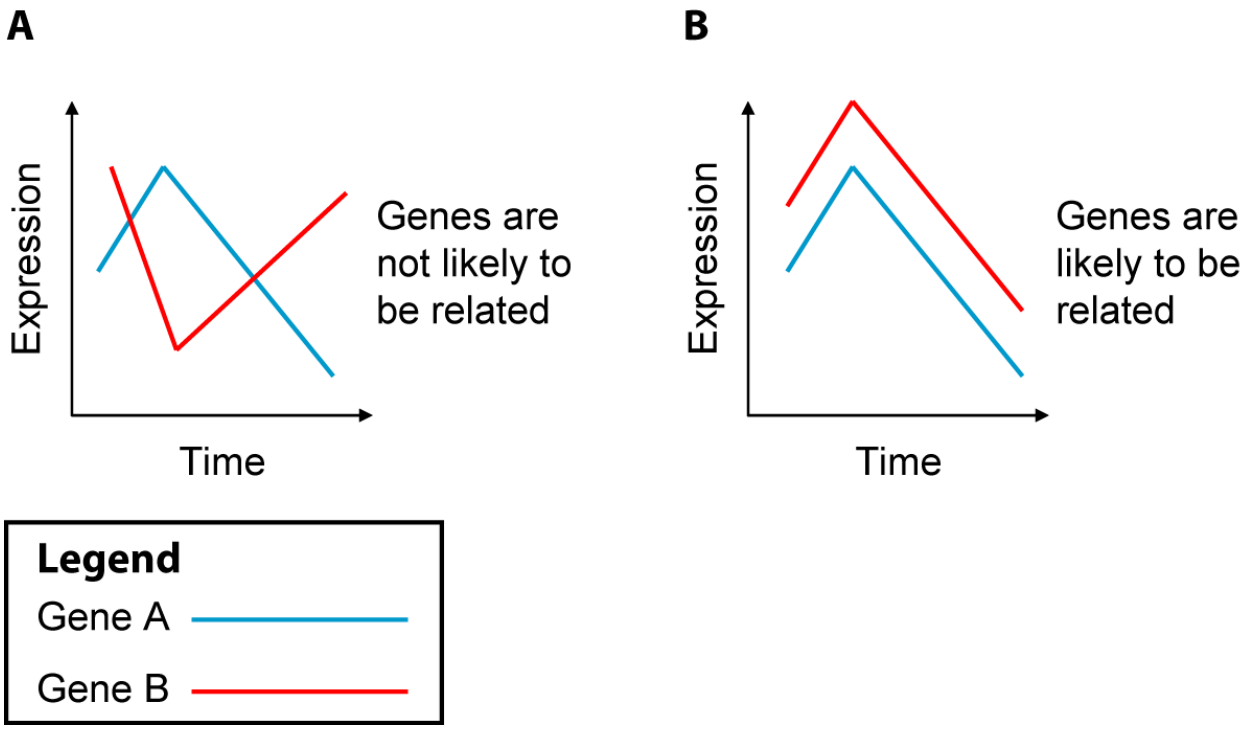

Figure 5 – Clustering of Genes.
This figure illustrates the general idea behind the gene expression data clustering. In this example, the expression 2 genes (Gene A in blue and Gene B in red) over time are represented. **A)** Both gene A and gene B have opposite expression patterns over time, thus they are assumed to be unlikely related to each other. In contrast, in **B)** both genes are expressed at the same level during the same time interval; therefore they are likely involved in the same biological pathway.

Many different clustering algorithms have been used to interpret microarray data. K-means (Hartigan and Wong, 1979; Tavazoie et al., 1999), hierarchical (de Hoon et al., 2004; Eisen et al., 1998) and self-organizing maps (SOM) (Tamayo et al., 1999) algorithms are widely used in the analysis of transcript profiling data sets (Reviewed in Chen et al., 2002; Thalamuthu et al., 2006). All of these grouping strategies, with the exception to k-means, are classified as unsupervised machine learning techniques. Therefore, no prior knowledge of the order and structure of the gene interactions are necessary in the analysis. (Extensive explanation of clustering algorithms and mathematical concepts in Everitt et al., 2011; Hartigan, 1975).

1.10.1. CLuster Identification via Connectivity Kernels (CLICK)

CLuster Identification via Connectivity Kernels (CLICK) uses a different technique to cluster data from mRNA profiling experiments. Essentially, using a graph-theoretic approach (Detailed explanation in Hartuv and Shamir, 2000), the algorithm first identifies groups of genes with highly similar expression patterns (kernels) and then by iterative process optimizes the number of kernels by merging similar kernels together (Sharan et al., 2003; Sharan and Shamir, 2000). This method, unlike some classical clustering methods such as, K-means and SOM which require the user input for the number of clusters, does not require prior knowledge of the number of clusters that exist in the data set, eliminating the need for trial and error. Partitioning of genes with highly similar expression patterns into subgroups has certainly helped in the analysis and interpretation of RNA transcriptome data. Nevertheless, the size of these genes lists resulting from these analyses may range from hundreds to thousands. It is impractical to examine each gene within a cluster individually; therefore further bioinformatics analysis is required.

1.10.2. Gene Ontology (GO)

It is assumed that genes that have similar expression patterns over time or over conditions participate in the same pathway. The reasoning behind this assumption is that a biological process requires the interaction of many genes; a single gene cannot function in a pathway alone. Hence, it is assumed that if a biological pathway is perturbed by experimentation, the genes that participate in the pathway will be found together in the same (or similar) clusters of co-expressed genes. For example, if the cells in an experimental condition have a proliferation defect (biological pathway), the genes involved in cell cycle control would be expected to be expressed at lower levels. Gene Ontology (GO) is a classification method based on known or predicted biological function, location or molecular activity of gene products (Huang da et al., 2009a, b). The GO classification is often used to determine if the co-expressed genes that compose a cluster are related. For example, if a cluster contains 15% of genes associated with cell cycle control as opposed to 2% found in the whole genome (termed the background), the category is termed enriched within that gene cluster. The Fisher's exact test statistic is used to determine whether a given enrichment level is statistically significant, considering the enrichment within the background set of genes as described in the previous example.

1.10.3. Gene Set Enrichment Analysis (GSEA)

Gene Set Enrichment Analysis (GSEA) is another bioinformatics tool used to identify genes with similar expression profiles and allowing us to search for enrichment of gene function within a microarray dataset. Unlike typical clustering methods which analyze only a subset of genes (i.e. filtered by statistical significance and fold change), GSEA instead makes use of

the entire gene expression data set and ranks all fold changes from top to bottom. This means the most down-regulated gene is ranked first and the most up-regulated gene is ranked last. The algorithm then proceeds to determine if certain groups of related genes are accumulated together near the top or the bottom of the ranked list of genes. In this case, the relatedness of genes is contained within a collection of gene sets. A few examples of gene sets could be "Positive regulation of cell cycle", containing genes such as the cyclins, "Wnt signaling pathway", comprising genes such as Wnt1, Fzd7 and beta-catenin, and "Genes up-regulated by ionizing radiation in fibroblasts", which would contain p53, p21 and Rad50. The molecular signature database (MSigDB) is a vast collection of over 3000 different gene sets, collated from various sources such as individual publication and gene expression data sets (Subramanian et al., 2005).

1.10.4. Whole Genome rVista Gene Promoter Analysis

Another method to characterize a set of genes resulting from a clustering analysis is to examine the promoters for common DNA regulatory elements. Whole Genome rVista is a web-based tool used to search for DNA regulatory elements that are enriched within a gene list (Dubchak and Ryaboy, 2006). The assumption behind this technique is that genes participating in the same biological pathway are regulated by a unique number of DNA sequence specific TFs. Therefore, if a cluster of genes have similar gene expression patterns, it is assumed that these genes will have within their promoters the DNA sequence corresponding to a common transcription factor binding site. In order to filter out non-informative sequence information, the analysis is performed only on DNA sequences that are conserved between human and mouse genomes. Here, it is assumed that if a DNA sequence

is conserved, it is therefore more likely to be functionally relevant (Elnitski et al., 2005; Frazer et al., 2004; Schwartz et al., 2003a; Schwartz et al., 2003b).

In summary, Six1 and Six4 TFs are known to be involved in skeletal myogenesis. However, their function within the myogenic transcriptional network is still unknown. This work aims at uncovering the roles of the Six factors during myogenesis through the combined use of RNAi, genome wide transcriptome analysis by microarray and computational techniques as illustrated in Figure 6.

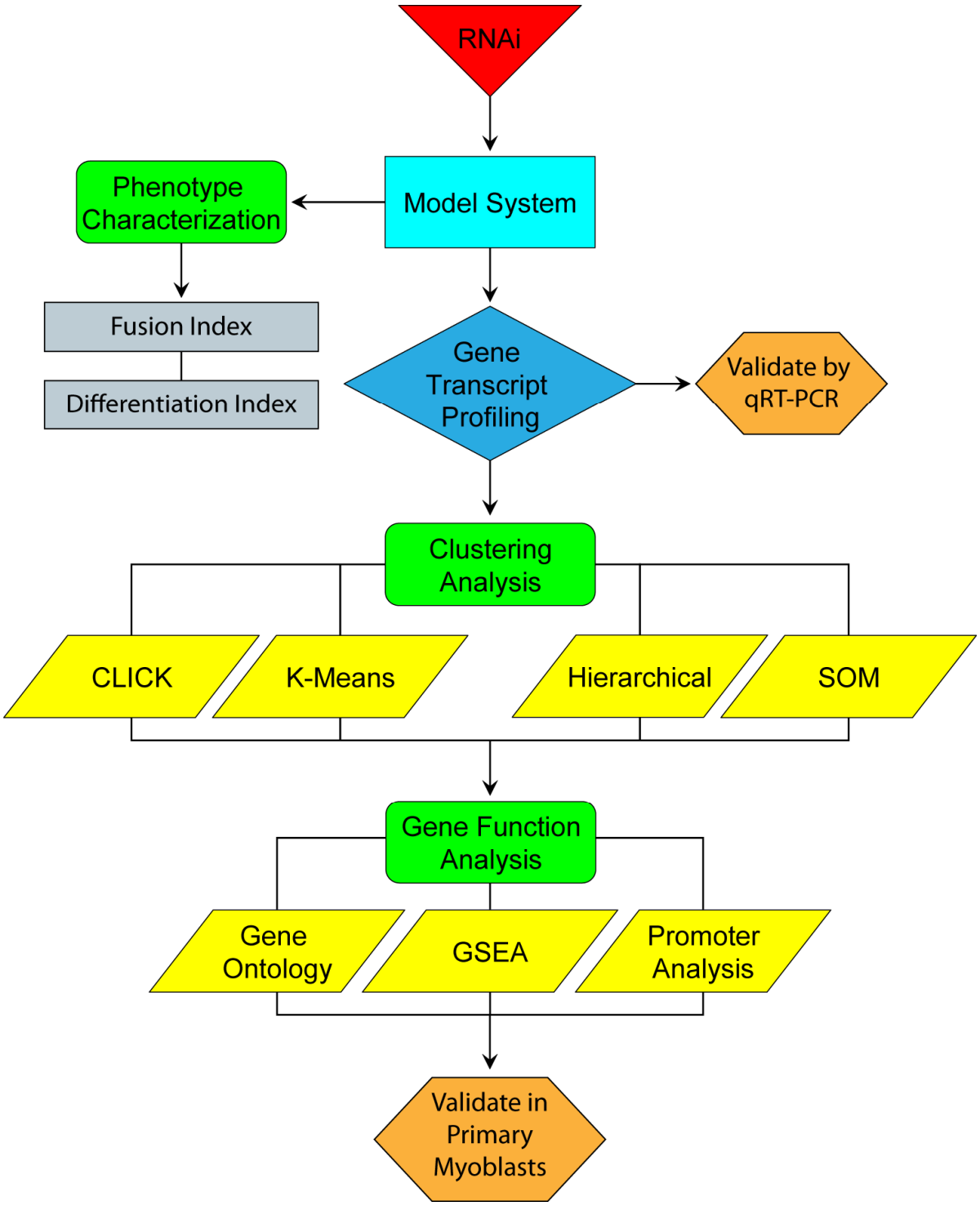

Figure 6 – Overview of the Functional Genomics Approach to Characterize the Role of Six Factors in Skeletal Myogenesis.
This flow chart illustrates the overall steps in characterizing the role of Six TFs during skeletal myogenesis by combining 3 functional genomics approach: RNAi, genome wide transcript profiling and bioinformatics.

2. Chapter 2 – Six Factors are Required for Myoblast Differentiation

2.1. Introduction

Skeletal myogenesis is a highly complex and regulated process controlled by a number of TFs. These DNA sequence specific proteins control the ordered expression of genes that are unique to muscle both in time and location. Myogenic Regulatory Factors (MRFs) (Berkes and Tapscott, 2005; Tapscott, 2005) are known to be important transcriptional regulators in myogenesis in cooperation with other factors such as the Mef2 family of TFs (Molkentin et al., 1995, 1996). The Six family of TFs have recently been shown to be required for proper muscle development (Giordani et al., 2007; Grifone et al., 2005; Ozaki et al., 2001). Up until now little is known about their function and the target genes that they regulate during myogenesis. Here, to better characterize the roles of Six1 and Six4, I used a combination of functional genomics approaches, RNAi directed against Six1 and Six4 to knock-down their expression coupled to genome wide transcript profiling and bioinformatics analyses. This work is aimed to better understand the role of Six1 and Six4 during myoblast differentiation.

2.2. Results

2.2.1. Six1 and Six4 are required for skeletal muscle differentiation

Taking into account that Six1 and the compound Six1/Six4 constitutive knock-out mice have severe defects in muscle development and that they are expressed in C2C12 (Figure 7), siRNA targeted against Six1, Six4, a mixture of both Six1 and Six4 siRNAs, or a control sequence were used to knock-down their expression in mouse myoblasts.

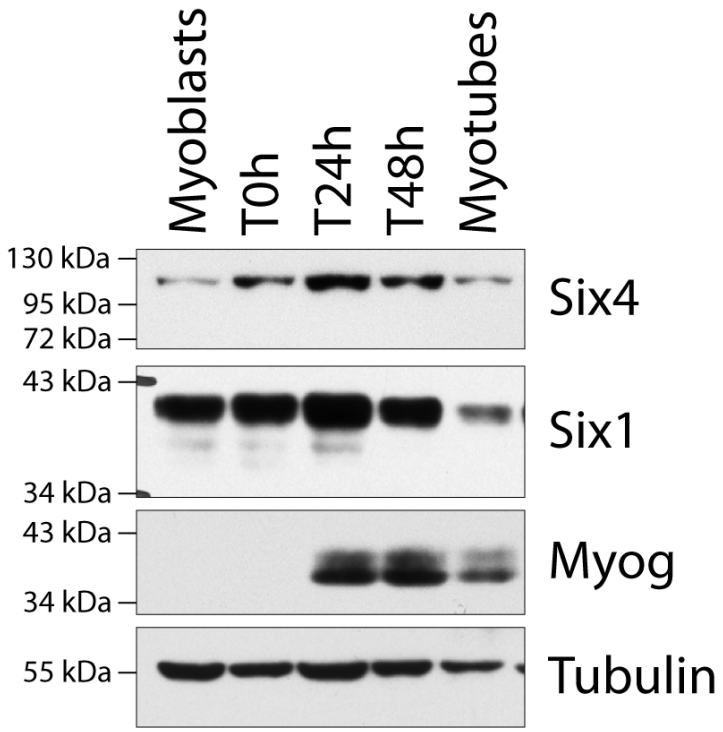

Figure 7 – Expression of Six1 and Six4 during C2C12 differentiation.
Whole cell extracts from a C2C12 time course were analyzed by western blot at different time points of C2C12 differentiation: MB – Myoblasts; T0 – Start of differentiation; T24 and T48 – 24h and 48h after the start of differentiation respectively MT – Myotubes after 96h of the start of differentiation. Myotubes were selectively harvested by use of diluted trypsin (1:20 with PBS).

The impact of these manipulations on the differentiation and fusion of myoblasts were then investigated. C2C12 myoblast cells were transfected with siRNA, allowed to proliferate to confluence and then induced to differentiate for 1 day and harvested for RNA and protein (Figure 8 A-1), or 3 days with an additional transfection 24h post differentiation before fixed for immunofluorescence to evaluate their fusion (Figure 8 A-2). Western blots performed one day after the start of differentiation indicate that knock-down of Six1 and Six4 were successful (Figure 8 B) and reproducible through densitometry quantification of five independent biological siRNA knock-down experiments (Figure 8 C). Interestingly, knock-down of Six1, Six4 or the combination of both reduced the expression of Myog, a key TF required for terminal muscle differentiation (Hasty et al., 1993; Nabeshima et al., 1993). This result was expected as it was previously reported that both Six1 and Six4 are able to activate expression of Myog by binding to the Mef3 motif upstream of the gene during embryogenesis (Spitz et al., 1998). Although regulation of Myog by Six1 has already been shown in an embryonic model, the novelty of this work is the regulation also occurs in C2C12, a model of adult muscle regeneration. In complementary experiments, siRNA-treated myoblasts were also allowed to differentiate in medium with reduced serum content for 3 days to evaluate the effect of Six1 and/or Six4 knock-down on myoblast differentiation and fusion (Figure 8 A-2). Staining for Myosin Heavy Chain (MHC), a marker of terminally differentiated muscle cells, and counterstaining with DAPI allowed for the assessment of the fusion (ratio of MHC positive-cells that have fused together into a myotube with at least two nuclei over the total number of nuclei) and differentiation (ratio of MHC-positive cells over the total number of cells) indices. The results show that myoblast differentiation and fusion indices were significantly impaired by knock-down of either or both factors (Figure 8 D and E), indicating that Six1 and Six4 are required for proper differentiation of myoblast cells.

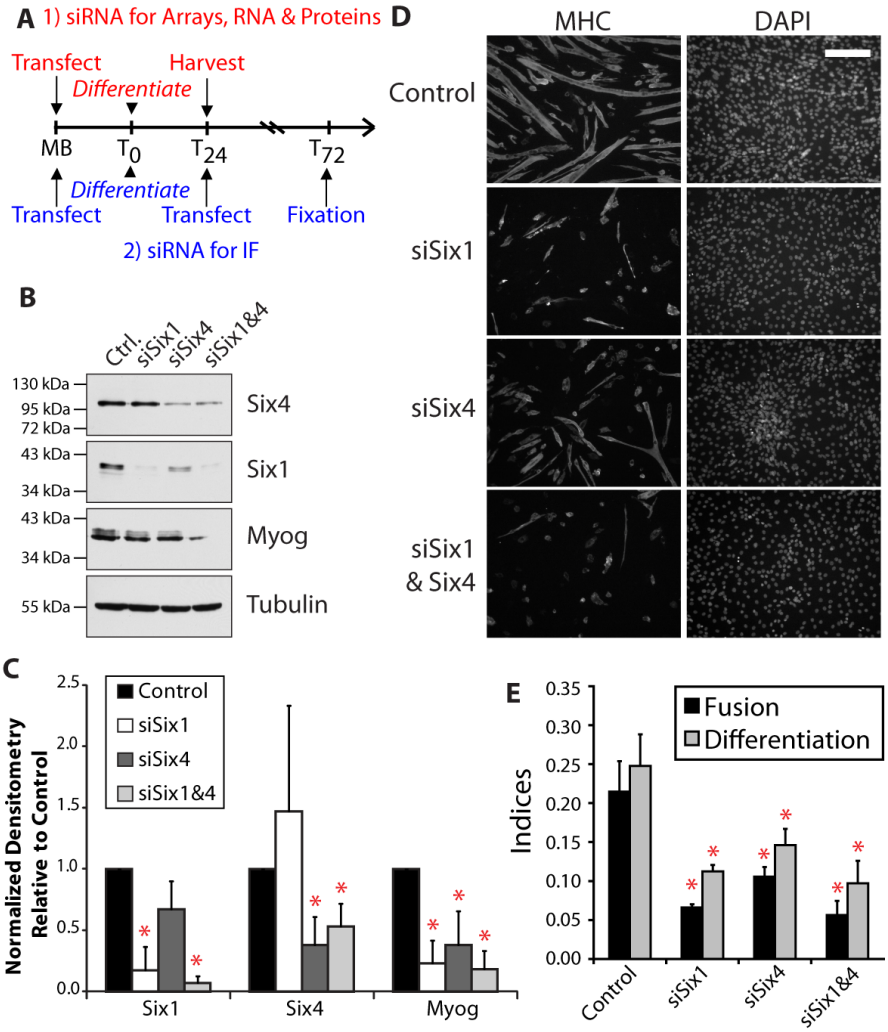

Figure 8 – Knock-down of Six1/Six4 impairs cell differentiation and fusion.
A) Timeline illustrating the time of transfection and harvest for transcript profiling (1 in Red) or fixation for immunofluorescence (2 in Blue) experiments. **B)** Western blots illustrating the protein levels of Six1, Six4 and Myog after a single treatment with siRNA 24h prior to the start of differentiation. Cells were harvested 24h after the start of differentiation. **C)** Western blot quantification of Six1, Six4 and Myog by densitometry of 5 independent siRNA transfection experiments. Measurements were normalized to the densitometry of Beta-Tubulin. Experiments performed as described in B). **D)** Representative microscopic fields from each siRNA conditions where staining of MHC, a marker for terminal muscle differentiation, shows impairment of C2C12 to fuse and differentiate. Cells were treated twice with siRNA, 24h before and after the start of differentiation. Cells were fixed for immunofluorescence at 72h after the start of differentiation. Scale bar: 200 microns. **E)** Fusion and differentiation indices of siRNA-treated C2C12 induced to differentiate for 72h showing impaired fusion and differentiation of myoblasts. Experiments performed as described in D). The data represent the average of 3 independent biological siRNA experiments. Error bars indicate the Standard Error over the Mean (SEM). * Represents statistically significant change by a two-tailed paired Student t-test ($p_{value} \leq 0.05$) compared to the control siRNA.

To rule out the possibility of off-target effects due to non-specificity of the siRNA duplex (i.e. the siRNA sequences used may target different mRNAs or genes other than Six1 and Six4) a pool of 3 different siRNA sequences (excluding the ones used in the previous experiment) were used and the knock-down resulted in comparable effects (Figure 9 A and B). Protein expression of Six1, Six4 and Myog were analyzed when C2C12 are treated with the control siRNA sequence or left untreated. The results indicate that Six and Myog expression are unaffected by control knock-down sequence (Figure 9 C). These results indicate the failure of myoblast cells to differentiate and fuse was unlikely to be caused by an off-target effect from the siRNA sequences and that both Six1 and Six4 regulate the process of muscle differentiation. The overall results show that Six factors, specifically Six1 and Six4, are both essential players in skeletal myogenesis.

An issue with RNAi technology is the possibility of inducing an interferon response, particularly with the use of vector based RNAi (Bridge et al., 2003; Sledz et al., 2003). To verify if the interferon response pathway was induced in C2C12 transfected with the different siRNA sequences, the expression levels of genes from the pathway were examined from the microarray data (Gene list from Sledz et al., 2003). Interestingly, out of 18 interferon genes, only 8 were detected as having a fluorescence signal over the background (Figure 10). Of these eight genes, none had significant expression changes Table S 2 in the Appendix). The siMyog duplex caused a notable upregulation of only one of those 18 genes. Since the possibility cannot be formally ruled out that the Myog transcription factor is involved in the regulation of this gene in myoblasts, it was concluded that none of the siRNA duplexes used induced the interferon response to a level higher than the control non-silencing siRNA duplex.

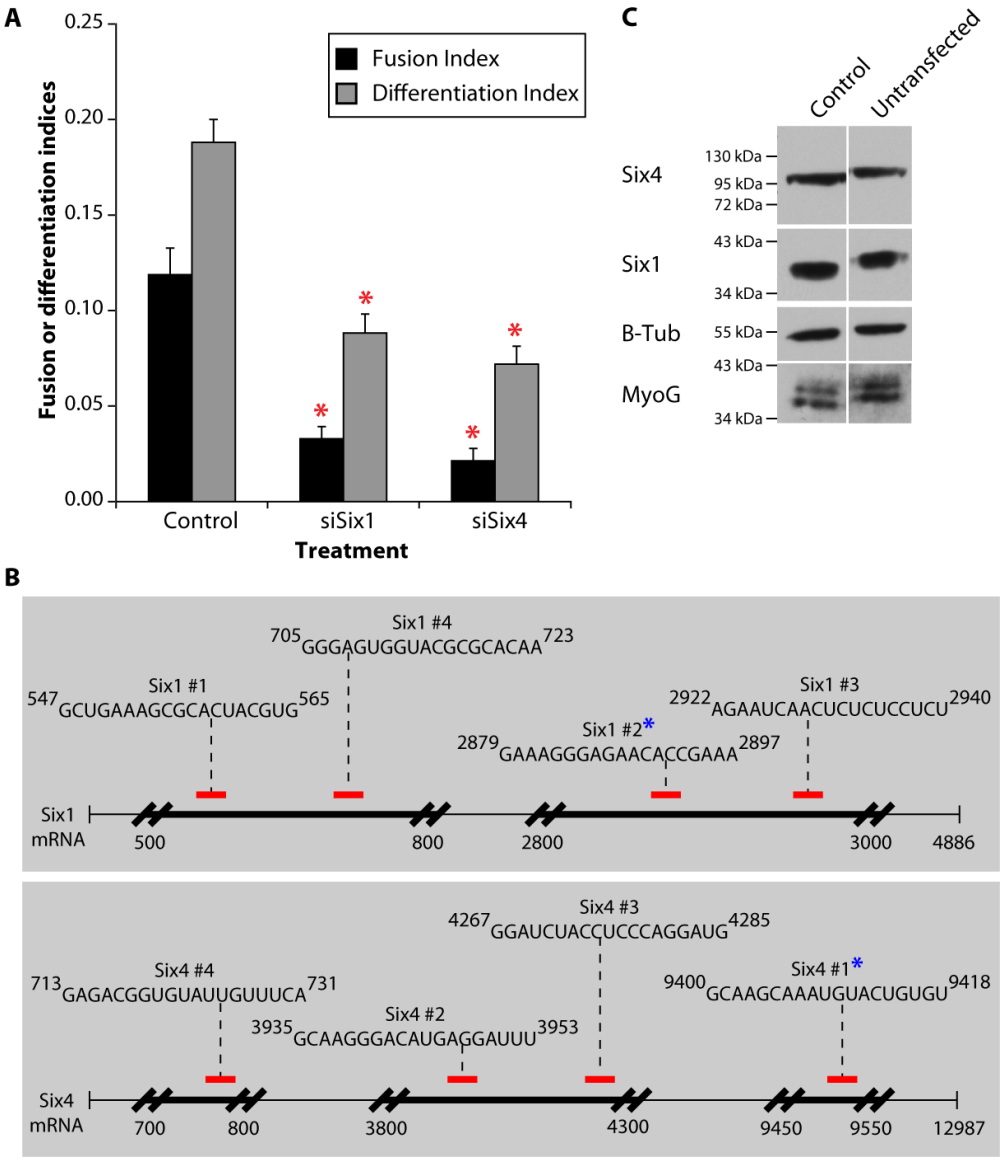

Figure 9 – Other siRNA sequences targeting Six1 and Six4 cause a comparable phenotype and the Control siRNA sequence does not affect Six1 and Six4 expression.
A) A mixture of 3 other siRNA sequences targeted to Six1 and Six4 were transfected twice into C2C12 cells, 24h prior and after start of differentiation, where the fusion and the differentiation index were scored after 72h of differentiation. The data represent the average of 3 independent experiments. Error bars indicate the Standard Error over the Mean (SEM). * Represents statistically significant change by a two-tailed paired Student t-test ($p_{value} \leq 0.05$) compared to the control siRNA. **B)** Graphic representation of siRNA sequences targets on Six1 and Six4 transcripts. * Represents single sequence used for all RNAi experiments, while the three other sequences were specifically used in this experiment (Section A of this figure) to verify specificity of siRNA sequences. siRNA binding positions are not to scale. **C)** Western blot of C2C12 cells were transfected once with siRNA 24h prior to differentiation control sequence or left untreated. The cells were allowed to differentiate for 24h and harvested of protein analysis.

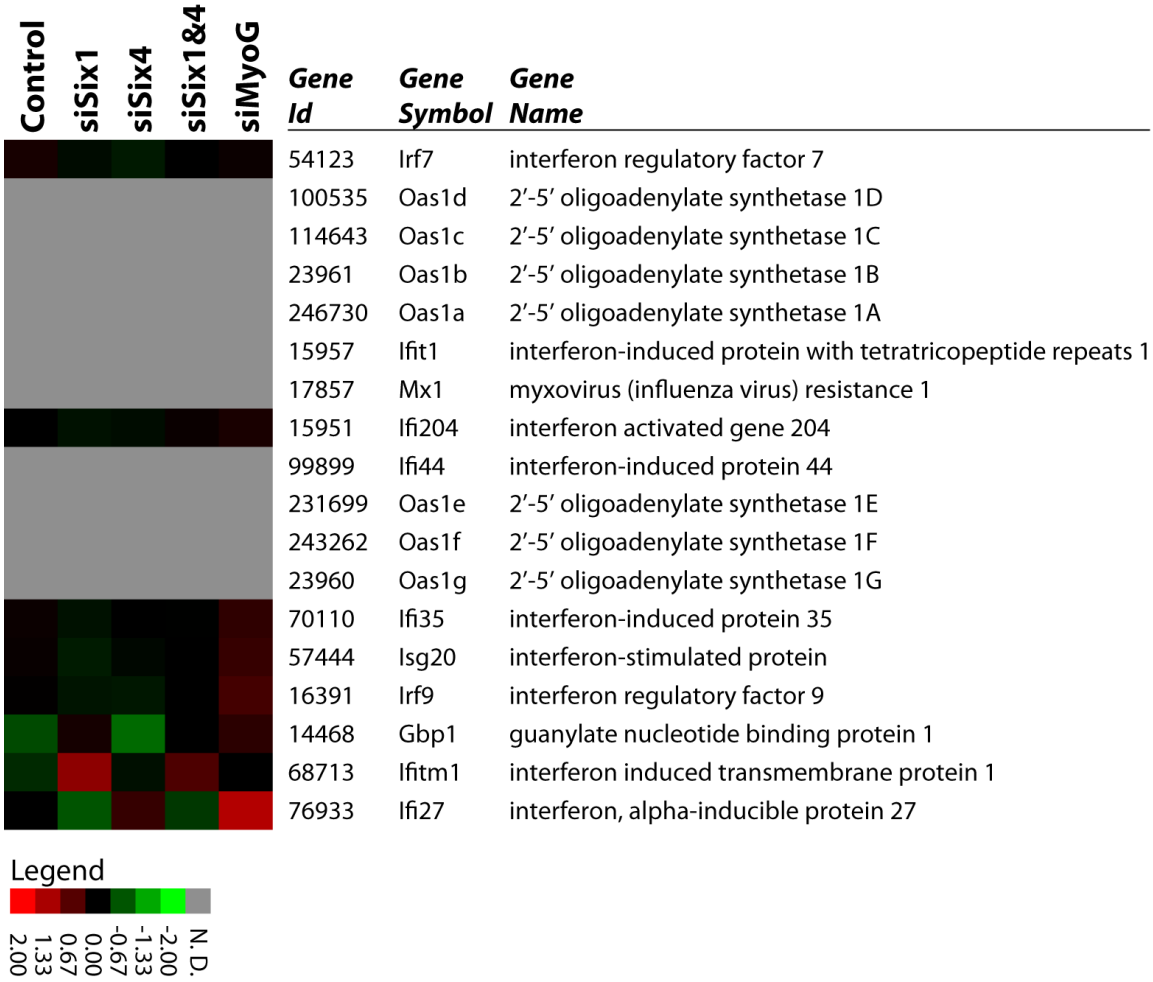

Figure 10 – Expression levels of interferon genes when C2C12 are treated with siRNA against Six1, Six4, Myog and a control non-specific sequence.
Heat-map representing the expression levels of interferon genes (list taken from Sledz et al., 2003) in the different RNAi mediated knock-down of Six1, Six4, Six1 and Six4, Myog and a control non-specific siRNA sequence. Color scale represents the normalized expression values in linear scale. Corresponding Entrez Gene Id, Gene Symbol and Gene names are indicated. N.D. Probed Not Detected. Normalized expression values can be found in Table S 2 in the Appendix.

2.2.2. Genome wide transcript profiling

To better establish the role of the Six proteins in skeletal muscle development, mRNA from the knock-down experiments were subjected to genome wide transcript profiling by microarray. The arrays comprise about 41000 DNA probes chemically immobilized to the glass array representing about 31200 unique genes. Four independent biological siRNA treated C2C12 replicates were Cyanine 3-CTP labelled using the Agilent Quick Amp labelling kit. To make the different microarray experiments comparable to each other, the quantiles normalization algorithm was applied to the entire data set. About 22800 were classified as "present" or meaning that they have fluorescence values above background in all four replicates of any of the 4 conditions. 6924 probes remained after filtering for statistical significance by ANOVA with multiple hypothesis testing correction using the Benjamini-Hochberg algorithm (Benjamini and Hochberg, 1995). A difference of expression of at least 1 in log base 2 (representing a 2 fold-change on a linear scale), either up or down-regulation was then applied to the statistically significant probes where only 1958 remained. These remaining probes were then analyzed by bioinformatics tools to search for groups of genes with similar expression patterns. Traditional clustering algorithms such as K-means, Hierarchical and SOMs were tested. However, the results from these algorithms were often difficult to interpret due to two main reasons: (1) they either did not look sufficiently homogeneous, or (2) they were sparsely populated (data not shown). The CLICK algorithm from the Expander software (Sharan et al., 2003; Sharan and Shamir, 2000; Ulitsky et al., 2010) was used to classify the 1958 genes. The algorithm was successfully able to classify the significant 2 fold change genes into 8 different clusters (Figure 11). Of these eight, Clusters #1 and #3 were of particular interest.

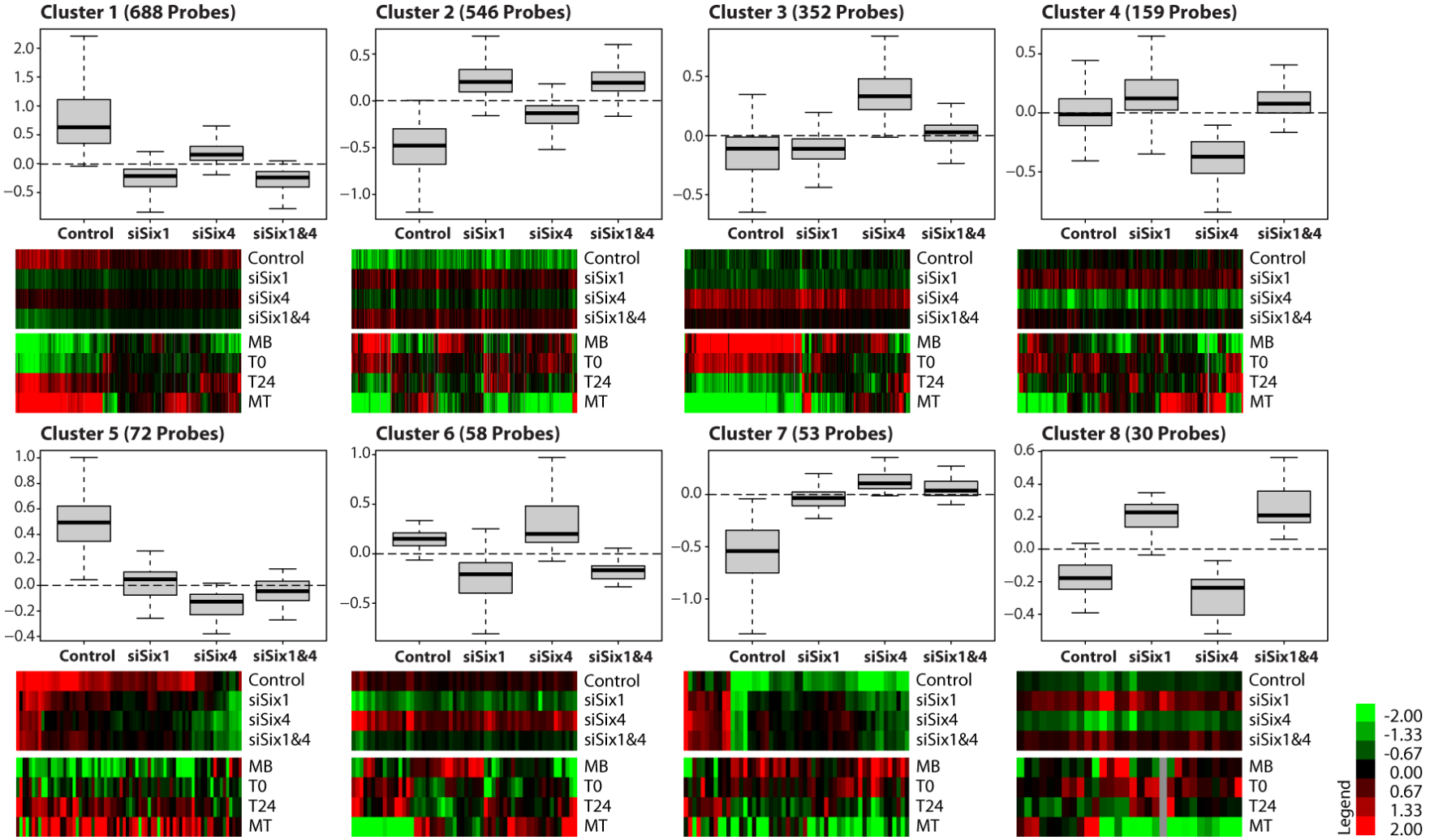

Figure 11 – CLICK Clustering Analysis of Transcript Profiling of C2C12 treated with siRNA 24h after differentiation. Genes that had an ANOVA < 0.05 by Benjamini-Hochberg correction and a fold change of above 2.0 (1958 probes) was submitted to CLICK Clustering Analysis and resulted in 8 clusters. Y-Axis of box plots represents the normalized expression values (Quantiles Normalization where in any given condition, the average of the entire data set is 0 with a standard deviation of 1). Whiskers represent 1.5 x Inner Quantile Range of the 1st and 3rd quantiles. The number of genes found within each cluster is indicated. The expression values of each cluster are also shown in the Heat Map format (NS; siSix1; siSix4 and siSix1&4). The corresponding genes were also aligned to the C2C12 differentiation time course (MB, T0, T24 and MT). N.D. Probe Not Detected.

Cluster #1 was interesting because it contained 688 probes that have a severe decrease in their expression when Six1 or both Six factors are knocked-down and a mild decline in mRNA levels when Six4 is knocked-down (Figure 12 A, B and Table S 3).

To better understand the functions of the genes regulated by Six1 and Six4 during normal C2C12 differentiation, whole genome transcript profiling was performed at four different time points during C2C12 differentiation. Total RNA was harvested at the myoblast (MB) stage; at the start of differentiation (T_0); 24 hours after the start of differentiation and differentiated myotubes (MT). Three independent biological time course experiments were used to perform gene expression analysis. The same labelling protocol was used and the same normalization algorithms were subsequently applied. The differentiation time course data were then aligned with the knock-down microarray data from Cluster #1 data. About half of the genes in Cluster #1 are expressed at low levels in myoblast but are up-regulated in differentiated cells (Figure 12 C), suggesting that these genes are required in differentiated cells or are necessary for the differentiation process to occur. Representative genes candidates that are up-regulated during C2C12 differentiation time course were analyzed by qRT-PCR and were confirmed to be expressed at high level during the later stages of differentiation (Figure 13). These results would strongly indicate that both Six1 and Six4 control a set of muscle function and differentiation genes. However, further analysis of the data set is required to corroborate this finding.

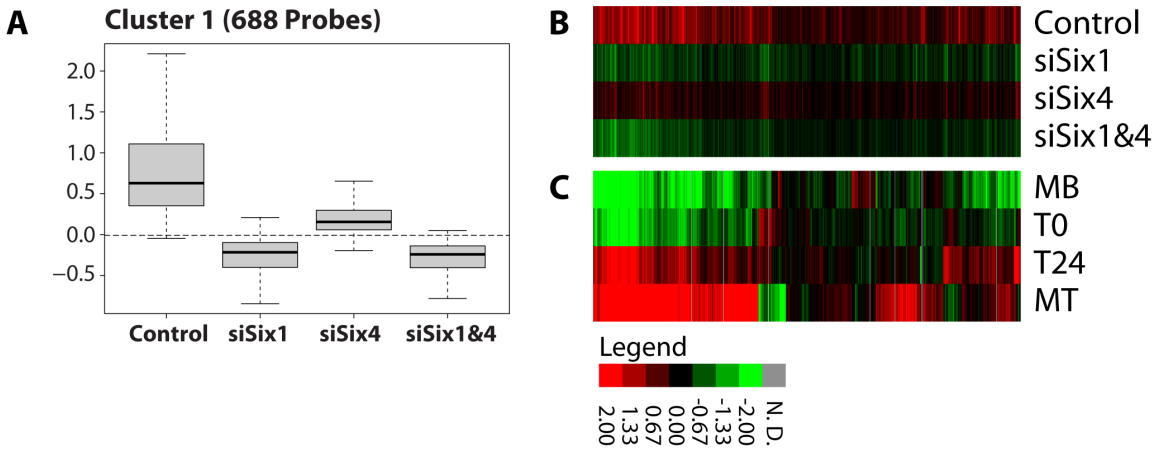

Figure 12 – Six1 and Six4 regulate genes that are normally induced during muscle differentiation – CLICK Cluster #1.

A) Box-plot representing the average expression profile of 688 probes over 4 siRNA treatments in C2C12 (siRNA against Six1, Six4, both or a non-silencing control sequence) grouped into Cluster #1. Whiskers represent 1.5 x Inner Quantile Range of the 1st and 3rd quantiles. For the full gene list see Table S 3 located in the Appendix. **B)** Heat-map representation of the 688 probes from Cluster #1 where Six1 and/or Six4 are knocked-down by RNAi. **C)** Probes from the C2C12 differentiation mRNA profiling experiment at four different time points (MB, T_0, T_{24} and MT) were aligned with the 688 probes from Cluster #1 to identify the expression trends of these genes during C2C12 differentiation. Color legend is given under both heat-maps and represents normalized expression values in linear scale.

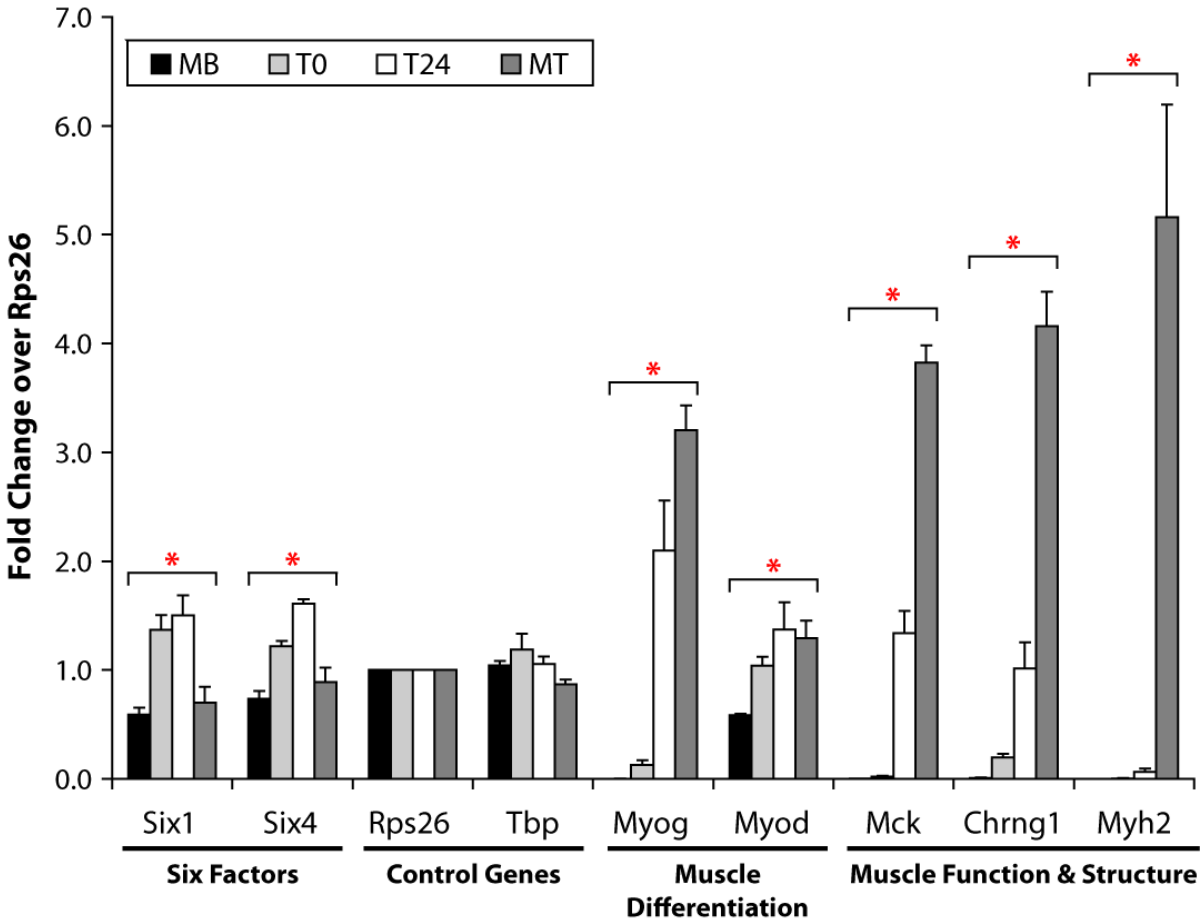

Figure 13 – Muscle differentiation and function genes are normally up-regulated during C2C12 differentiation.

Levels of mRNA during C2C12 differentiation at 4 different time points: Myoblasts (MB), 0h of differentiation (T0), 24h of differentiation (T24) and Myotubes (MT) were quantified by real-time PCR (qRT-PCR) on a number of muscle differentiation and muscle function genes. mRNA levels were normalized to the Rps26 gene as an invariant control and the TATA Binding Protein (Tbp) was used to assess the quality of the Rps26 control gene. The data represent the average of 3 independent biological time course experiments. Error bars indicate the Standard Error over the Mean (SEM). * Represents statistically significant change over the time course by single factor ANOVA analysis.

2.2.3. Genes regulated by Six1 participate in muscle function, differentiation and development

To further characterize the genes found within Cluster #1, the 688 genes were submitted to Gene Ontology (GO) analysis (Huang da et al., 2009a, b). This allowed the identification of the biological processes that are over-represented among the gene set. The bioinformatics analysis confirmed that genes constituting the cluster, where genes were down-regulated in the Six1 knock-down, participate in muscle cell development, function and differentiation above random expectations (Table 1) (See Table S 5 in the Appendix for the full list). A number of representative genes were tested by qRT-PCR and confirmed their down-regulation by the knock-down of Six1 and both Six1 and Six4 (Figure 14). As expected, the siRNA treatments caused a statistically significant decrease in mRNA expression levels of genes that play an essential role in muscle differentiation (Six1, Six4, Myog and Myod), genes implicated in muscle structure and function (Mck and Actn3), cholinergic receptors (Chrna1 and Chrng), calcium channels (Atp2a1 and Cacng1), and structural genes of muscle (Actn3 and Lbd3) (Figure 14). These results show that Six1 and Six1/Six4 plays a role in the regulation of muscle differentiation and muscle function genes.

GO Term	Count	Enrichment	p-Value[1]	Genes
Muscle development	23	4.81	5.67E-06	Myog, Myh1, Myod1, Six1
Muscle fiber development	13	7.09	1.48E-04	Chrna1, Myog, Ttn, Myod1, Six1
Muscle cell differentiation	11	5.76	8.44E-03	Ttn, Myod1, Tmod1, Cacna1s

Table 1 – Six1 and Six4 regulate muscle function and development genes.
Genes found in Cluster #1 were analyzed by Gene Ontology Analysis (GO). A subset of GO categories is shown in the table. For the full list see Table S 5 in the Appendix [1] p-Values were corrected by multiple hypothesis testing by Benjamini-Hochberg algorithm. The background set of genes used in the significance values was the entire mouse genome.

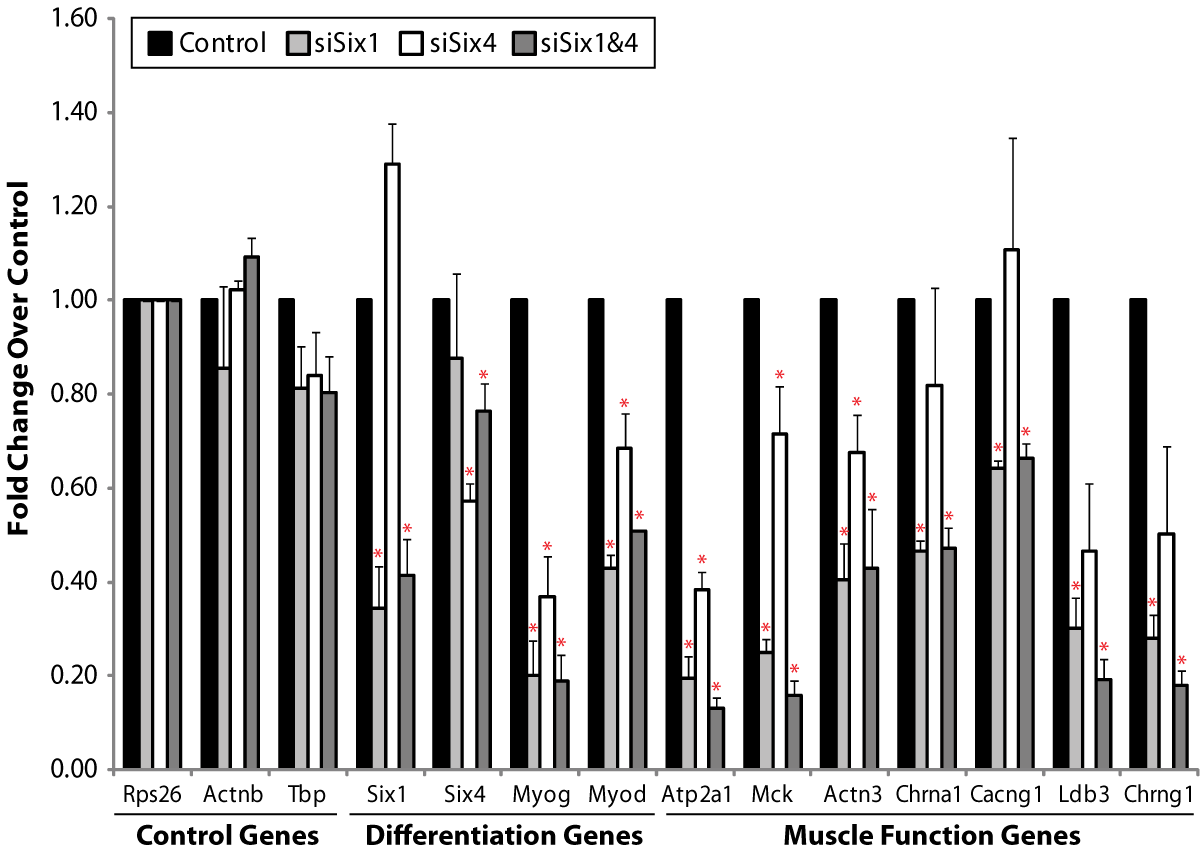

Figure 14 – Six1 and Six4 knock-down down-regulates genes associated with muscle differentiation and function.

Levels of mRNA in C2C12 treated with siRNA against Six1, Six4, both or Control sequence were quantified by real-time PCR (qRT-PCR) on a number of muscle differentiation and muscle function genes. mRNA levels were normalized to the Rps26 gene. Actin-Beta (Actnb) and the TATA Binding Protein (Tbp) were used to assess the quality of the Rps26 control gene. The data represent the average of at least 3 independent biological siRNA experiments. Error bars indicate the Standard Error over the Mean (SEM). * Represents statistically significant change by a two-tailed paired Student t-test ($p_{value} \leq 0.05$) compared to the Control siRNA.

Interestingly, the single knock-down of Six4 causes a less notable down-regulation than the Six1 or compound knock-down with both Six1 and Six4. On the other hand, the single Six1 and the Six1/Six4 combined knock-down show similar or more pronounced down-regulation of genes expressed during the terminal phase of differentiation. The results presented here indicate that Six TFs are required for proper induction and expression of genes necessary for proper skeletal muscle cell differentiation and function.

2.2.4. Genes regulated by Six1 contain promoters associated with muscle differentiation

Logically, genes participating within a biological pathway are very likely to be under the control of a common transcriptional regulator (Luan and Li, 2003). Since TFs bind to specific DNA consensus sequences, the genes in the cluster will most likely have enrichment for a particular DNA sequence recognized by the transcriptional regulator. A 5 kilobase pair (kb) promoter region upstream of genes in Cluster #1 was analysed with the Whole Genome rVista bioinformatics tool (Dubchak and Ryaboy, 2006). Most importantly, the program only considers TF binding sites that are conserved between the mouse and human genomes. The conservation between both species suggests that the TF binding site is functionally important (Elnitski et al., 2005; Frazer et al., 2004; Schwartz et al., 2003a; Schwartz et al., 2003b). The results showed a statistically significant over-representation of position weight matrices (PWM) matching the E-box and Mef2 motif, known to be bound by MRFs and Mef2 families of TFs (Table 2) (Andres et al., 1995; Blackwell and Weintraub, 1990; Fickett, 1996a, b; Li and Capetanaki, 1994). These observations therefore suggest that Six1 and Six4

are regulators of expression of muscle genes and they may work in cooperation with the MRFs and Mef2 TFs.

Promoter	Number of hits in the submitted regions	Total number of hits on genome	p-Value
MYOGENIN	567	14571	5.600E-08
HMEF2	12	113	2.122E-04

Table 2 – Genes regulated by Six1 and Six4 are characterized by the presence of phylogenetically conserved binding sites for muscle regulatory transcription factors.
Promoter analysis (5kb upstream) of genes regulated by Six1 and/or Six4 with Whole Genome rVista showing enrichment for muscle specific promoters.

To further authenticate the bioinformatics finding from the combined gene clustering with the gene functional and promoter analysis, an alternative computational analysis was used. Gene Set Enrichment Analysis (GSEA) is a bioinformatics tool used to identify genes with similar expression profiles and relate them to a biological process (Subramanian et al., 2005). This algorithm makes use of the entire gene expression data set, unlike the previous methods where pre-selection of co-expressed genes (probe fluorescence, statistical and fold change filtering) is required. Applying filters and removing probes from the analysis may cause bias in the interpretation of the data caused by removal of probes with minor fold changes. Although selecting genes with large fold changes can be useful in transcriptome analysis, it is also relevant to consider genes with minor fold changes. Genes that have small changes may also have a large impact on the expression of other genes. Also, when examining genes with low fold changes, the co-expression with other genes must also be considered in the context of a genome wide analysis. GSEA analysis of the genes down-regulated by the knock-down of Six1, as expected, results in the enrichment of gene sets termed with muscle function and TF binding motifs associated with muscle function and differentiation (Table 3) (The full GO list is available in Table S 7 in the Appendix).

Gene Set	Size	Enrichment Score	FDR
REACTOME STRIATED MUSCLE CONTRACTION	28	2.671	0.000
REACTOME MUSCLE CONTRACTION	44	2.507	0.000
MUSCLE_DEVELOPMENT	85	2.022	0.002
V$MEF2_02 (Mef2 Binding Site DNA Motif)	156	1.876	0.014
V$MYOD_Q6 (Myod Binding Site DNA Motif)	158	1.875	0.015
HEMOPOIETIC OR LYMPHOID ORGAN DEVELOPMENT	48	-2.058	0.012
HEMOPOIESIS	46	-2.074	0.010
TISSUE REMODELING	17	-2.159	0.005
BONE REMODELING	17	-2.167	0.007
REACTOME CLASS B2 SECRETIN FAMILY RECEPTORS	48	-2.230	0.002

Table 3 – GSEA Analysis of Control vs. siSix1.
Gene Set Enrichment Analysis of the entire data set comparing the Control siRNA treatment with the Six1 knock-down. A subset of GSEA categories are shown in the table. For the full list, see Table S 7. FDR corresponds to the False Discovery Rate. Red enrichment score indicates genes are expressed at higher levels in the control vs. the Six1 knock-down; blue enrichment score indicates genes are expressed at higher levels in the Six1 knock-down vs. the control.

Examination of the genes found within GSEA gene sets enriched when Six1 function is compromised reveals genes which overlap with the ones identified by CLICK and GO analysis. Interestingly, genes that are up-regulated by the knock-down of Six1 are involved in hemopoiesis and bone remodeling. The latter finding suggests that Six1 may repress key genes that confer other cell identity (i.e. hemopoiesis and bone) which allows expression of genes found in the C2C12 cells. Again, these findings support the hypothesis that both Six1 and Six4 regulate genes involved in myogenesis and they are required for proper development of terminally differentiated muscle cells.

In summary, the bioinformatics data analysis of genes that are regulated by Six1, determined by CLICK clustering, participate in muscle differentiation and muscle function. This finding was also confirmed by an independent bioinformatics analysis method. Overall, the results show that Six1 regulates genes that are involved in muscle function, differentiation and development.

2.2.5. Cooperation between Six transcription factors and Myogenin

To better understand the roles between the Six and the MRFs, RNAi was used to knock-down the expression of Myog. The C2C12 cells were transfected with siRNA directed against Six1, Six4, both Six1 and Six4, or Myog 24h prior to confluence and allowed to differentiate for 24h before harvest. Both protein western blots and mRNA quantification by qRT-PCR showed that the knock-downs for Six factors (Figure 8 and Figure 14) and Myog (Figure 15 B, C and D) were successful. To better differentiate between genes that are direct and indirect targets, microarray transcript profiling was performed and combined with Six1 ChIP-on-Chip binding data at 24h after the start of differentiation (Liu et al., 2010). Only genes bound by Six1 were retained in the analysis which yielded a set of 380 genes that are bound by Six1 and down-regulated by over 50% after Six1 loss of function. Similar to what was shown above (Figure 12) alignment of the data with the differentiation time course showed that the majority of these genes are induced during C2C12 differentiation (Figure 15A – Left Side of heatmap). In summary, these results show that Six1 directly binds and regulates genes essential in muscle differentiation and function.

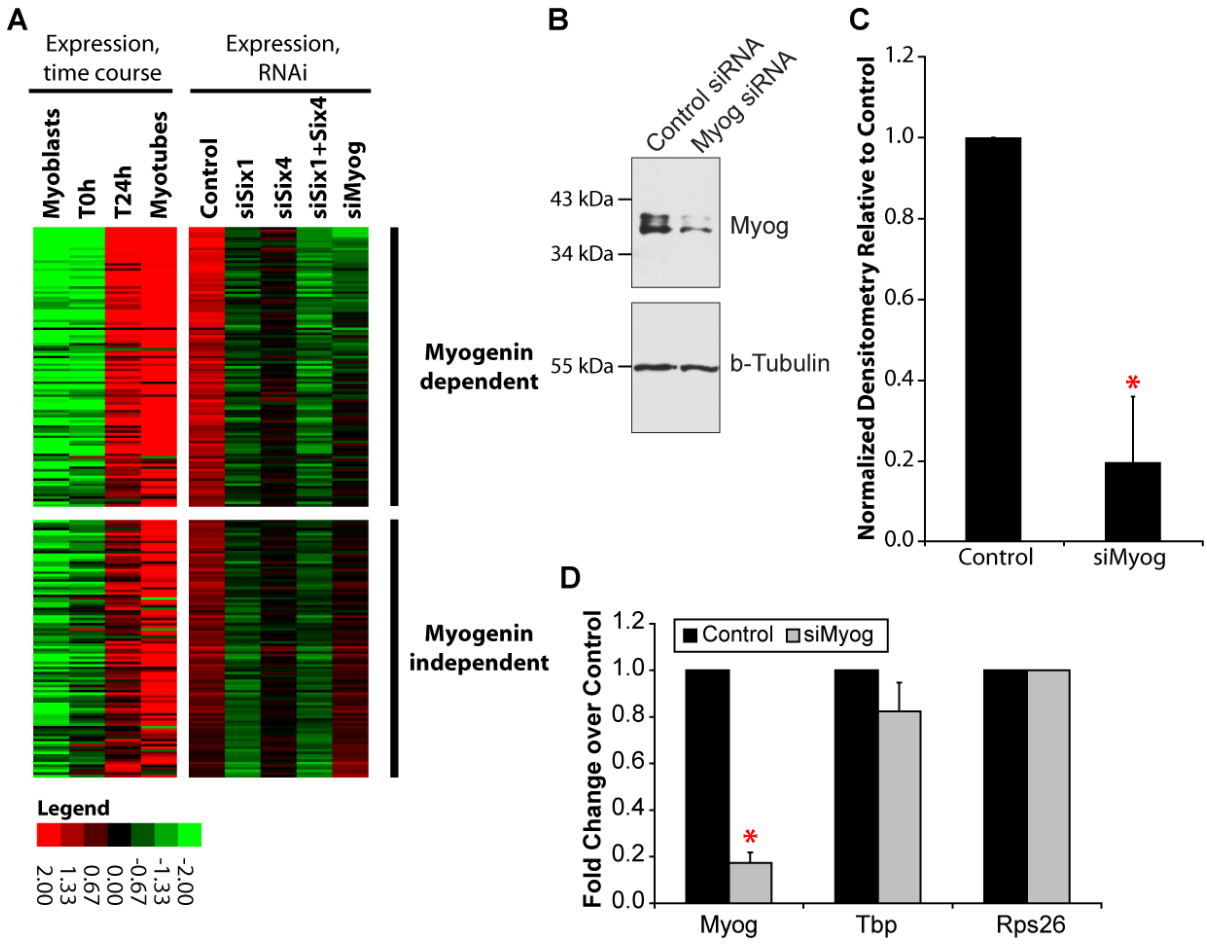

Figure 15 – Six Factors Cooperate with Myogenin.
A) Heat-map representing the expression levels of Six1 target genes (ChIP-on-Chip data) affected by knock-down of Six1 (≥2 fold change up or down) during normal differentiation (Expression Time Course – Left hand side) aligned with data from the knock-down of Six1, Six4, Six1 and Six4 or Myog (Expression RNAi – Right hand side). Genes are ranked from top to bottom in decreasing order of down-regulation after myogenin knock-down. Those for which the knock-down of Myog gave ≥50% reduction in expression are placed in the top portion (Myog-dependent), otherwise they lie on the bottom portion (Myog independent). Color Legend indicated under the heat-map. **B)** Protein level analysis of C2C12 cells treated with siRNA against Myog or a control sequence. B-Tubulin was used as a loading control. **C)** Western blot quantification of Myog by densitometry of 4 independent siRNA transfection experiments. Measurements were normalized to the densitometry of Beta-Tubulin. **D)** mRNA expression levels of Myog was quantified by qRT-PCR. mRNA levels were normalized to the Rps26 gene. The Tata Binding Protein (Tbp) was used to assess the quality of the Rps26 control gene. The data represents the average of 3 independent biological siRNA experiments. Error bars indicate the Standard Error over the Mean (SEM). * Represents statistically significant change by a two-tailed paired Student t-test ($p_{value} \leq 0.05$) compared to the Control siRNA.

Transcript profiling data from the Myog knock-down treated cells was combined with the Six loss of function data to better understand the roles that Six TFs play during skeletal myogenesis. In fact, this analysis allowed the distinction of genes that are bound and regulated directly by Six1 and genes that are indirectly regulated by Six1 (i.e. Six1 regulates the expression of Myog and Myog in turn regulates the gene). Alignment of the Myog knock-down data with the Six data distinguishes two categories of Six1 targets: (1) those that depend on Myog function for efficient induction during differentiation (i.e. they are affected by siMyog) and (2) those that are independent of Myog, targets that are affected by the Six knock-down, but not affected to a significant extent by the loss of Myog function (Figure 15A – Column siMyog). The Myog-dependent targets are therefore genes where both Six1 and Myog are required to be functional to induce their expression. These results not only demonstrate that Six1 and Six4 are necessary for muscle differentiation, but they also imply cooperation between Six1/Six4 and MRFs to allow terminal differentiation of muscle cells.

2.3. Chapter Conclusions

Overall the results presented in this chapter show that Six1 and Six4 TFs are required for proper skeletal myogenesis. Their knock-down leads to impairment of myoblast differentiation and fusion. The observed phenotype is partially attributed to lower expression of Myog, a key transcription factor required for terminal differentiation, and other genes expressed in differentiated muscle cells directly regulated by Six1. Interestingly, using a combination of functional genomics techniques, such as RNAi mediated knock-down of both Six1 and Six4 transcription factors, gene expression analysis and clustering, have further shown necessity of these TFs for myogenesis. Additionally, combining the gene transcript data from the Six1 and Six4 knock-down with the Myog RNAi data has revealed two distinct gene groups. The first group of genes where Six1 cooperates with Myog to regulate their expression and the second group contains genes that are regulated uniquely by Six1 independently from Myog. In summary, in cooperation with Myog, Six1 regulates genes that are required for proper skeletal myogenesis.

3. Chapter 3 – Six4 and the Cell Cycle: A Novel Role for Six4 during Myogenesis

3.1. Introduction

The process by which a precursor muscle cell differentiates into a mature muscle fiber is a complex sequence of events involving the expression of genes characterizing skeletal muscle and repression of genes conferring other tissue identity. Muscle cells originate from somites which arise from segmentation of the paraxial mesoderm. Once cells are committed to the muscle lineage, they migrate from the somites to the limb buds in the embryo (Tajbakhsh et al., 1998; Tajbakhsh and Buckingham, 2000; Tajbakhsh et al., 1996; Tajbakhsh et al., 1997) (Reviewed in Buckingham et al., 2003). Muscle progenitor cells, now named myoblast cells, are characterized by their high proliferative rate which leads to the generation of a pool of cells necessary to form the future mature muscle fibre. At the onset of terminal differentiation, the myoblasts withdraw from the cell cycle and begin expressing genes that are essential for myotube formation and function (Andres and Walsh, 1996). This important switch in the gene expression program fails to occur in rhabdomyosarcoma, a type of tumour where cells acquire a phenotype reminiscent of incompletely differentiated muscle cells that undergo continuous proliferation (Newton et al., 1995). Through the use of combined functional genomics approaches, RNAi against Six1 and/or Six4 with genome wide transcript profiling, it was discovered that Six4 specifically regulates the expression of cell cycle genes in myoblasts.

3.2. Results

3.2.1. Genome wide transcript profiling

As previously discussed in Chapter 2, genome wide transcript profiling was performed on C2C12 cells where Six1, Six4 or both TFs were targeted for knock-down by RNAi. The resulting data were analyzed by CLICK clustering to find genes with a common expression pattern which are likely to have similar function or implicated in the same biological pathway (Sharan et al., 2003; Sharan and Shamir, 2000; Ulitsky et al., 2010). The CLICK algorithm categorized the microarray probes with a significant 2 fold change in either up- or down-regulation into 8 distinct groups (Figure 11). Of these 8 gene clusters, 2 of them are of special interest, genes from Cluster #1 (discussed in Chapter 2) are down-regulated by the Six1 and double Six1/Six4 knock-down and mildly by the Six4 knock-down, whereas in Cluster #3, genes are specifically up-regulated by the Six4 knock-down. The average profile of genes from Cluster #3 is shown in Figure 16 A and B.

To further characterize the function of these 352 genes, they were aligned to the gene expression data from the C2C12 differentiation time course (MB, T_0, T_{24} and MT) to see how these genes are expressed during normal myoblast differentiation. Surprisingly, aligning the genes from the time course showed that nearly half (170 probes) of the probes up-regulated by the Six4 knock-down are normally expressed at higher levels in the myoblasts than in cell undergoing differentiation (Figure 16 C and Table S 4). Bioinformatic analysis was then performed to identify biological processes that might be enriched within this cluster of genes.

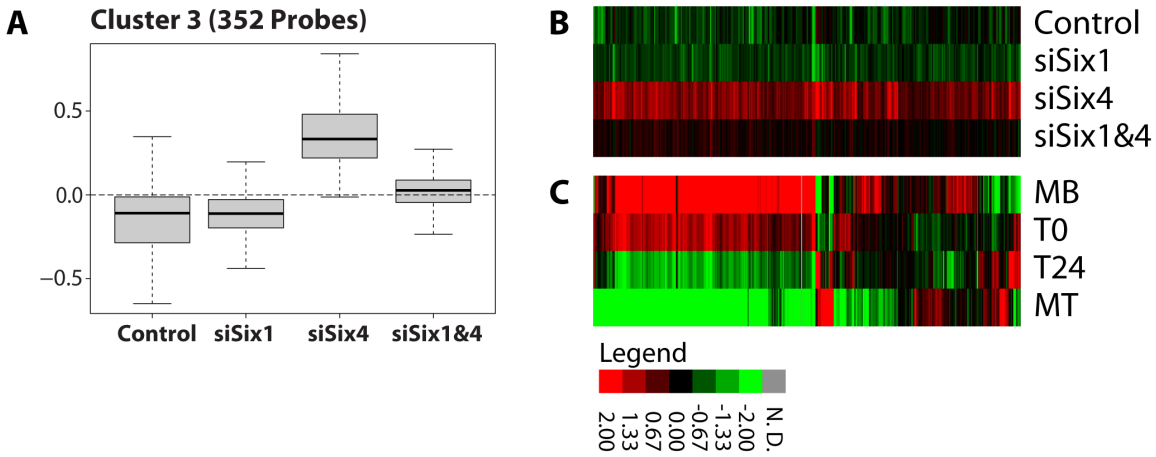

Figure 16 – Six4 regulates genes that are normally repressed during muscle differentiation – CLICK Cluster #3.

A) Box-plot representing the average expression profile of 352 probes over 4 siRNA treatments in C2C12 cells (siRNA against Six1, Six4, both or a control sequence) grouped into Cluster #3. Whiskers represent 1.5 x Inner Quantile Range of the 1st and 3rd quantiles. For the full gene list see Table S 4 located in the Appendix. **B)** Heat-map representation of the 352 probes from Cluster #3 where Six1 and/or Six4 are knocked-down by RNAi. **C)** Probes from the C2C12 differentiation mRNA profiling experiment at four different time points (MB, T_0, T_{24} and MT) were aligned with the 352 probes from Cluster #3 to identify the expression trends of these genes during C2C12 differentiation. Color legend is given under both heat-maps and represents normalized expression values in linear scale.

3.2.2. Genes regulated by Six4 participate in cell proliferation and the cell cycle

Gene Ontology (GO) analysis was performed on the genes found within Cluster #3 to identify their predominant biological functions (Huang da et al., 2009a, b). The data analysis revealed statistically significant enrichment of cell cycle related gene function above any random expectations (Table 4) (See Table S 6 in the Appendix for the full list).

GO Term	Count	Enrichment	p-Value[1]	Genes
M phase	40	10.71	9.65E-26	Bub1b, Cdc6, Ccna2, Cenpf
Cell cycle process	54	6.39	3.56E-25	Ccna2, Smc2, E2f1
Mitosis	34	11.61	1.25E-22	Bub1b, Cdc6, Smc2, Ccnb2
Regulation of cell cycle	18	3.39	3.34E-03	E2f1, Ccne1, Cdkn2d, Ccnb2

Table 4 – Six4 negatively regulates cell cycle function genes.
Genes found in Cluster #1 were analyzed by Gene Ontology Analysis (GO). A subset of GO categories is shown in the table. For the full list see Table S 6 in the Appendix. [1] p-Values were corrected by multiple hypothesis testing by Benjamini-Hochberg algorithm. The background set of genes used in the significance values was the entire mouse genome.

In fact, cell cycle genes are normally expressed at high levels in the myoblasts because of their high proliferation rate, but as differentiation progresses, these genes are turned off and repressed. At the onset of differentiation, the myoblasts exit the cell cycle and begin expression of genes required for myotube formation and function (Andres and Walsh, 1996). Several representative genes from different cell cycle categories were tested by qRT-PCR to examine their expression over a time course of skeletal muscle differentiation: cell cycle control genes (E2f1 and Foxm1 both transcription factors essential for the cell cycle); mitosis genes (Bub1b and Cenpe) and cyclin genes (Ccnd1). The mRNA quantification confirms the well known fact that genes involved in the control of the cell cycle, mitosis and cyclin genes are repressed as differentiation progresses (Figure 17) (Andres and Walsh, 1996).

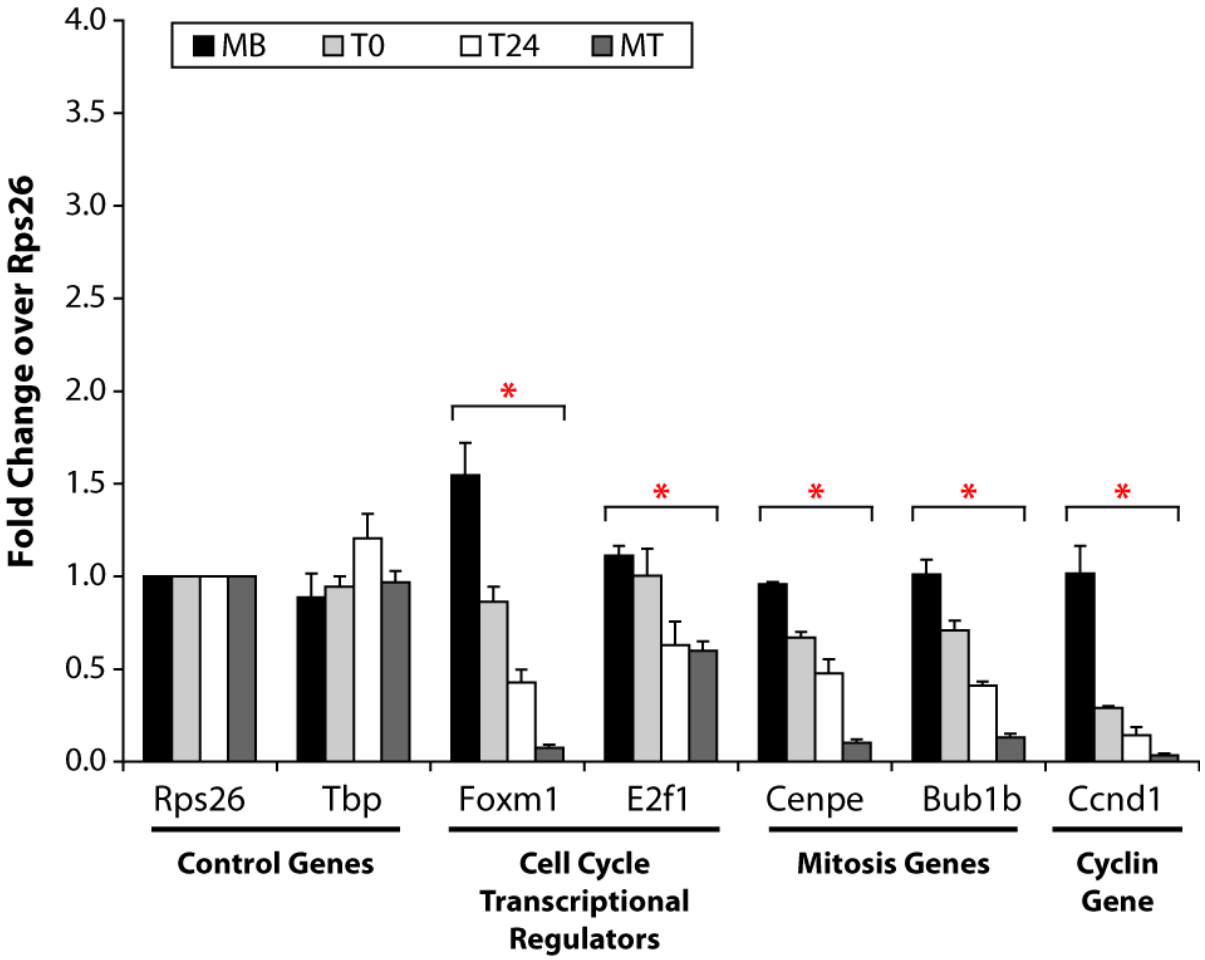

Figure 17 – Genes associated with cell proliferation are normally down-regulated during C2C12 differentiation.
Levels of mRNA during C2C12 differentiation at 4 different time points: Myoblasts (MB), 0h of differentiation (T0), 24h of differentiation (T24) and Myotubes (MT) were quantified by real-time PCR (qRT-PCR) on a number of cell cycle control, mitosis and cyclin genes. mRNA levels were normalized to the Rps26 gene as an invariant control and the TATA Binding Protein (Tbp) was used to assess the quality of the Rps26 control gene. The data represent the average of 3 independent biological time course experiments. Error bars indicate the Standard Error of the Mean (SEM). * Represents statistically significant change over the time course by single factor ANOVA analysis.

In summary, the results suggest that Six4 may function in part during myoblast differentiation by promoting cell cycle exit, possibly by repressing the expression of genes involved in proliferation.

3.2.3. Genes regulated by Six4 contain promoter elements associated with cell proliferation and the cell cycle

Promoter analysis was performed on the genes found in Cluster #3 under the assumption that genes sharing the same biological function would be under the control of one or more common transcriptional regulators, and therefore they may share specific DNA sequence elements within their regulatory regions (Dubchak and Ryaboy, 2006). The region 5kb upstream from the transcriptional start site (TSS) of those 352 genes was analyzed with the web-based program Whole Genome rVista (Dubchak and Ryaboy, 2006). Of note, the program only considers TF binding sites that are conserved between the mouse and human genome. The conservation between both species suggests that the TF binding site is functionally important (Elnitski et al., 2005; Frazer et al., 2004; Schwartz et al., 2003a; Schwartz et al., 2003b).

The analysis showed that a statistically significant enrichment for different variants of the E2F and Dp protein pair binding sites (p-value $< 5.14 \times 10^{-4}$) which are well known to be associated with the control of genes regulating cell cycle progression (Cam and Dynlacht, 2003; Helin et al., 1993; Takahashi et al., 2000). This finding is consistent with the prior result that knock-down of Six4 causes a specific up-regulation of genes involved in the cell cycle.

To further validate the finding that loss of function of Six4 causes a specific up-regulation of cell cycle genes, a second bioinformatics analysis was performed with the GSEA to group genes with similar expression patterns and indentify if they function within a common biological process (Subramanian et al., 2005). The analysis revealed that genes down-regulated by the knock-down of Six4 are mostly involved in muscle structure and function (Table 5) (The full GO list is available in Table S 8 in the Appendix). This is consistent with a similar observation made with the Six1 knock-down (genes from Cluster #1, discussed in Chapter 2). However, genes that are up-regulated by the same RNAi treatment are components of the cell cycle pathway, particularly the mitotic phase of the cell cycle (Table 5). This independent clustering analysis corroborates the results shown in the previous GO analysis and suggests that Six4 plays a role in the repression of cell cycle genes or a delay in cell cycle exit at the onset of skeletal muscle differentiation. However, the mechanism responsible for this observation is still unclear.

Gene Set	Size	Enrichment Score	FDR
REACTOME STRIATED MUSCLE CONTRACTION	28	2.436	0.000
STRUCTURAL CONSTITUENT OF MUSCLE	21	2.332	0.000
REACTOME MUSCLE CONTRACTION	44	2.259	0.000
SPINDLE	33	-2.133	0.000
M PHASE OF MITOTIC CELL CYCLE	64	-2.242	0.000
MITOSIS	62	-2.301	0.000

Table 5 – GSEA Analysis of Control vs. siSix4
Gene Set Enrichment Analysis of the entire data set comparing the Control siRNA treatment with the Six4 knock-down. A subset of GO categories is shown in the table. For the full list, see Table S 8. FDR corresponds to the False Discovery Rate. Red enrichment score indicates genes are expressed at higher levels in the control vs. the Six4 knock-down; blue enrichment score indicates genes are expressed at higher levels in the Six4 knock-down vs. the control.

The data analysis of genome wide transcript profiling technique may result in a high false positive rate due to limitations of quantification inherent in the technology (Gusnanto et al., 2007). Therefore, an alternative technique such as real-time quantitative reverse-transcription PCR (qRT-PCR) is often used to confirm findings from microarray data. A certain number of gene candidates representing different cell cycle categories were tested by qRT-PCR. Both cell cycle functions: mitosis genes (Cenpf, Cenpe, Bub1b and Smc2), cyclin genes (Ccnd1, Ccnb2 and Ccne1) and Cdc6 were specifically up-regulated in the Six4 knock-down with statistical significance (Figure 18). Genes known to regulate the cell cycle, E2f1 and Foxm1, were also quantified by qRT-PCR and they showed a statistically significant increase in the mRNA levels. It was also noticed that the Six1/Six4 compound knock-down also up-regulated these same genes, but to a lower extent when compared to the single knock-down of Six4. As for the expression levels of cell cycle genes in the Six1 knock-down, they remained unchanged as the microarray results indicated. Overall, these finding seem to indicate that Six4 plays a role in the withdrawal of C2C12 from the cell cycle.

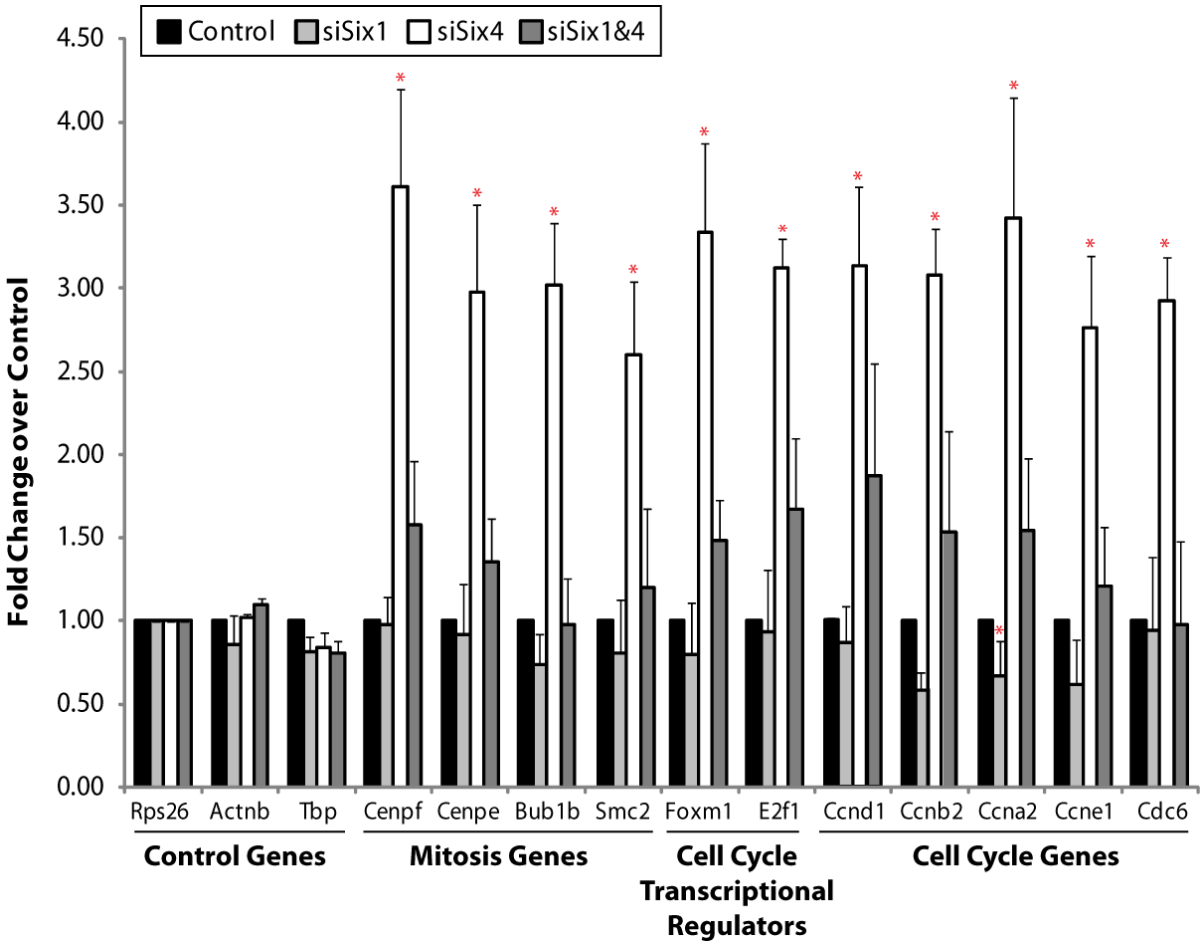

Figure 18 – Knock-down of Six4 induces the expression of cell cycle genes.
Levels of mRNA in C2C12 treated with siRNA against Six1, Six4, both or Control sequence were quantified by real-time PCR (qRT-PCR) on a number of cyclin, mitosis and cell cycle control genes. mRNA levels were normalized to the Rps26 gene. Actin-Beta (Actnb) and the TATA Binding Protein (Tbp) were used to assess the quality of the Rps26 control gene. The data represents the average of at least 3 independent biological siRNA experiments. Error bars indicate the Standard Error over the Mean (SEM). * Represents statistically significant change by a two-tailed paired Student t-test ($p_{value} \leq 0.05$) compared to the control siRNA.

3.2.4. Knock-down of Six4 in differentiating myoblasts causes an increase in the number of cells undergoing DNA replication

In-depth analysis of the genome wide transcript profiling data has revealed that the knock-down of Six4 causes a specific up-regulation of genes categorized as cell cycle and mitosis genes. To further confirm this finding, a functional experiment was designed and performed to validate this phenotype. A bromodeoxyuridine (BrdU) incorporation assay was used to evaluate the percentage of C2C12 cells that have gone through DNA replication in S phase during the pulse. Knowing the Six4 knock-down causes an up-regulation of cell cycle genes, an increase in the number of cells that have incorporated the BrdU label is expected. Quantification of the number of BrdU positive nuclei shows that Six4 knock-down causes a statistically significant increase in BrdU incorporation (Figure 19). This indicates a greater number of myoblasts transited through the S phase of the cell cycle during the BrdU pulse period. This result suggests that the down-regulation of Six4 in C2C12 causes not only an increase in the mRNA of cell cycle genes, but also has the consequence of increasing the number of cells that have gone through S phase. Other experiments aimed at complementing the BrdU results were attempted. For example, performing growth curves of knock-down cells and BrdU incorporation assays in primary myoblasts. However these experiments have proven unsuccessful due to technical difficulties.

Overall these findings not only confirm the initial experimental hypothesis of Six factors participating in skeletal muscle differentiation and function, but also implicate Six factors in other biological functions such as the cell cycle.

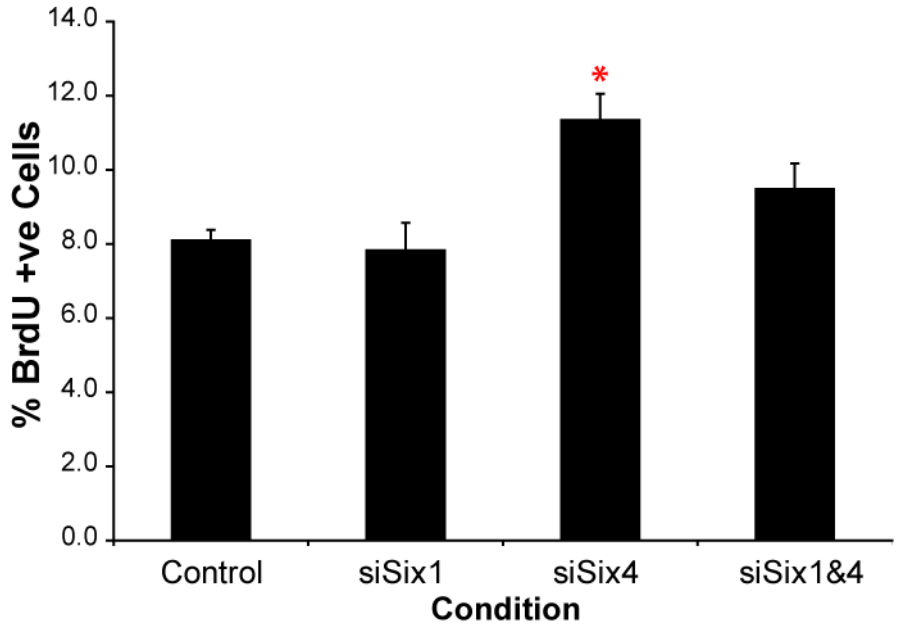

Figure 19 – Knock-down of Six4 increases incorporation of BrdU in C2C12 myoblasts. C2C12 cells were treated once 24h before treatment with siRNA against Six1, Six4, both or the control sequence. Cells were allowed to reach confluence and differentiate for 24h and were pulsed with BrdU at a final concentration of 10 μM for 4h prior to fixation. The percentage of BrdU positive cells over the total number of nuclei was evaluated. The data represent the average of 3 independent siRNA experiments. Error bars indicate the Standard Error over the Mean (SEM). * Represents statistically significant change by paired two-sided Student t-test ($p_{value} \leq 0.05$) compared to the Control siRNA.

3.2.5. Possible Mechanism of Six4 Regulation of Cell Cycle Genes

The data presented up to this point suggest that Six4 functions in part by participating in the repression of cell cycle genes. However, the mechanisms underlying this observation are still uncharacterized. There are at least 5 scenarios, represented in Figure 20, where Y is a key regulator of cell cycle genes and X is a regulator of Y:

1. Six4 directly *represses* proliferation genes;
2. Six4 directly *represses* Y which is an *activator* of cell cycle genes;
3. Six4 directly *activates* Y which is a *repressor* of cell cycle genes;
4. Six4 indirectly regulates cell cycle genes by *activation* of X, a *repressor* of Y, where Y is a direct *activator* of cell cycle genes;
5. Six4 indirectly regulates cell cycle genes by *repression* of X, an *activator* of Y, where Y is a direct *activator* of cell cycle genes

Because preliminary ChIP-on-chip data for Six4 in C2C12 cells do not support the idea of a direct binding of Six4 at a significant number of the cell cycle gene promoters being considered (I. Chakroun and A. Blais, unpublished results), it is hypothesized that induction of the cell cycle genes by the loss of function of Six4 is mediated through an indirect mechanism. Because a large number of up-regulated genes are predicted transcriptional targets of E2f1 and Foxm1, one reasonable scenario is that during normal myogenesis Six4 directly represses an activator of the cell cycle. Two possible candidates for the activator of cell cycle genes are E2F1 and Foxm1.

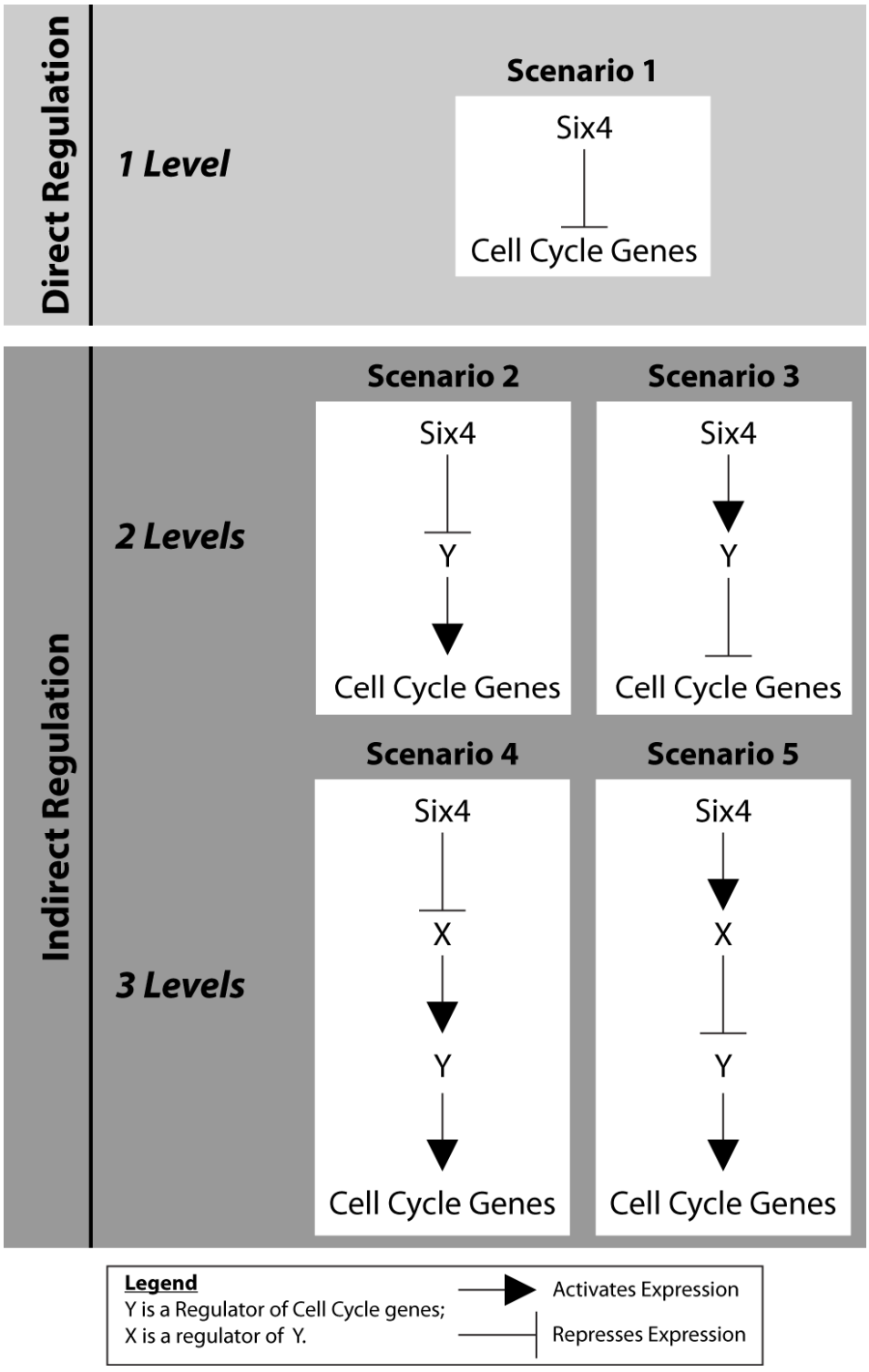

Figure 20 – Model of Six4 regulation of the Cell Cycle.
Model representing 5 possible scenarios where Six4 regulates genes of the cell cycle. This is not an exhaustive list of possible scenarios and other putative scenarios exist. See text for details

E2f1 is known to be an activator of transcription of genes essential for the progression of the cell cycle including CyclinE and CyclinA (Dyson, 1998; Harbour and Dean, 2000; Trimarchi and Lees, 2002) which are both up-regulated by the Six4 knock-down (Figure 18).

Alternatively, Foxm1 is known to control many steps in the cell cycle such as G1 to S transition checkpoint, S-phase and Mitosis (Laoukili et al., 2005). Foxm1 target genes include cyclins (CyclinA2, CyclinB, CyclinD2 and CyclinE) and mitosis genes (CenpF and CyclinB) (Alvarez et al., 2001; Krupczak-Hollis et al., 2004; Laoukili et al., 2005; Wang et al., 2005; Wang et al., 2002).

The gene promoter analysis did not yield any TFs binding sites for Foxm1. Possible explanations for these results include: (1) the Foxm1 TF binding site is not well characterized, therefore the search yields no enrichment for the site or (2) the site is characterized, but it is not conserved between mice and humans. This latter explanation is due to the design of Whole Genome rVista, the tool when searching for TF binding site enrichment only retains sites that are conserved between mice and humans (Dubchak and Ryaboy, 2006; Elnitski et al., 2005; Frazer et al., 2004; Schwartz et al., 2003a; Schwartz et al., 2003b).

Interestingly, the transcript levels of most Foxm1 and E2f1 target genes are up-regulated specifically in the Six4 RNAi treatment (Figure 18). Together these results suggest that both E2f1/Dp1 pair and Foxm1 may be important players in the up-regulation of cell cycle genes when Six4 function is lost.

3.2.6. Foxm1 is Up-Regulated when Six4 is Knocked-Down in Primary Myoblasts

C2C12 cells were used in all of the results presented so far. It has proven to be a valuable tool to better characterize the role of several TFs in the myogenic process (Asp et al., 2009; Asp et al., 2011; Aziz et al., 2010; Benhaddou et al., 2011; Blais et al., 2005; Cao et al., 2006; Cao et al., 2010; Iezzi et al., 2004; Rampalli et al., 2007; Seenundun et al., 2010; van Oevelen et al., 2010). However, the use of C2C12 may not be ideal to study the mechanisms that control the cell cycle, due to the fact that these cells have become spontaneously immortal through unknown random mutagenic event(s). One or many mutations may have occurred to cell cycle genes in order to maintain the cell line in this state (Pajcini et al., 2010). To assess the physiological relevance of the cell cycle phenotype found in the C2C12 cell line, it is essential to test if the Six4 loss of function causes the same phenotype in cells that are not immortalized, such as primary myoblasts. The use of primary myoblasts would overcome the previous problem, because they are muscle precursor cells that are freshly isolated from mouse muscles (Rando and Blau, 1994) and performing experiments after a limited number of population doublings essentially nullifies the likelihood of immortalization.

The main challenge with primary myoblasts is their well-known tendency to differentiate precociously; even in conditions of high growth factor concentration (A. Chu unpublished observations and Asp et al., 2011). To reduce this possibility a very high concentration of growth factors is added to the primary cultures to keep the cells in the proliferative state preventing them from untimely differentiation.

Given that primary cultures are better models to study cell cycle mechanisms, RNAi experiments were performed in the same manner as in C2C12 cells. Efficient knock-down of Six1, Six4 or both was achieved. Myog levels were also down-regulated in the Six4 and double knock-down. However, the Six1 knock-down only mildly reduces the expression of Myog (Figure 21A and B). This is in contrast to the experiments performed in C2C12, where knock-down of Six1 and Six4 have comparable potencies (Figure 8). Inconsistency of the results may be accounted for by the variable nature of primary myoblasts used for these experiments. This may be caused by slight variations between primary culture preparation or responsiveness of the cells to RNAi treatment. Western blots performed on these RNAi-treated primary cells show that Foxm1 expression is induced by the Six4 knock-down only (Figure 21 A and B), suggesting that Foxm1 may be the potential cell cycle regulator repressed by Six4 during normal skeletal myogenesis. However, E2f1 has not been ruled out as a possible candidate responsible for the Six4 cell cycle phenotype.

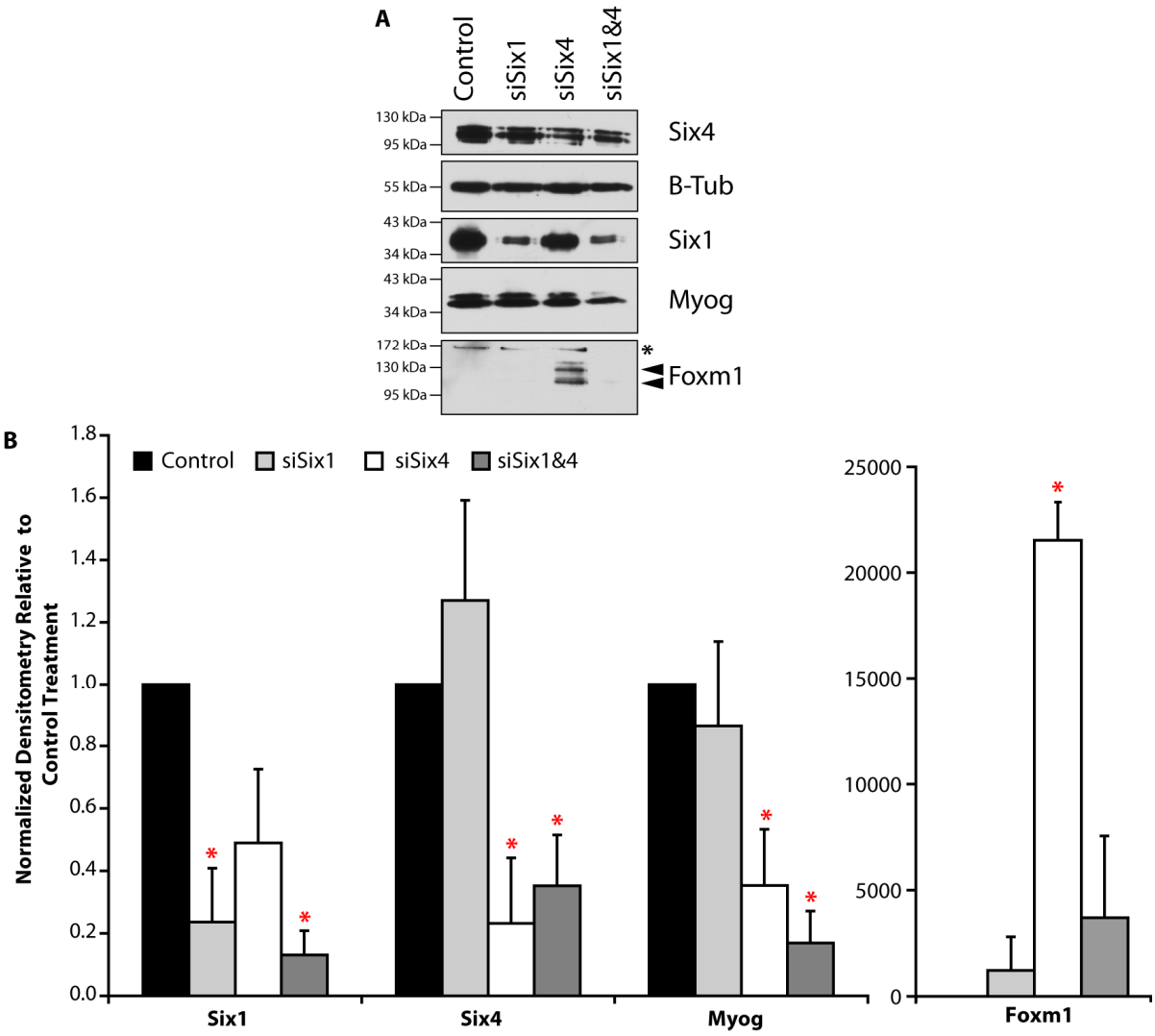

Figure 21 – Knock-down of Six4 causes an up-regulation of Foxm1 in Primary Myoblasts.

A) Primary myoblast cells were treated once with siRNA duplexes against Six1, Six4, both or the control sequence 24h prior to the start of differentiation. The cells were allowed to reach confluence, then switched to the differentiation medium for 24h and were harvested for analysis by western blot. Of note, Six4 specifically causes the up-regulation of Foxm1 a key regulator in the cell cycle. * Represents a non specific band recognized by the antibody. Both arrows heads designate the Foxm1 protein, the upper band being the phosphorylated form of Foxm1 (Laoukili et al., 2008). **B)** Western blot quantification of Six1, Six4, Myog and Foxm1 by densitometry of 3 independent siRNA transfection experiments in primary myoblasts. Measurements were normalized to the densitometry of Beta-Tubulin. * Represents statistically significant change by paired two-sided Student t-test ($p_{value} \leq 0.05$).

3.3. Chapter Conclusions

Overall, the results shown in this chapter indicate that Six4 plays a role in the regulation and the repression of cell cycle genes at the onset of differentiation. The gene expression analysis combined with different bioinformatic techniques revealed that genes up-regulated when Six4 is knocked-down by RNAi are related to the cell cycle. Complementary experiments also show that knock-down of Six4 causes an increase in the proliferation rate, further supporting the findings from transcript profiling experiments. Furthermore, the increase in the expression of cell cycle genes and subsequently the proliferation rate of cells treated with siRNA against Six4 may possibly be caused by the high expression levels of Foxm1.

4. Chapter 4 – Discussion

The main goal of this work was to characterize and clarify the role of Six1 and Six4 TFs in postnatal skeletal myogenesis, while also considering the contributions of the MRFs family of TFs. Using a combination of large scale functional genomics approaches and bioinformatics tools, the findings within this thesis provide a valuable insight into the genes that both these TFs regulate, whether directly or indirectly. Additionally, this work highlights that both these transcription factors can regulate a number of common muscle function and differentiation genes. However, each TF may also regulate a separate number of genes independently, as the results from this work highlight.

4.1. Requirement of Six Factors for Proper Myoblast Differentiation

Six1 knock-down in C2C12 resulted in impaired differentiation and fusion of myoblast cells, being more severe with the combined Six1/Six4 knock-down. These findings are consistent with the embryonic muscle developmental models in mice where muscle hypoplasia takes place in the $Six1^{-/-}$ animals and increases in severity when both Six1 and Six4 are knocked-out (Giordani et al., 2007; Grifone et al., 2005; Laclef et al., 2003). In both studies, loss of Six1 and both Six1/Six4 leads to lower expression of key markers of muscle differentiation such as Myog and Myod (Grifone et al., 2005; Laclef et al., 2003).

Knock-down of Six4 alone in myoblasts also results in impaired differentiation and fusion, but to a lesser extent than the Six1 and combined Six1/Six4 knock-down. The results are at odds with the findings in the Six4 null mice where no muscle development phenotype was observed (Ozaki et al., 2001). The differences can be explained in a number of ways. First, the sensitivity of the methods used by Ozaki *et al.* compared to this study. The authors of the

Six4 knock-out study reported that Myog and a terminal differentiation marker Atp1a1 (a sodium, potassium channel) are correctly expressed as determined by northern analysis and in situ hybridization (Myog only). On the other hand, this study uses a quantitative PCR, which is a more sensitive method to assay mRNA levels. In fact, the effect of Six4 knock-down in C2C12 is mild and if a similar effect occurs in the Six4 homozygous null mouse (mild effect from loss of Six4) in an *in vivo* model, the phenotypic outcome may not be clear or not observed in a mouse model. Another explanation for the discrepancy between the results reported here and the ones described by Ozaki and colleagues is the difference in the model system used. In the Six4 knock-out mouse model, the contribution of other Six family members can compensate for the loss of Six4. The most likely candidate is Six5, since it is the only other family member with an extended C-terminal transcriptional activation domain. Six5 is expressed at sufficient levels in the embryonic model to compensate for the loss of Six4 (Fougerousse et al., 2002; Ozaki et al., 2001). In contrast, in the C2C12 cell line expression of Six5 is low, therefore insufficient to compensate for the loss of Six4 (Liu et al., 2010). Another difference between both studies is the context under which the function of Six4 is studied. The work presented here, investigated the role of Six4 in a regeneration model. On the other hand, Ozaki and colleagues examined the role of Six4 during embryonic development. They did not challenge the adult mice with muscle injury to see if regeneration would be impaired in the absence of Six4.

4.2. Towards a Possible Six1 and Six4 Mechanism of Action during Skeletal Muscle Differentiation

The exact mechanism by which Six TFs activate or represses gene transcription, particularly of muscle specific genes, is still widely uncharacterized. One could hypothesize that both Six and MRF TFs recruit different proteins to facilitate transcription of muscle function and differentiation genes (Fuda et al., 2009). Chromatin remodelers have been shown to play an important role in the expression of terminal muscle genes. Myod is known to recruit PCAF, a histone acetyltransferase through, interaction with p300 (Puri et al., 1997a; Puri et al., 1997b), and Myog has been shown to recruit RBP3, a core subunit of the RNA polymerase II complex (Corbi et al., 2002). It was recently reported that Six4 is required for recruitment of UTX, a demethylase enzyme (Agger et al., 2007; Lan et al., 2007), to the Myog and Mck genes promoters. Interaction of UTX and Six4 is necessary to remove the Histone H3 Lysine 27 trimethylation (H3K27me3) (Seenundun et al., 2010), a histone mark associated with transcriptional repression (Barski et al., 2007; Boyer et al., 2006; Lee et al., 2006; Roh et al., 2006). The removal of the H3K27me3 subsequently allows the chromatin to be accessible for gene transcription. Importantly, upon Six4 knock-down, UTX recruitment is impaired; this prevents H3K27me3 demethylation on target genes and results in impaired gene expression of Myog and Mck. This suggests that Six4 may play a role in recruitment of chromatin remodelers to key genes involved in terminal myoblast differentiation. Results from this study show that knock-down of Six4 leads to lower expression levels of terminal muscle differentiation markers such as Myog and Mck, which partially support the Six4-UTX mechanism. However, the contribution of Six1 in the recruitment of UTX or other

chromatin remodelers to the Myog and Mck promoters is an important question that needs to be addressed.

Given the possibility that Six4 and UTX can be recruited to the Myog and Mck loci, one could wonder if UTX is recruited to all, or a subset of Six4 target genes. The question of UTX recruitment can also be expanded to Six1 gene targets. More importantly, are Six1 regulated genes targeted by UTX? And if so, do they overlap with the Six4-UTX targets? Also, are genes unique to Six1 targeted by UTX, or are genes regulated by Six1 and Myog (Myog independent vs. dependent genes) also targeted by UTX? To determine if UTX is recruited to all or a partial number of Six4 gene targets, performing ChIP-on-Chip or ChIP-Seq for both Six TF and the demethylase would reveal some information on the Six-UTX binding across the genome. Additionally, performing a ChIP for Six4 subsequently followed by a ChIP for UTX (ChIP-reChIP) could be used to confirm recruitment to the same locus and allow distinction between unique Six and UTX targets.

The implications of UTX targeting different Six4 genes remain unclear if it happens at all. However, some scenarios can be postulated. If UTX targets all Six4 regulated genes, this could imply that recruitment of UTX is necessary for demethylation of histones and allows transcription of Six4 target genes. On the hand, if UTX only targets a subset of Six4 regulated genes, this implies that UTX is not essential for expression of Six4 target genes. Furthermore, attention should be given to the timing of demethylation by the Six4-UTX complex during myogenesis (i.e. demethylation occurs during early vs. late differentiation). In sum, further work is required to characterize the role of Six4 and UTX, in addition to defining the role of Six1 in this mechanism, during myoblast differentiation.

4.3. Combinatorial use of functional genomics techniques and bioinformatics reveal known biological functions

Making use of a systems biology approach to better characterize the role of Six1 and Six4 in skeletal muscle differentiation, creates a set of unique challenges. The main difficulty is the analysis of the large amount of data generated from these studies. To reduce the complexity of large genome wide studies, it is necessary to make use of computational or bioinformatics tools. In this study, the use of different bioinformatics methods for clustering analysis combined with gene ontology analysis, gene set enrichment analysis and gene promoter analysis have provided a better understanding of the genes regulated by Six1 and Six4.

As expected, CLICK clustering combined with GO analysis showed genes down-regulated by the knock-down of Six1 or both Six1 and Six4 participate in muscle differentiation, development and function. These results are consistent with a recent study where whole genome transcript profiling was performed on muscle from mice carrying the Six1/Six4 double null mutation revealing a down regulation of genes involved in muscle structure, signal transduction and transcription (Niro et al., 2010; Richard et al., 2011). Additionally, both authors show that Six1 and Six4 are specifically required to induce fast-type muscle genes. Comparing the whole transcript profiling experiments from this study with the ones performed in the knock-out mice model would be insightful. Probes corresponding to fast- and slow-type muscle fibers were queried from the C2C12 differentiation time course and the Six1/Six4 RNAi treated myoblasts (Figure 22 and Table S 9) (Gene lists from Niro et al., 2010; Richard et al., 2011).

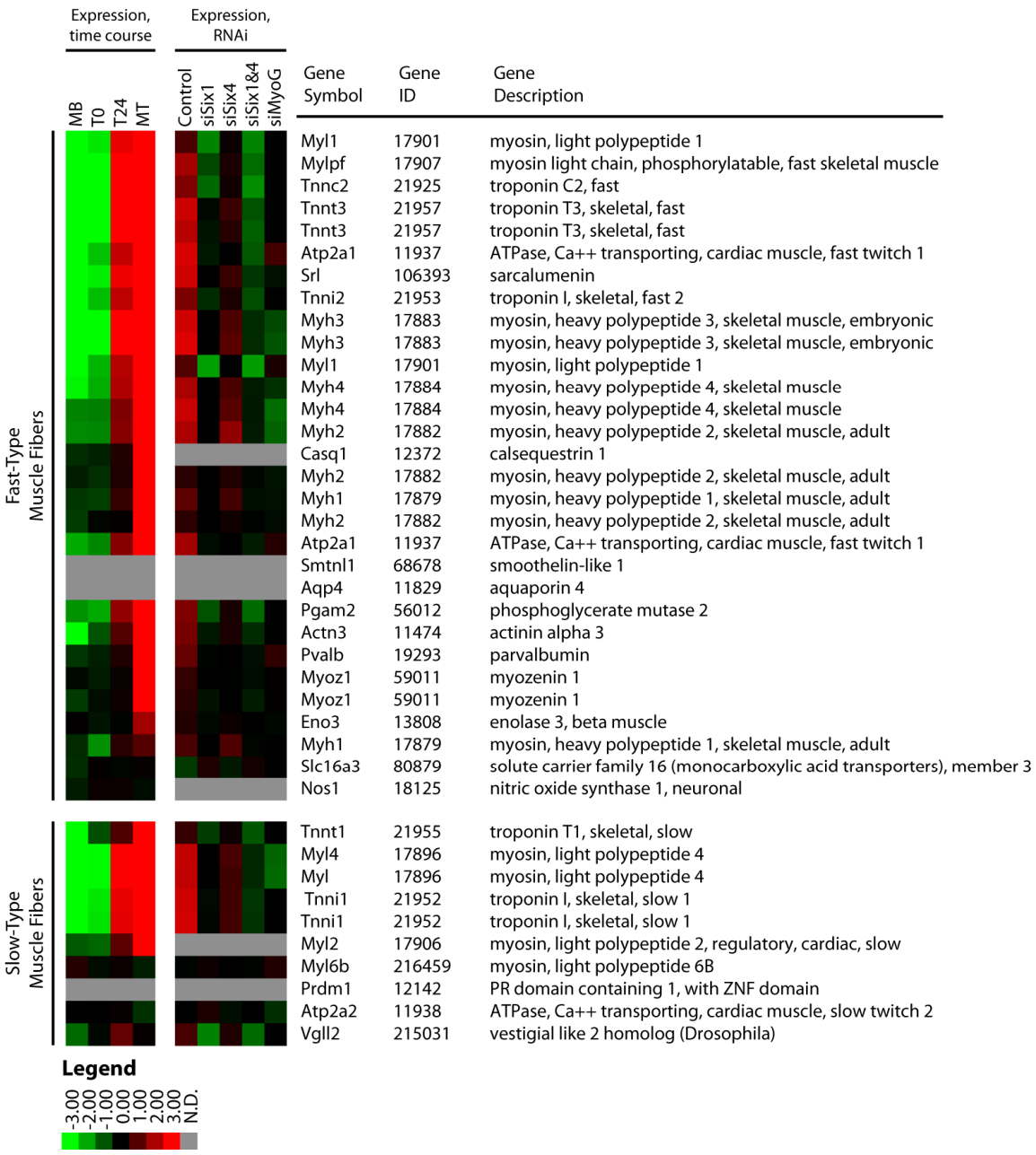

Figure 22 – Six Transcription Factors and Expression of Fast- and Slow-Type Muscle Genes in Myoblasts.

Heat-map representing the expression levels of Fast- and Slow-Type muscle fibers during a C2C12 differentiation time course and siRNA mediated knock-down of Six1, Six4, both Six1 and Six4, Myog and a control non-targeting sequence. Gene Symbols, Gene ID and Gene Descriptions are indicated. Gene list taken from (Niro et al., 2010; Richard et al., 2011). Fold changes for RNAi treatments are provided in Table S 9 found in the Appendix

The microarray data indicate that both types of muscle fibers are expressed at low levels in myoblasts and expression is significantly increased in differentiated myotubes as expected. Upon examination of knock-down microarray data, RNAi treatment for Six1 causes a noticeable decrease in fast- and slow-type muscle fibers. Interestingly, myoblasts treated with siRNA against Six4 have a mild decrease in expression of both fast- and slow-fibers compared to the Six1 knock-down. Furthermore, knock-down of Six1 and Six4 together causes a more significant decrease in fast- and slow- muscle type genes. Intriguingly, knock-down of Myog causes a similar gene expression profile as the Six1 RNAi treatment. Overall, Six1 and Myog seem to regulate both fast- and slow- muscle fiber types in myoblasts. However, it is to note that both studies were done in different model systems (i.e. in *vivo* vs. in *vitro*) and different physiological context (i.e. embryonic development vs. regeneration model) making comparisons between them more complex.

Additional analysis of the genes regulated by the Six1 and Six1/Six4 loss of function reveal that these genes contain DNA sequence elements enriched for the E-box and Mef2 motifs, known to be found in promoters of muscle differentiation and function genes (Andres et al., 1995; Blackwell and Weintraub, 1990; Fickett, 1996a, b; Li and Capetanaki, 1994). However, previous studies indicate that Six TF bind to Mef3 sites, which are not enriched in this analysis (Blais et al., 2005; Spitz et al., 1998). There are a few possible explanations to reconcile this discrepancy. One concerns the sensitivity of the bioinformatics tool used for this study. It is possible that the Mef3 PWM is too stringent and cannot be found within the promoter sequence of the genes investigated. A solution could involve using a more flexible Mef3 Position Weigh Matrix (PWM) or using a different analytical software such as CisGenome (Ji et al., 2006; Ji and Wong, 2005; Zhou and Wong, 2004).

Clustering of gene expression data and gene functional analysis are not the only methods to decipher the functions of genes found within a list. In fact, combining gene expression data with ChIP-on-Chip genome wide binding data and another TF knock-down involved with terminal differentiation (i.e. Myog) also provides valuable information on Six1 and Six4 function (Blais et al., 2005). The ChIP-on-Chip data set allows the distinction between genes directly regulated by Six1 and indirect gene targets. About 380 genes were found to be down-regulated due to knock-down of Six1. Among these genes 165 are directly bound by Six1. In fact, it is possible that Six1 directly binds to, and regulates additional genes, due to technical restriction of microarrays, these genes were not identified. Gene expression profiling and ChIP-on-Chip arrays contain a limited number of probes and are unable to cover the entire genome with their probe sets.

Additionally, the Six1/Six4 knock-down data and Six1 ChIP data sets were combined with the Myog knock-down transcript profiling data. This allows the distinction of genes that are directly bound and regulated by Six1 and genes bound by Six1 that are modulated by the knock-down due to impaired expression of Myog. The results show a subset of genes directly regulated by Six1 and not by Myog, whereas a second group are regulated by both Six1 and Myog. These results suggest that control of myogenesis occurs through a combinatorial fashion between the different TFs involved in muscle differentiation.

An alternative method to analyze genome wide transcript profiling is to use the GSEA analysis tool. As expected, the analysis reveals a significant number of genes being down-regulated by the Six1 knock-down correspond to genes sets or gene functional categories related to skeletal muscle structure, function, development and differentiation. Additionally, the same data analysis also show enrichment for TF DNA binding sites found in promoters

of muscle genes. Overall, the combination of different bioinformatics analytical tools has simplified the task of interpreting the large datasets and has confirmed previous knowledge of Six1 and Six4 function, where they play an important role in skeletal myogenesis.

4.4. Combinatorial use of functional genomics techniques and bioinformatics reveal novel biological functions

Although this study further strengthens the role of Six1 and Six4 in the myogenic process, the bioinformatics analysis also reveals other novel roles for both these TFs. For example, GSEA analysis showed genes up-regulated by the Six1 knock-down are enriched for genes sets involved in other cell types. Interestingly, hematopoiesis and bone remodelling are gene sets that are significantly enriched. In fact, it has been reported that C2C12 can be converted into osteoblasts (Bragdon et al., 2010; Katagiri et al., 1994; Okubo et al., 1999). However, this only occurs upon addition or exogenous expression of Bone Morphogenetic Protein-2 (BMP2), therefore C2C12 cells do not normally undergo transdifferentiation. The results presented here may support the idea that C2C12 have the potential (express the necessary proteins) to transdifferentiate into the osteoblast cell lineage. It is also possible that during C2C12 differentiation, expression of Six1 is required for repression of bone remodelling genes, thus explaining the increase in expression of bone remodelling genes. Therefore, the knock-down of Six1 would cause an up-regulation of osteoblast genes. The hypothesis may be valid, since mesenchymal cells are known to have the ability to differentiate into a variety of cell types including muscle cells (Ferrari et al., 1998; Mafi et al., 2011; Wakitani et al., 1995) where repression of the osteoblast genes would be required for muscle lineage commitment (Hayashi et al., 2008; Yang et al., 2007). The origin of myoblasts and osteoblasts also hints to this mechanism. Both cells types originate from the mesoderm germ

layer, specifically, the dermomyotome and the myotome for myoblast and the sclerotome for the osteoblasts (Gilbert, 2000; Huang et al., 2000). Studies investigating the roles of Shh and BMP-4 signalling revealed that both signalling molecules act in an antagonistic manner to restrict development of cartilage in the correct location in the embryo (Watanabe et al., 1998). Interestingly, experimental work involving explants and grafting of quail to chick embryos, indicate that ablation of the ectoderm subsequently affects the development of the dermomyotome and results in severe malformations of the ribs (Huang et al., 2000). The report suggests that muscle precursors from the dermomyotome and the myotome may contribute to bone development. In fact, Six1 and compound Six1/Six4 null mice display rib malformations. The rib phenotype may be a result from improper development of the dermomyotome and the myotome due to Six TFs loss (Giordani et al., 2007; Grifone et al., 2005; Laclef et al., 2003). However, this is at odds with the gene expression data from the RNAi mediated knock-down of Six1 in myoblasts, where bone remodelling genes are up-regulated. One could hypothesize that the up-regulation can occur in the mice models, however the levels are insufficient to restore proper bone development in the Six1 or compound Six1/Six4 null mice.

Another interesting finding in the GSEA computational analysis is the finding that knock-down of Six1 also causes an up-regulation of genes involved in hematopoiesis. Myoblasts have not been shown to be able to transdifferentiate into hematopoietic cells. It was initially claimed that myogenic cells were capable of differentiating into hematopoietic cells (Jackson et al., 1999). However, lineage analysis of hematopoietic cells originating from muscle cells showed that muscle derived hematopoietic cells originate from bone marrow and not from the muscle (Geiger et al., 2002; Kawada and Ogawa, 2001; McKinney-Freeman et al., 2002).

Again, as it is for the case of bone remodelling genes, there is a possibility that Six1 is required for the repression of hematopoietic genes in the C2C12 cell line. This hypothesis is supported by the fact that both muscle progenitor and hematopoietic cells originate from the mesoderm. Overall, the GSEA algorithm found two functional gene categories, bone remodelling and hematopoiesis, as up-regulated in the Six1 knock-down. This suggests that Six1 represses their expression during myogenesis.

Interestingly, GO and Whole Genome rVista bioinformatics analysis does not reveal any enrichment for TF binding sites for osteoblasts (i.e. CBFa1 or Runx2) (Cui et al., 2003; Lee et al., 2000a; Lian et al., 1998) and hematopoietic cells (i.e. Scl or Tal1, Gata2, Aml1 and Lmo2) (Reviewed in Zhu and Emerson, 2002). The inability of these algorithms to identify these genes can be due to a number of causes. First, clustering techniques often identify general changes in expression patterns over a number of conditions. For example, CLICK clustering was performed on a dataset of 4 conditions, while GSEA only compares two conditions at a time. Therefore, it is possible that a small group of hematopoiesis genes up-regulated in the Six1 knock-down may not have been significant or sufficient enough for the clustering algorithm to create a dedicated cluster for these genes. As previously discussed, GSEA analyzes each condition individually rather than identifying a general expression pattern across different RNAi conditions. This allowed for the identification of gene sets that are up-regulated by Six1 knock-down, such as bone remodeling and hematopoiesis. However, additional work is required to validate the findings where knock-down of Six1 in C2C12 causes an up-regulation of hematopoietic and bone remodeling genes. Experiments can include validation by qPCR of a number of hematopoietic and osteoblast candidate genes or to monitor Alkaline Phosphotase activity (ALP) and osteocalcin production, both used to

characterize osteoblasts (Katagiri et al., 1994; Okubo et al., 1999). To determine if hematopoietic cells are formed when Six1 is knocked-down in C2C12 cells, sorting by flow cytometry for hematopoietic markers such as Sca1 and CD45 (Geiger et al., 2002; Jackson et al., 1999; McKinney-Freeman et al., 2002) could be a method to verify this finding. It is also possible that Six1 knock-down is necessary, but not sufficient for transdifferentiation of myoblast cells into other lineages. Potentially, C2C12 cells could transdifferentiate with better efficiency if Six1 was knocked-down in addition to ectopic expression of BMP2. Overall, further work is still required to authenticate the results from the GSEA regarding the genes with increased expression when Six1 is knocked-down. Also, this further underlines the importance of using a number of different computational tools to analyze large data sets, due to strengths and weaknesses from different data processing algorithms.

4.5. The Role of Six4 in the Cell Cycle during Skeletal Muscle Differentiation

Arguably the most surprising finding from this study is the possible role of Six4 in regulating cell cycle genes. In fact, both the clustering analysis combined with the gene function/category and GSEA analysis indicate that knock-down of Six4 causes an up-regulation of cell cycle genes. Experimental results suggest that Six4 knock-down causes in increase in cell cycle gene expression and an increased number of C2C12 incorporating BrdU. Gene promoter analysis performed on genes that were identified to be up-regulated by the Six4 knock-down showed enrichment for the E2f/Dp pair of TF. The results further supports the findings from the gene functional analysis by GO and GSEA. This leads to the hypothesis that the E2f family of TF or Foxm1 may play an important role in the Six4 knock-down cell cycle phenotype. In fact, the expression levels of E2f1 and Foxm1 mRNA, both key regulators of the cell cycle, were up-regulated in the Six4 knock-down. Furthermore,

protein levels of Foxm1 were up-regulated specifically in the Six4 knock-down. However, promoter analysis did not reveal any Foxm1 DNA binding site enrichment. Even if the consensus site for Foxm1 has been characterized to be TAAACA, it has been hypothesized that other members of the Forkhead family of TF can bind to different variants of the consensus site (Korver et al., 1997; Littler et al., 2010; Ye et al., 1997). Additionally, the studies identified the DNA binding site using recombinant and purified proteins with chemically synthesized DNA, which does not correspond to the physiological function of Foxm1. These reasons may contribute to the failure in identifying Foxm1 binding sites. Other possibilities also include the lack of conservation of the Foxm1 DNA binding site; Foxm1 binds to an uncharacterized DNA sequence or binds to a sequence outside of the 5 kb analysis window. Finally, it is also plausible that Foxm1 binds indirectly to DNA via interaction with other TFs which bind DNA directly.

With the previous results, it was postulated that Foxm1 is responsible for the up-regulation of cell cycle genes in the Six4 knock-down. Supporting this hypothesis, a recent study in P19 cells showing that Foxm1 is important for the maintenance of pluripotency of embryonic carcinoma cells (Xie et al., 2010). The study showed that knock-down of Foxm1 in P19 cells caused spontaneous differentiation of the cells. On the other hand, if Foxm1 expression was maintained, expression of pluripotency markers such as Oct4 and Nanog can also be maintained. In addition, over expression of Foxm1 in fibroblasts, where it is not normally expressed, induced expression of the same pluripotency markers. This suggests that Six4 may be involved in regulation of the cell cycle and perhaps the stem cell like behaviour of satellite cells through Foxm1. The mechanism governing the up-regulation of Foxm1 is still uncharacterized. However, a recent article showed that Foxm1 is required for the

translocation and accumulation of β-Catenin in the nucleus resulting in the activation of Wnt target genes such as Axin2, c-Myc and Lef1 (Zhang et al., 2011). More importantly, Lef1 (Reviewed in Clevers, 2006) has been shown to be involved in cell proliferation (Reya et al., 2000) and regulating CyclinD1 (Shtutman et al., 1999; Tetsu and McCormick, 1999). Recent reports from the literature do provide initial clues to a possible Foxm1 mechanism. Although, the mechanism by which Six4 up-regulates cell cycle genes is hypothesized to be mediated through the function of Foxm1, it is still possible that E2f1 is also involved in the mechanism. Further experiments are needed to investigate the involvement of E2f factors in the Six4 cell cycle phenotype.

To investigate the contribution of Foxm1 and E2f1 in the Six4 proliferation phenotype, one could knock-down the expression of Foxm1 and E2f1 in combination with Six4. In fact, to evaluate the role of E2f1 in the Six4 phenotype, knock-down of DP1, a binding partner of the E2f family of TFs (Helin et al., 1993), would abrogate the activity and compensation by E2f family members (Reviewed in Polager and Ginsberg, 2008). One would expect that either the knock-down of Foxm1 or E2f1 would abolish the up-regulation of cell cycle genes mRNA and the incorporation of BrdU when Six4 is knocked-down.

A recent study also showed that over expression of Six4 in satellite cells caused a reduction of proliferation, supporting the conclusion of Six4 exerting a negative influence on cell proliferation (Yajima et al., 2010). Interestingly, from the studies of Yajima *et al.*, knock-down of Six4 in satellite cells causes inhibition of differentiation with no change in the proliferation rate. Instead, knock-down of Six5 causes an increase in the proliferation rate. The same study also verified the proliferation increase in satellite cells from mice sorted by Fluorescence-Activated Cell Sorting (FACS). Oddly, the authors show that satellite cells of

$Six4^{+/-}$ $Six5^{-/-}$ mice exhibit an increase in proliferation and not the $Six5^{-/-}$. In fact, the $Six5^{-/-}$ knock-out mouse has already been generated. The Six5 homozygous null mouse was found to develop cataracts but skeletal muscle development was not impaired (Klesert et al., 2000). Therefore, the mice were never examined for any cell cycle or proliferation defect in skeletal muscle. It is interesting to note that Yajima and colleagues do report that $Six4^{-/-}$ $Six5^{-/-}$ mice are never born and $Six4^{-/-}$ $Six5^{+/-}$ mice are rarely born, suggesting that Six4 alone is dispensable (Ozaki et al., 2001) but mice do require at least Six4 or Six5 (Yajima et al., 2010) for embryonic development. This report is at odds with the findings of this work, since knock-down of Six4 causes an increase in cell cycle genes and cell proliferation by BrdU pulse experiment. Again, both studies have been performed in two different model systems: C2C12 which are immortalized myoblasts and satellite cells isolated by FACS. Cells sorted by FACS often enrich for a specific population of cells and may not represent a heterogeneous population of cells. It is also possible that satellite cells have a compensation mechanism that an isolated model system lacks. In fact, Six5 is not expressed at appreciable levels in C2C12 making it difficult to assess the contribution of this TF (Liu et al., 2010). It is plausible that in the C2C12 cells, knock-down of Six4 causes a proliferation phenotype due to the lack of compensation by Six5 which is absent. On the other hand, in the mouse model, Six4 haploinsufficiency seems to be necessary for the increase in proliferation to occur in isolated satellite cells from the knock-out mice. This suggests that Six4 partially plays a role in proliferation. However, careful characterization of the $Six5^{-/-}$ satellite cells and mouse muscle regeneration are still required to validate this hypothesis.

Overall, this work has shown that both Six1 and Six4 are required for proper myoblast differentiation. In addition, examination of Six4 knock-down shows that Six4 plays a role in

regulating cell cycle genes. Combination of Six1, Six4 and Myog knock-down gene transcript profiling data with chromatin immunoprecipitation data allowed distinction between genes regulated directly or indirectly by Six1. Of particular note, genes in the latter category, indirectly regulated by Six1, are regulated by the activity of Myog or Myod. This suggests that Six TFs not only act upstream of MRFs but also in cooperation with them during terminal muscle differentiation. Additionally, bioinformatics analysis has also revealed that Six4 is implicated in the regulation of cell cycle genes. More importantly, this result suggests that Six4 can regulate a different set of gene targets independent of Six1 through TFs such as Foxm1 or E2f1, although these link have not been clearly established.

The findings of this work would also seem to indicate that terminal differentiation and the withdrawal of myoblasts from the cell cycle are events linked by the function of Six factors. Interestingly, both TFs are expressed in myoblasts, however, expression of Six1 does not induce premature differentiation and expression of Six4 does not cause a reduction in the proliferation rate. This suggests that both Six factors play a very important role during the final steps of myogenesis, Six TFs induce the expression of muscle function or differentiation genes either directly or through activation of the MRFs. At the same time, Six4 represses the expression of Foxm1 or E2f1, either in a direct or indirect manner, which subsequently causes the withdrawal of myoblasts from the cell cycle (Figure 23).

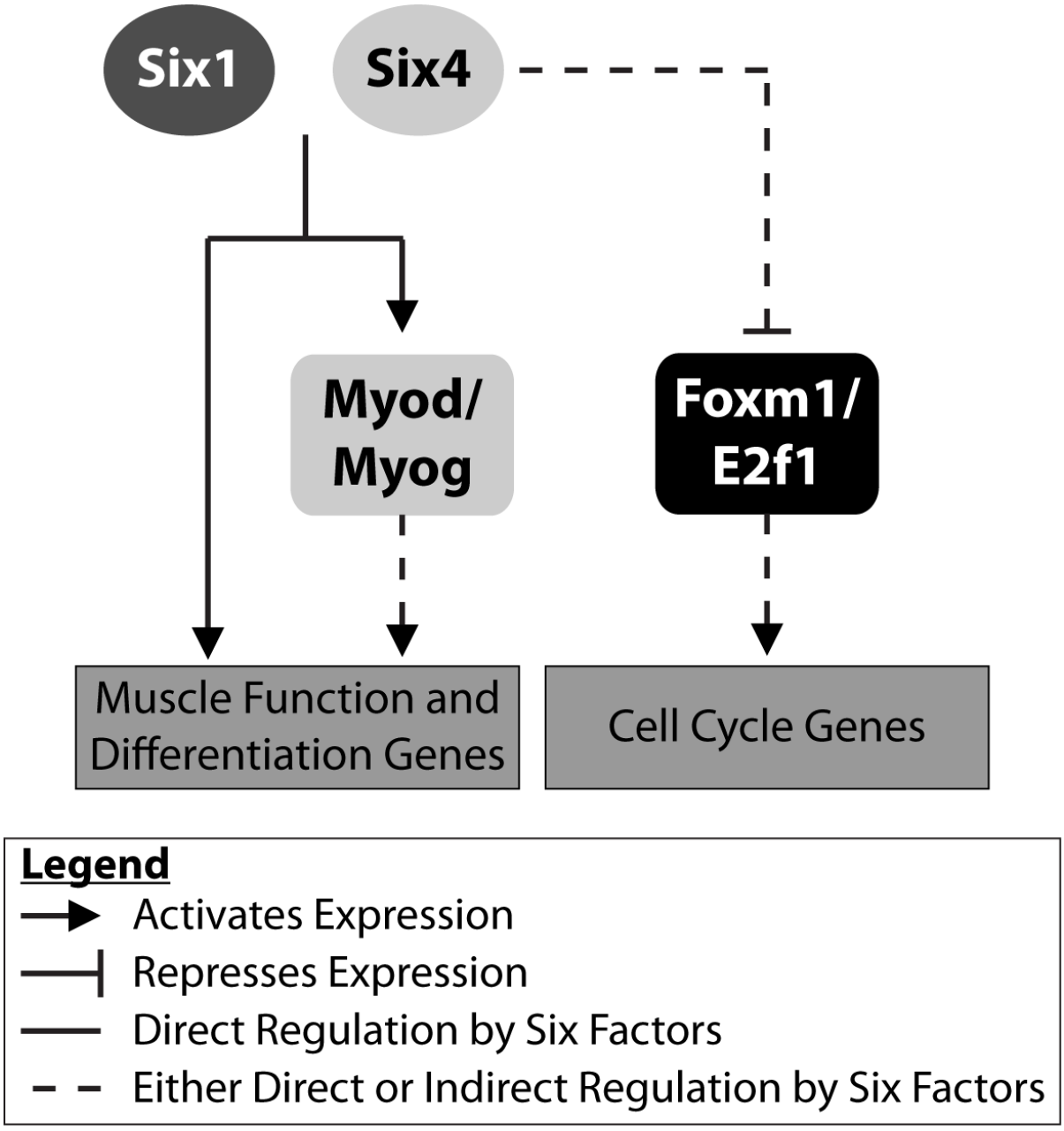

Figure 23 – Model of Six gene regulation and function during myoblast differentiation. Model illustrating the function of Six1 and Six4 during normal myoblast differentiation.

4.6. Further Characterizing the Role of Six1 and Six4 in Muscle Development

Although this project has focused on characterizing the role Six of TFs during terminal differentiation of myoblasts, other biological processes or systems may also play an important role in myogenesis. One such biological process includes changes in chromatin structure during terminal muscle differentiation. More precisely, histone post-translational modifications (PTM) have a large impact on gene expression since specific PTMs are associated with activation or repression of gene expression (Reviewed in Berger, 2007; Li et al., 2007). In line with the systems biology approach, the data presented in this work can be combined and analyzed with ChIP-Seq data of different histone marks. Such a study has already been performed to better understand the changes in the epigenetic landscape during C2C12 differentiation (Asp et al., 2011). It will be very informative to combine the data from the Six knock-down with the data from the epigenetic dynamics characterization study. It would allow us to associate dynamic changes in chromatin structure to gene activation and repression by Six1 and Six4 during myogenesis. By combining the data from the Six factors knock-down and the ChIP-on-Chip with the data from chromatin histone marks, it will lead to a better understanding of the molecular dynamics occurring during myogenesis.

Another factor to consider is the role of microRNA (miRNA) during skeletal myogenesis. miRNAs are generated from transcripts and result in small single stranded RNA (ssRNA) molecules of 22 nt in length that often adopt a hairpin shape (Reviewed in Chu and Rana, 2007; Kim et al., 2009). The miRNA functions like siRNA where the seed region targets the mRNA and degrades it, or inhibits translation of the transcript (Doench and Sharp, 2004). Recently, many different miRNAs have been shown to play an important role in skeletal myogenesis (Chen et al., 2006; Kim et al., 2006; Koning et al., 2011; Naguibneva et al.,

2006; Seok et al., 2011) and other biological processes (Chen et al., 2004; Giraldez et al., 2005; He et al., 2005; Lee et al., 1993b; Zhao et al., 2005) (Reviewed in Ge and Chen, 2011). One could potentially explore if any miRNAs are regulated by Six factors through miRNA profiling. The microarrays used in this study, only evaluate mRNAs from the mouse genome and not miRNAs. Additionally, it is also possible that Six factors regulate an uncharacterized miRNAs involved in muscle differentiation or cell cycle regulation. To find miRNAs regulating myogenesis, one could perform deep sequencing of short RNA molecules as a possible approach (Lu et al., 2005; Margulies et al., 2005; Morin et al., 2008). Alternatively, one could also perform miRNA profiling to identify small RNA molecules which are expressed or repressed during myogenesis and identify miRNAs which are essential for myogenesis. Another approach to uncover if miRNAs are indeed involved in the transcriptional regulation network with the Six factors, is to use computational methods to predict if genes regulated by either Six1, Six4 or both TF have an enrichment of miRNA seed sequences (Huang et al., 2007; Lai et al., 2003; Li et al., 2006; Nam et al., 2005). Additionally, these three techniques also lend themselves well to the systems biology approach, since they tackle the biological question using a holistic approach.

4.7. Chapter Conclusions

The work presented here illustrates the utility of a systems biology approach to solve a biological question. This method focuses on understanding an organism in its entirety which requires combining many different data sets from genome wide experiments. In this study, combining different data sets with computational techniques provided a better understanding the function of Six1 and Six4 with the MRFs. The systems approach also allowed the discovery of a novel function for Six4 during myogenesis. Nevertheless, this study has only combined a small number of systems biology data sets and has considered the system as a static model (single timepoint). To better understand the role of Six1 and Six4 during myogenesis, further work including collection of other type of data sets is still required. For example, investigating changes to miRNAs and chromatin marks of key genes in conjunction with the myogenic process as a dynamic process (i.e. changes in chromatin structure or gene regulated by Six factors over time) will provide a better understanding of the myogenic process.

Integration of these different data sets can possibly give invaluable insight into the transcriptional regulation network governing skeletal myogenesis which will eventually lead to development of therapies of MD patients. In fact, the data collected form this work can provide a better understanding of the molecular events that govern the myogenic process. This information can be used to control the proliferation and differentiation of myoblasts, both essential processes that need to be properly regulated for use in cell based therapies. For example, promoting self-renewal and preventing premature differentiation of muscle progenitor cells. Finally, the data suggest that Six4 is involved in the indirect regulation of the cell cycle process, raising the question of whether Six4 is involved in the development of

rhabdomyosarcomas, a form of cancer developing in muscle tissues. Further work is still required to better understand muscle diseases and uncovering the molecular mechanisms governing muscle development and regeneration are key for the development of therapies of affected MD patients.

5. Chapter 5 – Materials and Methods

5.1. Cell Culture

5.1.1. Tissue Cell Culture of C2C12 Myoblasts

The C2C12 (Yaffe and Saxel, 1977) cells from the American Type Culture Collection (ATCC) were grown in Growth Medium (GM) containing 88% Dulbecco's Modified Eagle's Medium (DMEM); 10% Fetal Bovine Serum (FBS); 2 mM Glutamine and 50 units of Penicillin/Streptomycin (P/S) until confluent. Once confluent, the cells were induced to differentiate by replacing with Differentiation Medium (DM) containing 96% DMEM; 2% Horse Serum; 2 mM Glutamine and 50 units of P/S. The cells were grown in a humidified water jacketed incubator at 37ºC with 5% CO_2.

5.1.2. Isolation and Culture of Primary Myoblasts from Mice

The isolation of primary myoblasts was described and characterized by Rando et al. (Rando and Blau, 1994) with modifications outlined below. The primary myoblasts were isolated from 90 day old female C57/B6 mice. Gastrocnemius, Tibialis Anterior and Quadricep muscles were pooled and digested with Collagenase I (Sigma) and Dispase II (Roche) at a final concentration of 625 μg/mL for 1.5 - 2.0 hours at 37°C. Cells were then diluted with DMEM and passed through a 70 μm nylon mesh filter (BD Falcon) to remove undigested connective tissues and large cell clumps. Cells were then rinsed twice with DMEM and resuspended in plating medium containing: 90% DMEM; 10% Donor Equine Serum and 5 ng/mL of bFGF (Peprotech). Cells were then pre-plated twice for 1h in a 10 cm tissue culture treated dish (Corning) to remove fast-attaching fibroblasts and non adhering cells were

finally transferred to Matrigel (BD Biosciences) coated plastic dishes and allowed to adhere to the plate for 48h.

Once the cells began proliferating 48h after isolation, cells were cultured in DMEM (American Type Culture Collection or ATCC) supplemented with 20% Fetal Bovine Serum (HyClone), 10% Donor Serum (HyClone) and 1% Penicillin/Streptomycin (HyClone). Growth Medium (GM) was supplemented with 10 ng/μL of Basic Fibroblast growth factor (bFGF) and 2 ng/μL of Basic Hepatocyte growth factor (bHGF) (Peprotech). The cells were grown in a humidified water jacketed incubator at 37ºC with 5% CO_2.

5.2. RNA interference and RNA work

5.2.1. siRNA Transfections

C2C12 or primary myoblasts cells were grown in 6 well plates (Corning) and were transfected with siRNA when the cells reached a confluency of 75 to 80%. Cells were then allowed to reach confluence and differentiate for 24h. For the immunofluorescence experiment, cells were treated as previously described, but to improve knock-down of target genes, cells were transfected with the same siRNA duplexes 24h after the start of differentiation. Cells were then allowed to continue differentiate for an additional 48h.

siRNA duplexes (Dharmacon) were resuspended and diluted as described by the manufacturer to a final concentration of 20 μM. siRNA sequences are available in . For 1 well of a 6 well plate, 5 μL of 20 μM siRNA (for double knock-down Six1 and Six4, 2.5 μL of each siRNA was used) was mixed with 45 μL of OptiMEM (Invitrogen) and the transfection reagent was prepared with 5 μL of Lipofectamine 2000 (Invitrogen) with 95 μL of OptiMEM. Both the siRNA and transfection reagent (50 and 100 μL each respectively)

were mixed together and incubated at room temperature for 20 minutes. Cells were then rinsed twice with Serum Free Medium (SFM) supplemented with L-Glutamine and 850 µL SFM was added to the cells. After incubation of the transfection reagent with the siRNA, 150 µL of the mix was added to each well. 4 hours later, the transfection mixture was replaced with GM without P/S.

5.2.2. RNA Extraction

RNA was extracted and purified with the Absolutely RNA® Miniprep Kit (Stratagene) following the manufacturer recommendations. The cells were rinsed twice with PBS and then lysed with 600 µL of lysis buffer with 0.7% of β-Mercaptoethanol. RNA was treated with DNaseI supplied with the kit. RNA isolation by Trizol reagent (Invitrogen) was performed according to the manufacturer's protocol. Concentration and quality of the extracted RNA were determined by using a Nano-Drop ND-1000 spectrophotometer (Nano-Drop). Quality of the RNA was determined by a non denaturing agarose gel electrophoresis.

5.2.3. Reverse-transcription PCR (RT-PCR)

1000 ng of RNA were submitted to reverse-transcription reaction using the SuperScript First-Strand Synthesis System for RT-PCR (Invitrogen) following the manufacturer's protocol. The reverse transcription was primed using random hexamers provided in the kit. After the reverse-transcription, 30 µL of 10 mM Tris pH 8.0 was added to the reaction mixture to bring the final volume to 50 µL.

5.2.4. Real-Time or Quantitative Reverse Transcription PCR (qRT-PCR)

Quantification of RNA transcript was done by SYBR green with ROX normalization method. The qPCR reactions were performed with a master mix containing for a 10 µL reaction: 1µL of 10X Qiagen HotStarTaq Buffer (Qiagen); 1 µL of 1.5% Triton X-100 (Fisher); 0.2 µL of dNTP (Fisher); 0.3 µL of ROX at concentration of 3 µM (Invitrogen); 1 µL of primer pairs at 5 uM (Operon); 0.05 µL of SYBR Green I used at a dilution of 1:500 in Dimethyl sulfoxide (DMSO) (Invitrogen) and 0.08 µL of HotStarTaq at 5 units/µL (Qiagen). 2 µL of 1:10 diluted cDNA was added to the qPCR master mixes. Reactions were performed on the MX3000P platform with the MxPro Software (Stratagene). Oligonucleotide sequences are provided in Table S 1. Primer pairs were validated by running a standard curve of a mixture of cDNA from all conditions in the experiment diluted to 1:1; 1:6; 1:36 and 1:216. Relative quantification was performed using the efficiency corrected method (Pfaffl, 2001). Genes were normalized using the Rps26 gene. The TATA Binding Protein (TBP) gene, assumed to be invariant in different experimental conditions, was also quantitated to ensure that normalization with Rps26 did not skew the results.

5.2.5. Microarray Gene Expression Profiling

To obtain cells suitable for gene expression profiling during a differentiation time course, the cells were grown in 10 cm tissue cell culture dishes (Corning). The cells were seeded at a density of 2.4 million cells in a 10 cm tissue culture dish. Cells were harvested at the myoblast stage (MB) at a confluency of 80%. When the confluency of the plate was 100%, the C2C12 cells were induced to differentiate in DM. Cells were harvested at start of differentiation, 24 and 96 hours after start of differentiation. Myotubes were selectively

harvested with a 1:20 dilution of trypsin (Invitrogen) with sterile phosphate buffered saline (PBS). Diluted trypsin is used here to specifically detach only myotubes from the tissue culture plate leaving the reserve cells behind.

Gene Expression Profiling was performed using the One-Color Microarray Gene Expression Platform from Agilent Technologies. The 4x44k Whole Mouse Genome Oligonucleotide microarray was used in the course of these experiments. To obtain labeled cDNA, the Quick Amp Labeling Kit, One Color kit was used (Agilent Technologies). Following the manufacturer's protocol, 300 ng of RNA was labelled and used for each microarray. Three replicates were obtained using 1 slide with 4 microarrays per replicate. One slide contained the following samples: cells at the Myoblast Stage (MB), T_0, T_{24}, and Myotubes (MT). For the siRNA treatments, four replicates were obtained one slide contained the non-specific siRNA (NS) treatment, the siRNA sequence targeting Six1; Six4; both or Myog. The slides were hybridized at 65ºC for about 17 hours. After washing the slides following the recommended protocol, the slides were scanned with the Agilent DNA Microarray Scanner at a resolution of 5 μm with extended dynamic range (XDR). After scanning, the intensities of the spots were extracted and preliminary normalization was performed using the program Feature Extraction Version 10.1.1 (Agilent Technologies). After extraction, gene expression was analyzed using GeneSpring GX Version 10.0 (Agilent Technologies). The data were log base 2 transformed and normalized using the quantiles algorithm. Microarray data can be downloaded from the Gene Expression Omnibus database under the accession number GSE19988.

5.3. Gene Expression Computational Analysis

5.3.1. Clustering Analysis (CLICK) and Heat-Maps

Probes retained for the cluster analysis were filtered with a change of at least 2-fold in either direction (higher than 1 or lower than -1, in log base 2) and ANOVA p-value < 0.05 corrected for multiple testing with the Benjamini-Hochberg algorithm. The probes were then analyzed with the CLICK clustering algorithm in the Expander5 software (Sharan et al., 2003; Sharan and Shamir, 2000) with the homogeneity set to default. Heat-maps were generated by using Cluster 3 (de Hoon et al., 2004) and Java TreeView (Saldanha, 2004). Since the goal was to identify changes in gene expression patterns in the siRNA treated cells, both the siRNA and the timecourse data sets were normalized separately to avoid the influences of large changes in gene expression from the time course data.

5.3.2. Gene Set Enrichment Analysis (GSEA)

All of the probes from the gene expression data from the siRNA knock-down of Six1 and Six4 were used in the Gene Set Enrichment Analysis (GSEA) software package (Mootha et al., 2003; Subramanian et al., 2005). Raw data from the microarrays were loaded into GSEA and ran with the following parameters: (1) Number of Permutations: 1000, (2) Permutation Type: Gene Set, (3) Enrichment Statistic: Weighted and (4) Metric for Ranking: Signal2Noise. Pair wise comparisons were Control vs. siSix1 and Control vs. siSix4. The entire molecular signature database (MsigDB) version 3.7 was used in the GSEA analysis.

5.3.3. Gene Ontology (GO)

Gene lists from the CLICK analysis were submitted to Gene Ontology Analysis with the default parameters of the website (http://david.abcc.ncifcrf.gov/). The background gene set was set to all genes of the mouse genome. The analysis used only functional annotation chart with the Gene Ontology Biological Processes Categories level 3, 4 and 5 (Huang da et al., 2009a, b).

5.3.4. Whole Genome rVista

Gene lists from the CLICK clustering analysis were submitted to the Whole Genome rVista website (http://genome.lbl.gov/vista/index.shtml). Using the Mouse transcription factor binding sites assembly (mm8), Gene Symbols were submitted and the software searched for transcription factors binding sites 5 kb upstream of genes (Dubchak and Ryaboy, 2006).

5.3.5. Box-Plots

Normalized gene expression data corresponding to 8 CLICK clusters were imported into R version 2.13.2 (R Development Core Team, 2011). Box plots were generated using the *boxplot()* function with the option *outline=FALSE*.

5.4. Protein Work

5.4.1. Western Blot Analysis

Proteins from the Absolutely RNA® Miniprep Kit (Stratagene) were extracted from the first flowthrough using acetone precipitation at -20ºC. The proteins were then resuspended in 100 µL of resolubilization solution (6M Urea; 0.1% SDS; 20 mM Tris pH 6.8). Proteins were

assayed using the BCA Protein Assay Kit™ (Pierce) and 20 μg of protein were loaded on a 10% SDS-PAGE as previously described (Laemmli, 1970). The proteins were transferred from the SDS-PAGE gel to a polyvinylidene fluoride (PVDF) membrane (Millipore) following previously described procedure (Burnette, 1981). Detection of the antibodies was done using a secondary antibody conjugated to a horse raddish peroxidase (HRP) at a dilution of 1:10000 coupled to an antibody recognizing rabbit or mouse antibodies (Fisher). Primary antibodies are listed in the following section.

5.4.2. Primary Antibodies

Antibody	Dilution	Source or Product Number	Manufacturer
Six1	1 : 250	Purified from immunized rabbits	Open Biosystems
Six4	1 : 1000	Purified from immunized rabbits	Open Biosystems
Beta-Tubulin	1 : 1000	Hybridoma Cells Clone 9E7	DHSB
Myogenin	1 : 1	Hybridoma Cells Clone F5D	DHSB
Anti-BrdU	1: 4000	Ascites Fluid from Clone G3G4	DHSB
MHC	1: 1	Hybridoma Cells Clone MF20	DHSB
Foxm1	1 : 25	SC-500	Santa Cruz

5.4.3. Western Blot Densitometry Quantification

X-ray films of western blots were digitized on a flat bed scanner using the transparency mode with a resolution of 300 dpi. Images were imported into Adobe Photoshop version 13.0 and saved as unaltered TIFF files. Densitometry measurements were performed with the ImageJ software (Abramoff et al., 2004). Densitometry values were measured and normalized across samples to the densitometry values of beta-tubulin and normalized to the control non-silencing siRNA, in knock-down experiments (set at 1).

5.5. Immunofluoresecence Work

5.5.1. Immunofluorescence of myotubes

The C2C12 cells were grown in slide flasks (Nunc) and were transfected with siRNA as described above. The cells were transfected again at T24 with siRNA to improve knock-down of target genes. The cells were fixed with -20ºC methanol for 5 minutes and air dried. Slides from the BrdU experiments were treated immediately after fixation with 4N HCl for 15 minutes at room temperature to denature the DNA. The HCl was removed by 3 washes with PBS 3 minutes each. The slides were blocked for 1h with PBST-BSA made with: PBS; 3% BSA; 0.1% Triton X-100 (Fisher). After blocking, the slides were incubated with the primary antibody: Myosin Heavy Chain – Clone MF20 from the Developmental Studies Hybridoma Bank (DSHB) for 2h diluted 1:1 in PBST-BSA; Anti-BrdU – Clone G3G4 (DSHB) for 1h diluted 1:4000 in PBST-BSA. After washing, the secondary antibody conjugated to FITC (Jackson) was incubated with the slides for 1h diluted in PBST-BSA solution at a 1:300 dilution. Once washed and dried, 3 drops of Prolong® Gold Antifade reagent with DAPI (Invitrogen) were added and the slides were sealed with cover slips. Images were taken on a Leica DMI6000 B epifluoresecnce microscope at a magnification of 10x. Images were acquired with the Velocity software version 5.0.2 (Improvision). Ten randomly selected fields were taken on each slide. The TIFF images of each color channel were merged and intensity adjusted using Adobe Photoshop version 13.0. Figures were created with Illustrator version 13.0 (Adobe).

5.5.2. BrdU Pulse Labeling

C2C12 cells were grown in slide flasks (Nunc) and treated with siRNA as described previously. 20 hours after the start of differentiation, cells were treated with 10 μM of BrdU for 4h. At the end of the pulse (24 hours in DM), cells were fixed with formalin. Denaturing of the DNA was done with 4N HCl for 10 minutes at room temperature. Samples were processed as described in the immunofluorescence section. Counting of nuclei and BrdU positive nuclei was automated with the use of Cell Profiler (Carpenter et al., 2006; Kamentsky et al., 2011).

6. References

Abramoff, M.D., Magalhaes, P.J., and Ram, S.J. (2004). Image Processing with ImageJ. Biophotonics International *11*, 36-42.

Agger, K., Cloos, P.A., Christensen, J., Pasini, D., Rose, S., Rappsilber, J., Issaeva, I., Canaani, E., Salcini, A.E., and Helin, K. (2007). UTX and JMJD3 are histone H3K27 demethylases involved in HOX gene regulation and development. Nature *449*, 731-734.

Alvarez, B., Martinez, A.C., Burgering, B.M., and Carrera, A.C. (2001). Forkhead transcription factors contribute to execution of the mitotic programme in mammals. Nature *413*, 744-747.

Amarzguioui, M., Rossi, J.J., and Kim, D. (2005). Approaches for chemically synthesized siRNA and vector-mediated RNAi. FEBS Lett *579*, 5974-5981.

Ando, Z., Sato, S., Ikeda, K., and Kawakami, K. (2005). Slc12a2 is a direct target of two closely related homeobox proteins, Six1 and Six4. FEBS J *272*, 3026-3041.

Andres, V., Cervera, M., and Mahdavi, V. (1995). Determination of the consensus binding site for MEF2 expressed in muscle and brain reveals tissue-specific sequence constraints. J Biol Chem *270*, 23246-23249.

Andres, V., and Walsh, K. (1996). Myogenin expression, cell cycle withdrawal, and phenotypic differentiation are temporally separable events that precede cell fusion upon myogenesis. J Cell Biol *132*, 657-666.

Asp, P., Acosta-Alvear, D., Tsikitis, M., van Oevelen, C., and Dynlacht, B.D. (2009). E2f3b plays an essential role in myogenic differentiation through isoform-specific gene regulation. Genes Dev *23*, 37-53.

Asp, P., Blum, R., Vethantham, V., Parisi, F., Micsinai, M., Cheng, J., Bowman, C., Kluger, Y., and Dynlacht, B.D. (2011). Genome-wide remodeling of the epigenetic landscape during myogenic differentiation. Proc Natl Acad Sci U S A *108*, E149-158.

Aziz, A., Liu, Q.C., and Dilworth, F.J. (2010). Regulating a master regulator: establishing tissue-specific gene expression in skeletal muscle. Epigenetics *5*, 691-695.

Bajard, L., Relaix, F., Lagha, M., Rocancourt, D., Daubas, P., and Buckingham, M.E. (2006). A novel genetic hierarchy functions during hypaxial myogenesis: Pax3 directly activates Myf5 in muscle progenitor cells in the limb. Genes Dev *20*, 2450-2464.

Banerji, J., Rusconi, S., and Schaffner, W. (1981). Expression of a beta-globin gene is enhanced by remote SV40 DNA sequences. Cell *27*, 299-308.

Barski, A., Cuddapah, S., Cui, K., Roh, T.Y., Schones, D.E., Wang, Z., Wei, G., Chepelev, I., and Zhao, K. (2007). High-resolution profiling of histone methylations in the human genome. Cell *129*, 823-837.

Bassaglia, Y., and Gautron, J. (1995). Fast and slow rat muscles degenerate and regenerate differently after whole crush injury. J Muscle Res Cell Motil *16*, 420-429.

Beauchamp, J.R., Heslop, L., Yu, D.S., Tajbakhsh, S., Kelly, R.G., Wernig, A., Buckingham, M.E., Partridge, T.A., and Zammit, P.S. (2000). Expression of CD34 and Myf5 defines the majority of quiescent adult skeletal muscle satellite cells. J Cell Biol *151*, 1221-1234.

Benhaddou, A., Keime, C., Ye, T., Morlon, A., Michel, I., Jost, B., Mengus, G., and Davidson, I. (2011). Transcription factor TEAD4 regulates expression of Myogenin and the unfolded protein response genes during C2C12 cell differentiation. Cell Death Differ.

Benjamini, Y., and Hochberg, Y. (1995). Controlling the False Discovery Rate - A Practical and Powerful Approach To Multiple Testing. Journal of the Royal Statistical Society Series B-Methodological *57*, 289-300.

Berger, S.L. (2007). The complex language of chromatin regulation during transcription. Nature *447*, 407-412.

Berkes, C.A., and Tapscott, S.J. (2005). MyoD and the transcriptional control of myogenesis. Semin Cell Dev Biol *16*, 585-595.

Bernstein, B.E., Mikkelsen, T.S., Xie, X., Kamal, M., Huebert, D.J., Cuff, J., Fry, B., Meissner, A., Wernig, M., Plath, K.*, et al.* (2006). A bivalent chromatin structure marks key developmental genes in embryonic stem cells. Cell *125*, 315-326.

Bernstein, E., Caudy, A.A., Hammond, S.M., and Hannon, G.J. (2001). Role for a bidentate ribonuclease in the initiation step of RNA interference. Nature *409*, 363-366.

Black, B.L., and Olson, E.N. (1998). Transcriptional control of muscle development by myocyte enhancer factor-2 (MEF2) proteins. Annu Rev Cell Dev Biol *14*, 167-196.

Blackwell, T.K., and Weintraub, H. (1990). Differences and similarities in DNA-binding preferences of MyoD and E2A protein complexes revealed by binding site selection. Science *250*, 1104-1110.

Bladt, F., Riethmacher, D., Isenmann, S., Aguzzi, A., and Birchmeier, C. (1995). Essential role for the c-met receptor in the migration of myogenic precursor cells into the limb bud. Nature *376*, 768-771.

Blais, A., Tsikitis, M., Acosta-Alvear, D., Sharan, R., Kluger, Y., and Dynlacht, B.D. (2005). An initial blueprint for myogenic differentiation. Genes Dev *19*, 553-569.

Bober, E., Franz, T., Arnold, H.H., Gruss, P., and Tremblay, P. (1994). Pax-3 is required for the development of limb muscles: a possible role for the migration of dermomyotomal muscle progenitor cells. Development *120*, 603-612.

Boyer, L.A., Plath, K., Zeitlinger, J., Brambrink, T., Medeiros, L.A., Lee, T.I., Levine, S.S., Wernig, M., Tajonar, A., Ray, M.K.*, et al.* (2006). Polycomb complexes repress developmental regulators in murine embryonic stem cells. Nature *441*, 349-353.

Bragdon, B., Thinakaran, S., Moseychuk, O., King, D., Young, K., Litchfield, D.W., Petersen, N.O., and Nohe, A. (2010). Casein kinase 2 beta-subunit is a regulator of bone morphogenetic protein 2 signaling. Biophys J *99*, 897-904.

Braun, T., Bober, E., Buschhausen-Denker, G., Kohtz, S., Grzeschik, K.H., and Arnold, H.H. (1989a). Differential expression of myogenic determination genes in muscle cells: possible autoactivation by the Myf gene products. EMBO J *8*, 3617-3625.

Braun, T., Buschhausen-Denker, G., Bober, E., Tannich, E., and Arnold, H.H. (1989b). A novel human muscle factor related to but distinct from MyoD1 induces myogenic conversion in 10T1/2 fibroblasts. EMBO J *8*, 701-709.

Braun, T., Rudnicki, M.A., Arnold, H.H., and Jaenisch, R. (1992). Targeted inactivation of the muscle regulatory gene Myf-5 results in abnormal rib development and perinatal death. Cell *71*, 369-382.

Brent, A.E., and Tabin, C.J. (2002). Developmental regulation of somite derivatives: muscle, cartilage and tendon. Curr Opin Genet Dev *12*, 548-557.

Brickey, W.J., and Greenleaf, A.L. (1995). Functional studies of the carboxy-terminal repeat domain of Drosophila RNA polymerase II in vivo. Genetics *140*, 599-613.

Bridge, A.J., Pebernard, S., Ducraux, A., Nicoulaz, A.L., and Iggo, R. (2003). Induction of an interferon response by RNAi vectors in mammalian cells. Nat Genet *34*, 263-264.

Brody, J.P., Williams, B.A., Wold, B.J., and Quake, S.R. (2002). Significance and statistical errors in the analysis of DNA microarray data. Proc Natl Acad Sci U S A *99*, 12975-12978.

Brown, C.E., Lechner, T., Howe, L., and Workman, J.L. (2000). The many HATs of transcription coactivators. Trends Biochem Sci *25*, 15-19.

Bryson-Richardson, R.J., and Currie, P.D. (2008). The genetics of vertebrate myogenesis. Nat Rev Genet *9*, 632-646.

Buckingham, M. (2001). Skeletal muscle formation in vertebrates. Curr Opin Genet Dev *11*, 440-448.

Buckingham, M., Bajard, L., Chang, T., Daubas, P., Hadchouel, J., Meilhac, S., Montarras, D., Rocancourt, D., and Relaix, F. (2003). The formation of skeletal muscle: from somite to limb. J Anat *202*, 59-68.

Burke, T.W., and Kadonaga, J.T. (1997). The downstream core promoter element, DPE, is conserved from Drosophila to humans and is recognized by TAFII60 of Drosophila. Genes Dev *11*, 3020-3031.

Burnette, W.N. (1981). "Western blotting": electrophoretic transfer of proteins from sodium dodecyl sulfate--polyacrylamide gels to unmodified nitrocellulose and radiographic detection with antibody and radioiodinated protein A. Anal Biochem *112*, 195-203.

Caldwell, C.J., Mattey, D.L., and Weller, R.O. (1990). Role of the basement membrane in the regeneration of skeletal muscle. Neuropathol Appl Neurobiol *16*, 225-238.

Cam, H., and Dynlacht, B.D. (2003). Emerging roles for E2F: beyond the G1/S transition and DNA replication. Cancer Cell *3*, 311-316.

Cao, Y., Kumar, R.M., Penn, B.H., Berkes, C.A., Kooperberg, C., Boyer, L.A., Young, R.A., and Tapscott, S.J. (2006). Global and gene-specific analyses show distinct roles for Myod and Myog at a common set of promoters. EMBO J *25*, 502-511.

Cao, Y., Yao, Z., Sarkar, D., Lawrence, M., Sanchez, G.J., Parker, M.H., MacQuarrie, K.L., Davison, J., Morgan, M.T., Ruzzo, W.L.*, et al.* (2010). Genome-wide MyoD binding in skeletal muscle cells: a potential for broad cellular reprogramming. Dev Cell *18*, 662-674.

Carlson, B.M. (1973). The regeneration of skeletal muscle. A review. Am J Anat *137*, 119-149.

Carpenter, A.E., Jones, T.R., Lamprecht, M.R., Clarke, C., Kang, I.H., Friman, O., Guertin, D.A., Chang, J.H., Lindquist, R.A., Moffat, J.*, et al.* (2006). CellProfiler: image analysis software for identifying and quantifying cell phenotypes. Genome Biol *7*, R100.

Carter, D., Chakalova, L., Osborne, C.S., Dai, Y.F., and Fraser, P. (2002). Long-range chromatin regulatory interactions in vivo. Nat Genet *32*, 623-626.

Carvajal, J.J., Cox, D., Summerbell, D., and Rigby, P.W. (2001). A BAC transgenic analysis of the Mrf4/Myf5 locus reveals interdigitated elements that control activation and maintenance of gene expression during muscle development. Development *128*, 1857-1868.

Chalkley, G.E., and Verrijzer, C.P. (1999). DNA binding site selection by RNA polymerase II TAFs: a TAF(II)250-TAF(II)150 complex recognizes the initiator. EMBO J *18*, 4835-4845.

Chamberlain, J.S. (2002). Gene therapy of muscular dystrophy. Hum Mol Genet *11*, 2355-2362.

Chen, C.Z., Li, L., Lodish, H.F., and Bartel, D.P. (2004). MicroRNAs modulate hematopoietic lineage differentiation. Science *303*, 83-86.

Chen, G.X., Jaradat, S.A., Banerjee, N., Tanaka, T.S., Ko, M.S.H., and Zhang, M.Q. (2002). Evaluation and comparison of clustering algorithms in analyzing es cell gene expression data. Stat Sin *12*, 241-262.

Chen, J.F., Mandel, E.M., Thomson, J.M., Wu, Q., Callis, T.E., Hammond, S.M., Conlon, F.L., and Wang, D.Z. (2006). The role of microRNA-1 and microRNA-133 in skeletal muscle proliferation and differentiation. Nat Genet *38*, 228-233.

Chen, R., Amoui, M., Zhang, Z., and Mardon, G. (1997). Dachshund and eyes absent proteins form a complex and function synergistically to induce ectopic eye development in Drosophila. Cell *91*, 893-903.

Cheung, P., and Lau, P. (2005). Epigenetic regulation by histone methylation and histone variants. Mol Endocrinol *19*, 563-573.

Cheyette, B.N., Green, P.J., Martin, K., Garren, H., Hartenstein, V., and Zipursky, S.L. (1994). The Drosophila sine oculis locus encodes a homeodomain-containing protein required for the development of the entire visual system. Neuron *12*, 977-996.

Chiogna, M., Massa, M.S., Risso, D., and Romualdi, C. (2009). A comparison on effects of normalisations in the detection of differentially expressed genes. BMC Bioinformatics *10*, 61.

Chong, L., and Ray, L.B. (2002). Whole-istic Biology. Science *295*, 1661.

Chu, C.Y., and Rana, T.M. (2007). Small RNAs: regulators and guardians of the genome. J Cell Physiol *213*, 412-419.

Chuaqui, R.F., Bonner, R.F., Best, C.J., Gillespie, J.W., Flaig, M.J., Hewitt, S.M., Phillips, J.L., Krizman, D.B., Tangrea, M.A., Ahram, M.*, et al.* (2002). Post-analysis follow-up and validation of microarray experiments. Nat Genet *32 Suppl*, 509-514.

Clevers, H. (2006). Wnt/beta-catenin signaling in development and disease. Cell *127*, 469-480.

Conaway, R.C., and Conaway, J.W. (2011). Function and regulation of the Mediator complex. Curr Opin Genet Dev *21*, 225-230.

Conboy, I.M., and Rando, T.A. (2002). The regulation of Notch signaling controls satellite cell activation and cell fate determination in postnatal myogenesis. Dev Cell *3*, 397-409.

Corbi, N., Di Padova, M., De Angelis, R., Bruno, T., Libri, V., Iezzi, S., Floridi, A., Fanciulli, M., and Passananti, C. (2002). The alpha-like RNA polymerase II core subunit 3 (RPB3) is involved in tissue-specific transcription and muscle differentiation via interaction with the myogenic factor myogenin. FASEB J *16*, 1639-1641.

Cosma, M.P. (2002). Ordered recruitment: gene-specific mechanism of transcription activation. Mol Cell *10*, 227-236.

Cossu, G., and Sampaolesi, M. (2004). New therapies for muscular dystrophy: cautious optimism. Trends Mol Med *10*, 516-520.

Cui, C.B., Cooper, L.F., Yang, X., Karsenty, G., and Aukhil, I. (2003). Transcriptional coactivation of bone-specific transcription factor Cbfa1 by TAZ. Mol Cell Biol *23*, 1004-1013.

Dalma-Weiszhausz, D.D., Warrington, J., Tanimoto, E.Y., and Miyada, C.G. (2006). The Affymetrix GeneChip® Platform: An Overview. In Methods in Enzymology, K. Alan, and O. Brian, eds. (Academic Press), pp. 3-28.

Davis, R.L., Weintraub, H., and Lassar, A.B. (1987). Expression of a single transfected cDNA converts fibroblasts to myoblasts. Cell *51*, 987-1000.

de Hoon, M.J., Imoto, S., Nolan, J., and Miyano, S. (2004). Open source clustering software. Bioinformatics *20*, 1453-1454.

de Laat, W., and Grosveld, F. (2003). Spatial organization of gene expression: the active chromatin hub. Chromosome Res *11*, 447-459.

Dharmadi, Y., and Gonzalez, R. (2004). DNA microarrays: experimental issues, data analysis, and application to bacterial systems. Biotechnol Prog *20*, 1309-1324.

Dietrich, S., Abou-Rebyeh, F., Brohmann, H., Bladt, F., Sonnenberg-Riethmacher, E., Yamaai, T., Lumsden, A., Brand-Saberi, B., and Birchmeier, C. (1999). The role of SF/HGF and c-Met in the development of skeletal muscle. Development *126*, 1621-1629.

Doench, J.G., and Sharp, P.A. (2004). Specificity of microRNA target selection in translational repression. Genes Dev *18*, 504-511.

Draghici, S., Khatri, P., Eklund, A.C., and Szallasi, Z. (2006). Reliability and reproducibility issues in DNA microarray measurements. Trends Genet *22*, 101-109.

Dubchak, I., and Ryaboy, D.V. (2006). VISTA family of computational tools for comparative analysis of DNA sequences and whole genomes. Methods Mol Biol *338*, 69-89.

Dudoit, S., Shaffer, J.P., and Boldrick, J.C. (2003). Multiple hypothesis testing in microarray experiments. Statistical Science *18*, 71-103.

Dynlacht, B.D., Hoey, T., and Tjian, R. (1991). Isolation of coactivators associated with the TATA-binding protein that mediate transcriptional activation. Cell *66*, 563-576.

Dyson, N. (1998). The regulation of E2F by pRB-family proteins. Genes Dev *12*, 2245-2262.

Edmondson, D.G., and Olson, E.N. (1989). A gene with homology to the myc similarity region of MyoD1 is expressed during myogenesis and is sufficient to activate the muscle differentiation program. Genes Dev *3*, 628-640.

Eisen, M.B., Spellman, P.T., Brown, P.O., and Botstein, D. (1998). Cluster analysis and display of genome-wide expression patterns. Proc Natl Acad Sci U S A *95*, 14863-14868.

Elbashir, S.M., Martinez, J., Patkaniowska, A., Lendeckel, W., and Tuschl, T. (2001). Functional anatomy of siRNAs for mediating efficient RNAi in Drosophila melanogaster embryo lysate. EMBO J *20*, 6877-6888.

Elledge, S.J., and Harper, J.W. (1994). Cdk inhibitors: on the threshold of checkpoints and development. Curr Opin Cell Biol *6*, 847-852.

Elnitski, L., Riemer, C., Burhans, R., Hardison, R., and Miller, W. (2005). MultiPipMaker: comparative alignment server for multiple DNA sequences. Curr Protoc Bioinformatics *Chapter 10*, Unit10 14.

Emery, A.E. (1991). Population frequencies of inherited neuromuscular diseases--a world survey. Neuromuscul Disord *1*, 19-29.

Emery, A.E. (2002). The muscular dystrophies. Lancet *359*, 687-695.

Everitt, B.S., Landau, S., Leese, M., and Stahl, D. (2011). Cluster Analysis, 5th edn (Chichester, John Wiley & Sons, Ltd).

Fan, J.B., Gunderson, K.L., Bibikova, M., Yeakley, J.M., Chen, J., Wickham Garcia, E., Lebruska, L.L., Laurent, M., Shen, R., and Barker, D. (2006). Illumina Universal Bead Arrays. In Methods in Enzymology, K. Alan, and O. Brian, eds. (Academic Press), pp. 57-73.

Ferrari, G., Cusella-De Angelis, G., Coletta, M., Paolucci, E., Stornaiuolo, A., Cossu, G., and Mavilio, F. (1998). Muscle regeneration by bone marrow-derived myogenic progenitors. Science *279*, 1528-1530.

Fickett, J.W. (1996a). Coordinate positioning of MEF2 and myogenin binding sites. Gene *172*, GC19-32.

Fickett, J.W. (1996b). Quantitative discrimination of MEF2 sites. Mol Cell Biol *16*, 437-441.

Fink, E., Fortin, D., Serrurier, B., Ventura-Clapier, R., and Bigard, A.X. (2003). Recovery of contractile and metabolic phenotypes in regenerating slow muscle after notexin-induced or crush injury. J Muscle Res Cell Motil *24*, 421-429.

Fire, A., Xu, S., Montgomery, M.K., Kostas, S.A., Driver, S.E., and Mello, C.C. (1998). Potent and specific genetic interference by double-stranded RNA in Caenorhabditis elegans. Nature *391*, 806-811.

Fischle, W., Wang, Y., Jacobs, S.A., Kim, Y., Allis, C.D., and Khorasanizadeh, S. (2003). Molecular basis for the discrimination of repressive methyl-lysine marks in histone H3 by Polycomb and HP1 chromodomains. Genes Dev *17*, 1870-1881.

Fletcher, J.E., and Jiang, M.S. (1993). Possible mechanisms of action of cobra snake venom cardiotoxins and bee venom melittin. Toxicon *31*, 669-695.

Forget, D., Robert, F., Grondin, G., Burton, Z.F., Greenblatt, J., and Coulombe, B. (1997). RAP74 induces promoter contacts by RNA polymerase II upstream and downstream of a DNA bend centered on the TATA box. Proc Natl Acad Sci U S A *94*, 7150-7155.

Fougerousse, F., Durand, M., Lopez, S., Suel, L., Demignon, J., Thornton, C., Ozaki, H., Kawakami, K., Barbet, P., Beckmann, J.S.*, et al.* (2002). Six and Eya expression during human somitogenesis and MyoD gene family activation. J Muscle Res Cell Motil *23*, 255-264.

Franz, T., Kothary, R., Surani, M.A., Halata, Z., and Grim, M. (1993). The Splotch mutation interferes with muscle development in the limbs. Anat Embryol (Berl) *187*, 153-160.

Frazer, K.A., Pachter, L., Poliakov, A., Rubin, E.M., and Dubchak, I. (2004). VISTA: computational tools for comparative genomics. Nucleic Acids Res *32*, W273-279.

Fuda, N.J., Ardehali, M.B., and Lis, J.T. (2009). Defining mechanisms that regulate RNA polymerase II transcription in vivo. Nature *461*, 186-192.

Ge, Y., and Chen, J. (2011). MicroRNAs in skeletal myogenesis. Cell Cycle *10*, 441-448.

Gehring, W.J., Affolter, M., and Burglin, T. (1994). Homeodomain proteins. Annu Rev Biochem *63*, 487-526.

Geiger, H., True, J.M., Grimes, B., Carroll, E.J., Fleischman, R.A., and Van Zant, G. (2002). Analysis of the hematopoietic potential of muscle-derived cells in mice. Blood *100*, 721-723.

Geller, S.C., Gregg, J.P., Hagerman, P., and Rocke, D.M. (2003). Transformation and normalization of oligonucleotide microarray data. Bioinformatics *19*, 1817-1823.

Gilbert, S.F. (2000). Developmental Biology, 6th edn (Sunderland MA, Sinauer Associates).

Gillies, S.D., Morrison, S.L., Oi, V.T., and Tonegawa, S. (1983). A tissue-specific transcription enhancer element is located in the major intron of a rearranged immunoglobulin heavy chain gene. Cell *33*, 717-728.

Gilroy, J., Cahalan, J.L., Berman, R., and Newman, M. (1963). Cardiac and pulmonary complications in Duchenne's progressive muscular dystrophy. Circulation *27*, 484-493.

Giordani, J., Bajard, L., Demignon, J., Daubas, P., Buckingham, M., and Maire, P. (2007). Six proteins regulate the activation of Myf5 expression in embryonic mouse limbs. Proc Natl Acad Sci U S A *104*, 11310-11315.

Giraldez, A.J., Cinalli, R.M., Glasner, M.E., Enright, A.J., Thomson, J.M., Baskerville, S., Hammond, S.M., Bartel, D.P., and Schier, A.F. (2005). MicroRNAs regulate brain morphogenesis in zebrafish. Science *308*, 833-838.

Gossett, L.A., Kelvin, D.J., Sternberg, E.A., and Olson, E.N. (1989). A new myocyte-specific enhancer-binding factor that recognizes a conserved element associated with multiple muscle-specific genes. Mol Cell Biol *9*, 5022-5033.

Goulding, M., Lumsden, A., and Paquette, A.J. (1994). Regulation of Pax-3 expression in the dermomyotome and its role in muscle development. Development *120*, 957-971.

Grifone, R., Demignon, J., Houbron, C., Souil, E., Niro, C., Seller, M.J., Hamard, G., and Maire, P. (2005). Six1 and Six4 homeoproteins are required for Pax3 and Mrf expression during myogenesis in the mouse embryo. Development *132*, 2235-2249.

Grimm, D., Streetz, K.L., Jopling, C.L., Storm, T.A., Pandey, K., Davis, C.R., Marion, P., Salazar, F., and Kay, M.A. (2006). Fatality in mice due to oversaturation of cellular microRNA/short hairpin RNA pathways. Nature *441*, 537-541.

Guang, S., Bochner, A.F., Burkhart, K.B., Burton, N., Pavelec, D.M., and Kennedy, S. (2010). Small regulatory RNAs inhibit RNA polymerase II during the elongation phase of transcription. Nature *465*, 1097-1101.

Guo, C., Sun, Y., Zhou, B., Adam, R.M., Li, X., Pu, W.T., Morrow, B.E., and Moon, A. (2011). A Tbx1-Six1/Eya1-Fgf8 genetic pathway controls mammalian cardiovascular and craniofacial morphogenesis. J Clin Invest *121*, 1585-1595.

Guo, H., Ingolia, N.T., Weissman, J.S., and Bartel, D.P. (2010). Mammalian microRNAs predominantly act to decrease target mRNA levels. Nature *466*, 835-840.

Guo, S., and Kemphues, K.J. (1995). par-1, a gene required for establishing polarity in C. elegans embryos, encodes a putative Ser/Thr kinase that is asymmetrically distributed. Cell *81*, 611-620.

Gusnanto, A., Calza, S., and Pawitan, Y. (2007). Identification of differentially expressed genes and false discovery rate in microarray studies. Curr Opin Lipidol *18*, 187-193.

Gutierrez, J.M., and Ownby, C.L. (2003). Skeletal muscle degeneration induced by venom phospholipases A2: insights into the mechanisms of local and systemic myotoxicity. Toxicon *42*, 915-931.

Hadchouel, J., Tajbakhsh, S., Primig, M., Chang, T.H., Daubas, P., Rocancourt, D., and Buckingham, M. (2000). Modular long-range regulation of Myf5 reveals unexpected heterogeneity between skeletal muscles in the mouse embryo. Development *127*, 4455-4467.

Hahn, S. (2004). Structure and mechanism of the RNA polymerase II transcription machinery. Nat Struct Mol Biol *11*, 394-403.

Hammond, S.M., Bernstein, E., Beach, D., and Hannon, G.J. (2000). An RNA-directed nuclease mediates post-transcriptional gene silencing in Drosophila cells. Nature *404*, 293-296.

Han, M., and Grunstein, M. (1988). Nucleosome loss activates yeast downstream promoters in vivo. Cell *55*, 1137-1145.

Hannon, G.J. (2002). RNA interference. Nature *418*, 244-251.

Harbour, J.W., and Dean, D.C. (2000). The Rb/E2F pathway: expanding roles and emerging paradigms. Genes Dev *14*, 2393-2409.

Hardiman, G. (2004). Microarray platforms--comparisons and contrasts. Pharmacogenomics *5*, 487-502.

Harris, J.B. (2003). Myotoxic phospholipases A2 and the regeneration of skeletal muscles. Toxicon *42*, 933-945.

Harris, J.B., Johnson, M.A., and Karlsson, E. (1975). Pathological responses of rat skeletal muscle to a single subcutaneous injection of a toxin isolated from the venom of the Australian tiger snake, Notechis scutatus scutatus. Clinical and Experimental Pharmacology and Physiology *2*, 383-404.

Hartigan, J.A. (1975). Clustering Algorithms (New York, Wiley).

Hartigan, J.A., and Wong, M.A. (1979). Algorithm AS 136: A K-Means Clustering Algorithm. Journal of the Royal Statistical Society Series C (Applied Statistics) *28*, 100-108.

Hartuv, E., and Shamir, R. (2000). A clustering algorithm based on graph connectivity. Inf Process Lett *76*, 175-181.

Hasty, P., Bradley, A., Morris, J.H., Edmondson, D.G., Venuti, J.M., Olson, E.N., and Klein, W.H. (1993). Muscle deficiency and neonatal death in mice with a targeted mutation in the myogenin gene. Nature *364*, 501-506.

Hayashi, M., Maeda, S., Aburatani, H., Kitamura, K., Miyoshi, H., Miyazono, K., and Imamura, T. (2008). Pitx2 prevents osteoblastic transdifferentiation of myoblasts by bone morphogenetic proteins. J Biol Chem *283*, 565-571.

He, L., Thomson, J.M., Hemann, M.T., Hernando-Monge, E., Mu, D., Goodson, S., Powers, S., Cordon-Cardo, C., Lowe, S.W., Hannon, G.J.*, et al.* (2005). A microRNA polycistron as a potential human oncogene. Nature *435*, 828-833.

Heanue, T.A., Reshef, R., Davis, R.J., Mardon, G., Oliver, G., Tomarev, S., Lassar, A.B., and Tabin, C.J. (1999). Synergistic regulation of vertebrate muscle development by Dach2, Eya2, and Six1, homologs of genes required for Drosophila eye formation. Genes Dev *13*, 3231-3243.

Heintzman, N.D., Stuart, R.K., Hon, G., Fu, Y., Ching, C.W., Hawkins, R.D., Barrera, L.O., Van Calcar, S., Qu, C., Ching, K.A.*, et al.* (2007). Distinct and predictive chromatin signatures of transcriptional promoters and enhancers in the human genome. Nat Genet *39*, 311-318.

Helin, K., Wu, C.L., Fattaey, A.R., Lees, J.A., Dynlacht, B.D., Ngwu, C., and Harlow, E. (1993). Heterodimerization of the transcription factors E2F-1 and DP-1 leads to cooperative trans-activation. Genes Dev *7*, 1850-1861.

Hirose, Y., and Manley, J.L. (1998). RNA polymerase II is an essential mRNA polyadenylation factor. Nature *395*, 93-96.

Hoheisel, J.D. (2006). Microarray technology: beyond transcript profiling and genotype analysis. Nat Rev Genet *7*, 200-210.

Huang da, W., Sherman, B.T., and Lempicki, R.A. (2009a). Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. Nucleic Acids Res *37*, 1-13.

Huang da, W., Sherman, B.T., and Lempicki, R.A. (2009b). Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat Protoc *4*, 44-57.

Huang, R., Zhi, Q., Schmidt, C., Wilting, J., Brand-Saberi, B., and Christ, B. (2000). Sclerotomal origin of the ribs. Development *127*, 527-532.

Huang, T.H., Fan, B., Rothschild, M.F., Hu, Z.L., Li, K., and Zhao, S.H. (2007). MiRFinder: an improved approach and software implementation for genome-wide fast microRNA precursor scans. BMC Bioinformatics *8*, 341.

Hubner, K., Sahle, S., and Kummer, U. (2011). Applications and trends in systems biology in biochemistry. FEBS J *278*, 2767-2857.

Iezzi, S., Di Padova, M., Serra, C., Caretti, G., Simone, C., Maklan, E., Minetti, G., Zhao, P., Hoffman, E.P., Puri, P.L.*, et al.* (2004). Deacetylase inhibitors increase muscle cell size by promoting myoblast recruitment and fusion through induction of follistatin. Dev Cell *6*, 673-684.

Ikeda, K., Watanabe, Y., Ohto, H., and Kawakami, K. (2002). Molecular interaction and synergistic activation of a promoter by Six, Eya, and Dach proteins mediated through CREB binding protein. Mol Cell Biol *22*, 6759-6766.

Jackson, K.A., Mi, T., and Goodell, M.A. (1999). Hematopoietic potential of stem cells isolated from murine skeletal muscle. Proc Natl Acad Sci U S A *96*, 14482-14486.

Jarve, A., Muller, J., Kim, I.H., Rohr, K., MacLean, C., Fricker, G., Massing, U., Eberle, F., Dalpke, A., Fischer, R.*, et al.* (2007). Surveillance of siRNA integrity by FRET imaging. Nucleic Acids Res *35*, e124.

Jarvinen, M. (1976). Healing of a crush injury in rat striated muscle. 3. A micro-angiographical study of the effect of early mobilization and immobilization on capillary ingrowth. Acta Pathol Microbiol Scand A *84*, 85-94.

Jean, D., Bernier, G., and Gruss, P. (1999). Six6 (Optx2) is a novel murine Six3-related homeobox gene that demarcates the presumptive pituitary/hypothalamic axis and the ventral optic stalk. Mech Dev *84*, 31-40.

Ji, H., Vokes, S.A., and Wong, W.H. (2006). A comparative analysis of genome-wide chromatin immunoprecipitation data for mammalian transcription factors. Nucleic Acids Res *34*, e146.

Ji, H., and Wong, W.H. (2005). TileMap: create chromosomal map of tiling array hybridizations. Bioinformatics *21*, 3629-3636.

Kagey, M.H., Newman, J.J., Bilodeau, S., Zhan, Y., Orlando, D.A., van Berkum, N.L., Ebmeier, C.C., Goossens, J., Rahl, P.B., Levine, S.S.*, et al.* (2010). Mediator and cohesin connect gene expression and chromatin architecture. Nature *467*, 430-435.

Kamentsky, L., Jones, T.R., Fraser, A., Bray, M.A., Logan, D.J., Madden, K.L., Ljosa, V., Rueden, C., Eliceiri, K.W., and Carpenter, A.E. (2011). Improved structure, function and compatibility for CellProfiler: modular high-throughput image analysis software. Bioinformatics *27*, 1179-1180.

Kassar-Duchossoy, L., Gayraud-Morel, B., Gomes, D., Rocancourt, D., Buckingham, M., Shinin, V., and Tajbakhsh, S. (2004). Mrf4 determines skeletal muscle identity in Myf5:Myod double-mutant mice. Nature *431*, 466-471.

Katagiri, T., Yamaguchi, A., Komaki, M., Abe, E., Takahashi, N., Ikeda, T., Rosen, V., Wozney, J.M., Fujisawa-Sehara, A., and Suda, T. (1994). Bone morphogenetic protein-2 converts the differentiation pathway of C2C12 myoblasts into the osteoblast lineage. J Cell Biol *127*, 1755-1766.

Kawada, H., and Ogawa, M. (2001). Bone marrow origin of hematopoietic progenitors and stem cells in murine muscle. Blood *98*, 2008-2013.

Kawakami, K., Ohto, H., Ikeda, K., and Roeder, R.G. (1996a). Structure, function and expression of a murine homeobox protein AREC3, a homologue of Drosophila sine oculis gene product, and implication in development. Nucleic Acids Res *24*, 303-310.

Kawakami, K., Ohto, H., Takizawa, T., and Saito, T. (1996b). Identification and expression of six family genes in mouse retina. FEBS Lett *393*, 259-263.

Kawakami, K., Sato, S., Ozaki, H., and Ikeda, K. (2000). Six family genes--structure and function as transcription factors and their roles in development. Bioessays *22*, 616-626.

Kayne, P.S., Kim, U.J., Han, M., Mullen, J.R., Yoshizaki, F., and Grunstein, M. (1988). Extremely conserved histone H4 N terminus is dispensable for growth but essential for repressing the silent mating loci in yeast. Cell *55*, 27-39.

Kennerdell, J.R., and Carthew, R.W. (1998). Use of dsRNA-mediated genetic interference to demonstrate that frizzled and frizzled 2 act in the wingless pathway. Cell *95*, 1017-1026.

Kenyon, K.L., Li, D.J., Clouser, C., Tran, S., and Pignoni, F. (2005). Fly SIX-type homeodomain proteins Sine oculis and Optix partner with different cofactors during eye development. Dev Dyn *234*, 497-504.

Kim, H.K., Lee, Y.S., Sivaprasad, U., Malhotra, A., and Dutta, A. (2006). Muscle-specific microRNA miR-206 promotes muscle differentiation. J Cell Biol *174*, 677-687.

Kim, T.H., Barrera, L.O., Zheng, M., Qu, C., Singer, M.A., Richmond, T.A., Wu, Y., Green, R.D., and Ren, B. (2005). A high-resolution map of active promoters in the human genome. Nature *436*, 876-880.

Kim, T.K., Lagrange, T., Wang, Y.H., Griffith, J.D., Reinberg, D., and Ebright, R.H. (1997). Trajectory of DNA in the RNA polymerase II transcription preinitiation complex. Proc Natl Acad Sci U S A *94*, 12268-12273.

Kim, V.N., Han, J., and Siomi, M.C. (2009). Biogenesis of small RNAs in animals. Nat Rev Mol Cell Biol *10*, 126-139.

Kissinger, C.R., Liu, B.S., Martin-Blanco, E., Kornberg, T.B., and Pabo, C.O. (1990). Crystal structure of an engrailed homeodomain-DNA complex at 2.8 A resolution: a framework for understanding homeodomain-DNA interactions. Cell *63*, 579-590.

Kitano, H. (2002). Systems Biology: A Brief Overview. Science *295*, 1662-1664.

Klesert, T.R., Cho, D.H., Clark, J.I., Maylie, J., Adelman, J., Snider, L., Yuen, E.C., Soriano, P., and Tapscott, S.J. (2000). Mice deficient in Six5 develop cataracts: implications for myotonic dystrophy. Nat Genet *25*, 105-109.

Knapp, J.R., Davie, J.K., Myer, A., Meadows, E., Olson, E.N., and Klein, W.H. (2006). Loss of myogenin in postnatal life leads to normal skeletal muscle but reduced body size. Development *133*, 601-610.

Knezetic, J.A., and Luse, D.S. (1986). The presence of nucleosomes on a DNA template prevents initiation by RNA polymerase II in vitro. Cell *45*, 95-104.

Kobayashi, M., Nishikawa, K., Suzuki, T., and Yamamoto, M. (2001). The homeobox protein Six3 interacts with the Groucho corepressor and acts as a transcriptional repressor in eye and forebrain formation. Dev Biol *232*, 315-326.

Koning, M., Werker, P., Bank, R.A., and Harmsen, M.C. (2011). MicroRNA-1 and microRNA-206 improve differentiation potential of human satellite cells: A novel approach for tissue engineering of skeletal muscle. Tissue Eng Part A.

Korver, W., Roose, J., and Clevers, H. (1997). The winged-helix transcription factor Trident is expressed in cycling cells. Nucleic Acids Res *25*, 1715-1719.

Kouzarides, T. (2007). Chromatin modifications and their function. Cell *128*, 693-705.

Krupczak-Hollis, K., Wang, X., Kalinichenko, V.V., Gusarova, G.A., Wang, I.C., Dennewitz, M.B., Yoder, H.M., Kiyokawa, H., Kaestner, K.H., and Costa, R.H. (2004). The mouse Forkhead Box m1 transcription factor is essential for hepatoblast mitosis and development of intrahepatic bile ducts and vessels during liver morphogenesis. Dev Biol *276*, 74-88.

Kumar, J.P. (2009). The sine oculis homeobox (SIX) family of transcription factors as regulators of development and disease. Cell Mol Life Sci *66*, 565-583.

Kutach, A.K., and Kadonaga, J.T. (2000). The downstream promoter element DPE appears to be as widely used as the TATA box in Drosophila core promoters. Mol Cell Biol *20*, 4754-4764.

Laclef, C., Hamard, G., Demignon, J., Souil, E., Houbron, C., and Maire, P. (2003). Altered myogenesis in Six1-deficient mice. Development *130*, 2239-2252.

Laemmli, U.K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature *227*, 680-685.

Lagha, M., Sato, T., Bajard, L., Daubas, P., Esner, M., Montarras, D., Relaix, F., and Buckingham, M. (2008). Regulation of skeletal muscle stem cell behavior by Pax3 and Pax7. Cold Spring Harb Symp Quant Biol *73*, 307-315.

Lagrange, T., Kapanidis, A.N., Tang, H., Reinberg, D., and Ebright, R.H. (1998). New core promoter element in RNA polymerase II-dependent transcription: sequence-specific DNA binding by transcription factor IIB. Genes Dev *12*, 34-44.

Lai, E.C., Tomancak, P., Williams, R.W., and Rubin, G.M. (2003). Computational identification of Drosophila microRNA genes. Genome Biol *4*, R42.

Lan, F., Bayliss, P.E., Rinn, J.L., Whetstine, J.R., Wang, J.K., Chen, S., Iwase, S., Alpatov, R., Issaeva, I., Canaani, E*., et al.* (2007). A histone H3 lysine 27 demethylase regulates animal posterior development. Nature *449*, 689-694.

Laoukili, J., Alvarez, M., Meijer, L.A., Stahl, M., Mohammed, S., Kleij, L., Heck, A.J., and Medema, R.H. (2008). Activation of FoxM1 during G2 requires cyclin A/Cdk-dependent relief of autorepression by the FoxM1 N-terminal domain. Mol Cell Biol *28*, 3076-3087.

Laoukili, J., Kooistra, M.R., Bras, A., Kauw, J., Kerkhoven, R.M., Morrison, A., Clevers, H., and Medema, R.H. (2005). FoxM1 is required for execution of the mitotic programme and chromosome stability. Nat Cell Biol *7*, 126-136.

Larkin, M.A., Blackshields, G., Brown, N.P., Chenna, R., McGettigan, P.A., McWilliam, H., Valentin, F., Wallace, I.M., Wilm, A., Lopez, R*., et al.* (2007). Clustal W and Clustal X version 2.0. Bioinformatics *23*, 2947-2948.

Lassar, A.B., Davis, R.L., Wright, W.E., Kadesch, T., Murre, C., Voronova, A., Baltimore, D., and Weintraub, H. (1991). Functional activity of myogenic HLH proteins requires hetero-oligomerization with E12/E47-like proteins in vivo. Cell *66*, 305-315.

Lee, D.Y., Hayes, J.J., Pruss, D., and Wolffe, A.P. (1993a). A positive role for histone acetylation in transcription factor access to nucleosomal DNA. Cell *72*, 73-84.

Lee, H.K., Hsu, A.K., Sajdak, J., Qin, J., and Pavlidis, P. (2004). Coexpression analysis of human genes across many microarray data sets. Genome Res *14*, 1085-1094.

Lee, K.S., Kim, H.J., Li, Q.L., Chi, X.Z., Ueta, C., Komori, T., Wozney, J.M., Kim, E.G., Choi, J.Y., Ryoo, H.M.*, et al.* (2000a). Runx2 is a common target of transforming growth factor beta1 and bone morphogenetic protein 2, and cooperation between Runx2 and Smad5 induces osteoblast-specific gene expression in the pluripotent mesenchymal precursor cell line C2C12. Mol Cell Biol *20*, 8783-8792.

Lee, M.L., Kuo, F.C., Whitmore, G.A., and Sklar, J. (2000b). Importance of replication in microarray gene expression studies: statistical methods and evidence from repetitive cDNA hybridizations. Proc Natl Acad Sci U S A *97*, 9834-9839.

Lee, R.C., Feinbaum, R.L., and Ambros, V. (1993b). The C. elegans heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14. Cell *75*, 843-854.

Lee, T.I., Jenner, R.G., Boyer, L.A., Guenther, M.G., Levine, S.S., Kumar, R.M., Chevalier, B., Johnstone, S.E., Cole, M.F., Isono, K.*, et al.* (2006). Control of developmental regulators by Polycomb in human embryonic stem cells. Cell *125*, 301-313.

Leung, Y.F., and Cavalieri, D. (2003). Fundamentals of cDNA microarray data analysis. Trends in Genetics *19*, 649-659.

Li, B., Carey, M., and Workman, J.L. (2007). The role of chromatin during transcription. Cell *128*, 707-719.

Li, H., and Capetanaki, Y. (1994). An E box in the desmin promoter cooperates with the E box and MEF-2 sites of a distal enhancer to direct muscle-specific transcription. EMBO J *13*, 3580-3589.

Li, S.C., Pan, C.Y., and Lin, W.C. (2006). Bioinformatic discovery of microRNA precursors from human ESTs and introns. BMC Genomics *7*, 164.

Li, X., Oghi, K.A., Zhang, J., Krones, A., Bush, K.T., Glass, C.K., Nigam, S.K., Aggarwal, A.K., Maas, R., Rose, D.W.*, et al.* (2003). Eya protein phosphatase activity regulates Six1-Dach-Eya transcriptional effects in mammalian organogenesis. Nature *426*, 247-254.

Li, X., Perissi, V., Liu, F., Rose, D.W., and Rosenfeld, M.G. (2002). Tissue-specific regulation of retinal and pituitary precursor cell proliferation. Science *297*, 1180-1183.

Lian, J.B., Stein, G.S., Stein, J.L., and van Wijnen, A.J. (1998). Transcriptional control of osteoblast differentiation. Biochem Soc Trans *26*, 14-21.

Lin Shiau, S.Y., Huang, M.C., and Lee, C.Y. (1976). Mechanism of action of cobra cardiotoxin in the skeletal muscle. J Pharmacol Exp Ther *196*, 758-770.

Littler, D.R., Alvarez-Fernandez, M., Stein, A., Hibbert, R.G., Heidebrecht, T., Aloy, P., Medema, R.H., and Perrakis, A. (2010). Structure of the FoxM1 DNA-recognition domain bound to a promoter sequence. Nucleic Acids Res *38*, 4527-4538.

Liu, Y., Chu, A., Chakroun, I., Islam, U., and Blais, A. (2010). Cooperation between myogenic regulatory factors and SIX family transcription factors is important for myoblast differentiation. Nucleic Acids Res *38*, 6857-6871.

Lopez-Rios, J., Tessmar, K., Loosli, F., Wittbrodt, J., and Bovolenta, P. (2003). Six3 and Six6 activity is modulated by members of the groucho family. Development *130*, 185-195.

Lorch, Y., LaPointe, J.W., and Kornberg, R.D. (1987). Nucleosomes inhibit the initiation of transcription but allow chain elongation with the displacement of histones. Cell *49*, 203-210.

Louie, M.C., Yang, H.Q., Ma, A.H., Xu, W., Zou, J.X., Kung, H.J., and Chen, H.W. (2003). Androgen-induced recruitment of RNA polymerase II to a nuclear receptor-p160 coactivator complex. Proc Natl Acad Sci U S A *100*, 2226-2230.

Lu, C., Tej, S.S., Luo, S., Haudenschild, C.D., Meyers, B.C., and Green, P.J. (2005). Elucidation of the small RNA component of the transcriptome. Science *309*, 1567-1569.

Luan, Y., and Li, H. (2003). Clustering of time-course gene expression data using a mixed-effects model with B-splines. Bioinformatics *19*, 474-482.

Luger, K., Mader, A.W., Richmond, R.K., Sargent, D.F., and Richmond, T.J. (1997). Crystal structure of the nucleosome core particle at 2.8 A resolution. Nature *389*, 251-260.

Mafi, R., Hindocha, S., Mafi, P., Griffin, M., and Khan, W.S. (2011). Sources of adult mesenchymal stem cells applicable for musculoskeletal applications - a systematic review of the literature. Open Orthop J *5 Suppl 2*, 242-248.

Margulies, M., Egholm, M., Altman, W.E., Attiya, S., Bader, J.S., Bemben, L.A., Berka, J., Braverman, M.S., Chen, Y.J., Chen, Z. *et al.* (2005). Genome sequencing in microfabricated high-density picolitre reactors. Nature *437*, 376-380.

Marioni, J.C., Mason, C.E., Mane, S.M., Stephens, M., and Gilad, Y. (2008). RNA-seq: an assessment of technical reproducibility and comparison with gene expression arrays. Genome Res *18*, 1509-1517.

Martinez, J., Patkaniowska, A., Urlaub, H., Luhrmann, R., and Tuschl, T. (2002). Single-stranded antisense siRNAs guide target RNA cleavage in RNAi. Cell *110*, 563-574.

Mauro, A. (1961). Satellite cell of skeletal muscle fibers. J Biophys Biochem Cytol *9*, 493-495.

McCracken, S., Fong, N., Rosonina, E., Yankulov, K., Brothers, G., Siderovski, D., Hessel, A., Foster, S., Shuman, S., and Bentley, D.L. (1997a). 5'-Capping enzymes are targeted to pre-mRNA by binding to the phosphorylated carboxy-terminal domain of RNA polymerase II. Genes Dev *11*, 3306-3318.

McCracken, S., Fong, N., Yankulov, K., Ballantyne, S., Pan, G., Greenblatt, J., Patterson, S.D., Wickens, M., and Bentley, D.L. (1997b). The C-terminal domain of RNA polymerase II couples mRNA processing to transcription. Nature *385*, 357-361.

McKinney-Freeman, S.L., Jackson, K.A., Camargo, F.D., Ferrari, G., Mavilio, F., and Goodell, M.A. (2002). Muscle-derived hematopoietic stem cells are hematopoietic in origin. Proc Natl Acad Sci U S A *99*, 1341-1346.

Meister, G., and Tuschl, T. (2004). Mechanisms of gene silencing by double-stranded RNA. Nature *431*, 343-349.

Michael, B. (2006). Sample Labeling: An Overview. In Methods in Enzymology, K. Alan, and O. Brian, eds. (Academic Press), pp. 222-237.

Minamoto, V.B., Grazziano, C.R., and Salvini, T.F. (1999). Effect of single and periodic contusion on the rat soleus muscle at different stages of regeneration. Anat Rec *254*, 281-287.

Miner, J.H., and Wold, B. (1990). Herculin, a fourth member of the MyoD family of myogenic regulatory genes. Proc Natl Acad Sci U S A *87*, 1089-1093.

Mintz, B., and Baker, W.W. (1967). Normal mammalian muscle differentiation and gene control of isocitrate dehydrogenase synthesis. Proc Natl Acad Sci U S A *58*, 592-598.

Molkentin, J.D., Black, B.L., Martin, J.F., and Olson, E.N. (1995). Cooperative activation of muscle gene expression by MEF2 and myogenic bHLH proteins. Cell *83*, 1125-1136.

Molkentin, J.D., Black, B.L., Martin, J.F., and Olson, E.N. (1996). Mutational analysis of the DNA binding, dimerization, and transcriptional activation domains of MEF2C. Mol Cell Biol *16*, 2627-2636.

Mootha, V.K., Lindgren, C.M., Eriksson, K.F., Subramanian, A., Sihag, S., Lehar, J., Puigserver, P., Carlsson, E., Ridderstrale, M., Laurila, E.*, et al.* (2003). PGC-1alpha-responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes. Nat Genet *34*, 267-273.

Moresi, V., Williams, A.H., Meadows, E., Flynn, J.M., Potthoff, M.J., McAnally, J., Shelton, J.M., Backs, J., Klein, W.H., Richardson, J.A.*, et al.* (2010). Myogenin and class II HDACs control neurogenic muscle atrophy by inducing E3 ubiquitin ligases. Cell *143*, 35-45.

Morin, R.D., O'Connor, M.D., Griffith, M., Kuchenbauer, F., Delaney, A., Prabhu, A.L., Zhao, Y., McDonald, H., Zeng, T., Hirst, M.*, et al.* (2008). Application of massively parallel sequencing to microRNA profiling and discovery in human embryonic stem cells. Genome Res *18*, 610-621.

Murre, C., McCaw, P.S., and Baltimore, D. (1989a). A new DNA binding and dimerization motif in immunoglobulin enhancer binding, daughterless, MyoD, and myc proteins. Cell *56*, 777-783.

Murre, C., McCaw, P.S., Vaessin, H., Caudy, M., Jan, L.Y., Jan, Y.N., Cabrera, C.V., Buskin, J.N., Hauschka, S.D., Lassar, A.B.*, et al.* (1989b). Interactions between heterologous helix-loop-helix proteins generate complexes that bind specifically to a common DNA sequence. Cell *58*, 537-544.

Nabeshima, Y., Hanaoka, K., Hayasaka, M., Esumi, E., Li, S., and Nonaka, I. (1993). Myogenin gene disruption results in perinatal lethality because of severe muscle defect. Nature *364*, 532-535.

Nadon, R., and Shoemaker, J. (2002). Statistical issues with microarrays: processing and analysis. Trends Genet *18*, 265-271.

Naguibneva, I., Ameyar-Zazoua, M., Polesskaya, A., Ait-Si-Ali, S., Groisman, R., Souidi, M., Cuvellier, S., and Harel-Bellan, A. (2006). The microRNA miR-181 targets the homeobox protein Hox-A11 during mammalian myoblast differentiation. Nat Cell Biol *8*, 278-284.

Nakajima, N., Horikoshi, M., and Roeder, R.G. (1988). Factors involved in specific transcription by mammalian RNA polymerase II: purification, genetic specificity, and TATA box-promoter interactions of TFIID. Mol Cell Biol *8*, 4028-4040.

Nam, J.W., Shin, K.R., Han, J., Lee, Y., Kim, V.N., and Zhang, B.T. (2005). Human microRNA prediction through a probabilistic co-learning model of sequence and structure. Nucleic Acids Res *33*, 3570-3581.

Newton, W.A., Jr., Gehan, E.A., Webber, B.L., Marsden, H.B., van Unnik, A.J., Hamoudi, A.B., Tsokos, M.G., Shimada, H., Harms, D., Schmidt, D.*, et al.* (1995). Classification of rhabdomyosarcomas and related sarcomas. Pathologic aspects and proposal for a new classification--an Intergroup Rhabdomyosarcoma Study. Cancer *76*, 1073-1085.

Nie, X., Xu, J., El-Hashash, A., and Xu, P.X. (2011). Six1 regulates Grem1 expression in the metanephric mesenchyme to initiate branching morphogenesis. Dev Biol *352*, 141-151.

Nigg, E.A. (1995). Cyclin-dependent protein kinases: key regulators of the eukaryotic cell cycle. Bioessays *17*, 471-480.

Niro, C., Demignon, J., Vincent, S., Liu, Y., Giordani, J., Sgarioto, N., Favier, M., Guillet-Deniau, I., Blais, A., and Maire, P. (2010). Six1 and Six4 gene expression is necessary to

activate the fast-type muscle gene program in the mouse primary myotome. Dev Biol *338*, 168-182.

Noble, D. (2006). The music of life : biology beyond the genome (Oxford ; New York, Oxford University Press).

Oelgeschlager, T., Chiang, C.M., and Roeder, R.G. (1996). Topology and reorganization of a human TFIID-promoter complex. Nature *382*, 735-738.

Ohto, H., Kamada, S., Tago, K., Tominaga, S.I., Ozaki, H., Sato, S., and Kawakami, K. (1999). Cooperation of six and eya in activation of their target genes through nuclear translocation of Eya. Mol Cell Biol *19*, 6815-6824.

Ohto, H., Takizawa, T., Saito, T., Kobayashi, M., Ikeda, K., and Kawakami, K. (1998). Tissue and developmental distribution of Six family gene products. Int J Dev Biol *42*, 141-148.

Okubo, Y., Bessho, K., Fujimura, K., Iizuka, T., and Miyatake, S. (1999). Expression of bone morphogenetic protein-2 via adenoviral vector in C2C12 myoblasts induces differentiation into the osteoblast lineage. Biochem Biophys Res Commun *262*, 739-743.

Oliver, G., Mailhos, A., Wehr, R., Copeland, N.G., Jenkins, N.A., and Gruss, P. (1995a). Six3, a murine homologue of the sine oculis gene, demarcates the most anterior border of the developing neural plate and is expressed during eye development. Development *121*, 4045-4055.

Oliver, G., Wehr, R., Jenkins, N.A., Copeland, N.G., Cheyette, B.N., Hartenstein, V., Zipursky, S.L., and Gruss, P. (1995b). Homeobox genes and connective tissue patterning. Development *121*, 693-705.

Ordahl, C.P., and Le Douarin, N.M. (1992). Two myogenic lineages within the developing somite. Development *114*, 339-353.

Orlowski, J., and Lingrel, J.B. (1988). Tissue-specific and developmental regulation of rat Na,K-ATPase catalytic alpha isoform and beta subunit mRNAs. J Biol Chem *263*, 10436-10442.

Orphanides, G., Lagrange, T., and Reinberg, D. (1996). The general transcription factors of RNA polymerase II. Genes Dev *10*, 2657-2683.

Ozaki, H., Watanabe, Y., Takahashi, K., Kitamura, K., Tanaka, A., Urase, K., Momoi, T., Sudo, K., Sakagami, J., Asano, M.*, et al.* (2001). Six4, a putative myogenin gene regulator, is not essential for mouse embryonal development. Mol Cell Biol *21*, 3343-3350.

Pajcini, K.V., Corbel, S.Y., Sage, J., Pomerantz, J.H., and Blau, H.M. (2010). Transient inactivation of Rb and ARF yields regenerative cells from postmitotic mammalian muscle. Cell Stem Cell *7*, 198-213.

Pfaffl, M.W. (2001). A new mathematical model for relative quantification in real-time RT-PCR. Nucleic Acids Res *29*, e45.

Pignoni, F., Hu, B., Zavitz, K.H., Xiao, J., Garrity, P.A., and Zipursky, S.L. (1997). The eye-specification proteins So and Eya form a complex and regulate multiple steps in Drosophila eye development. Cell *91*, 881-891.

Polager, S., and Ginsberg, D. (2008). E2F - at the crossroads of life and death. Trends Cell Biol *18*, 528-535.

Puri, P.L., Avantaggiati, M.L., Balsano, C., Sang, N., Graessmann, A., Giordano, A., and Levrero, M. (1997a). p300 is required for MyoD-dependent cell cycle arrest and muscle-specific gene transcription. EMBO J *16*, 369-383.

Puri, P.L., Sartorelli, V., Yang, X.J., Hamamori, Y., Ogryzko, V.V., Howard, B.H., Kedes, L., Wang, J.Y., Graessmann, A., Nakatani, Y*., et al.* (1997b). Differential roles of p300 and PCAF acetyltransferases in muscle differentiation. Mol Cell *1*, 35-45.

Quackenbush, J. (2002). Microarray data normalization and transformation. Nat Genet *32 Suppl*, 496-501.

Qureshi, S.A., and Jackson, S.P. (1998). Sequence-specific DNA binding by the S. shibatae TFIIB homolog, TFB, and its effect on promoter strength. Mol Cell *1*, 389-400.

R Development Core Team (2011). R: A Language and Environment for Statistical Computing (Vienna, Austria).

Rampalli, S., Li, L., Mak, E., Ge, K., Brand, M., Tapscott, S.J., and Dilworth, F.J. (2007). p38 MAPK signaling regulates recruitment of Ash2L-containing methyltransferase complexes to specific genes during differentiation. Nat Struct Mol Biol *14*, 1150-1156.

Rando, T.A., and Blau, H.M. (1994). Primary mouse myoblast purification, characterization, and transplantation for cell-mediated gene therapy. J Cell Biol *125*, 1275-1287.

Rao, D.D., Vorhies, J.S., Senzer, N., and Nemunaitis, J. (2009). siRNA vs. shRNA: similarities and differences. Adv Drug Deliv Rev *61*, 746-759.

Relaix, F., Montarras, D., Zaffran, S., Gayraud-Morel, B., Rocancourt, D., Tajbakhsh, S., Mansouri, A., Cumano, A., and Buckingham, M. (2006). Pax3 and Pax7 have distinct and overlapping functions in adult muscle progenitor cells. J Cell Biol *172*, 91-102.

Relaix, F., Polimeni, M., Rocancourt, D., Ponzetto, C., Schafer, B.W., and Buckingham, M. (2003). The transcriptional activator PAX3-FKHR rescues the defects of Pax3 mutant mice but induces a myogenic gain-of-function phenotype with ligand-independent activation of Met signaling in vivo. Genes Dev *17*, 2950-2965.

Relaix, F., Rocancourt, D., Mansouri, A., and Buckingham, M. (2005). A Pax3/Pax7-dependent population of skeletal muscle progenitor cells. Nature *435*, 948-953.

Rensink, W.A., and Hazen, S.P. (2006). Statistical issues in microarray data analysis. Methods Mol Biol *323*, 359-366.

Reya, T., O'Riordan, M., Okamura, R., Devaney, E., Willert, K., Nusse, R., and Grosschedl, R. (2000). Wnt signaling regulates B lymphocyte proliferation through a LEF-1 dependent mechanism. Immunity *13*, 15-24.

Richard, A.F., Demignon, J., Sakakibara, I., Pujol, J., Favier, M., Strochlic, L., Le Grand, F., Sgarioto, N., Guernec, A., Schmitt, A.*, et al.* (2011). Genesis of muscle fiber-type diversity during mouse embryogenesis relies on Six1 and Six4 gene expression. Dev Biol.

Ridgeway, A.G., and Skerjanc, I.S. (2001). Pax3 is essential for skeletal myogenesis and the expression of Six1 and Eya2. J Biol Chem *276*, 19033-19039.

Roh, T.Y., Cuddapah, S., Cui, K., and Zhao, K. (2006). The genomic landscape of histone modifications in human T cells. Proc Natl Acad Sci U S A *103*, 15782-15787.

Rudnicki, M.A., Braun, T., Hinuma, S., and Jaenisch, R. (1992). Inactivation of MyoD in mice leads to up-regulation of the myogenic HLH gene Myf-5 and results in apparently normal muscle development. Cell *71*, 383-390.

Rudnicki, M.A., Schnegelsberg, P.N., Stead, R.H., Braun, T., Arnold, H.H., and Jaenisch, R. (1993). MyoD or Myf-5 is required for the formation of skeletal muscle. Cell *75*, 1351-1359.

Saldanha, A.J. (2004). Java Treeview--extensible visualization of microarray data. Bioinformatics *20*, 3246-3248.

Santos-Rosa, H., Schneider, R., Bernstein, B.E., Karabetsou, N., Morillon, A., Weise, C., Schreiber, S.L., Mellor, J., and Kouzarides, T. (2003). Methylation of histone H3 K4 mediates association of the Isw1p ATPase with chromatin. Mol Cell *12*, 1325-1332.

Sauer, U., Heinemann, M., and Zamboni, N. (2007). Genetics. Getting closer to the whole picture. Science *316*, 550-551.

Schafer, K.A. (1998). The cell cycle: a review. Vet Pathol *35*, 461-478.

Schena, M., Shalon, D., Davis, R.W., and Brown, P.O. (1995). Quantitative monitoring of gene expression patterns with a complementary DNA microarray. Science *270*, 467-470.

Schiaffino, S., and Partridge, T. (2008). Skeletal Muscle Repair and Regeneration (Dordrecht, , NLD, Springer).

Schmidt, C., Bladt, F., Goedecke, S., Brinkmann, V., Zschiesche, W., Sharpe, M., Gherardi, E., and Birchmeier, C. (1995). Scatter factor/hepatocyte growth factor is essential for liver development. Nature *373*, 699-702.

Schwartz, S., Elnitski, L., Li, M., Weirauch, M., Riemer, C., Smit, A., Green, E.D., Hardison, R.C., and Miller, W. (2003a). MultiPipMaker and supporting tools: Alignments and analysis of multiple genomic DNA sequences. Nucleic Acids Res *31*, 3518-3524.

Schwartz, S., Kent, W.J., Smit, A., Zhang, Z., Baertsch, R., Hardison, R.C., Haussler, D., and Miller, W. (2003b). Human-mouse alignments with BLASTZ. Genome Res *13*, 103-107.

Schwarz, D.S., Hutvagner, G., Haley, B., and Zamore, P.D. (2002). Evidence that siRNAs function as guides, not primers, in the Drosophila and human RNAi pathways. Mol Cell *10*, 537-548.

Seenundun, S., Rampalli, S., Liu, Q.C., Aziz, A., Palii, C., Hong, S., Blais, A., Brand, M., Ge, K., and Dilworth, F.J. (2010). UTX mediates demethylation of H3K27me3 at muscle-specific genes during myogenesis. EMBO J *29*, 1401-1411.

Selbach, M., Schwanhausser, B., Thierfelder, N., Fang, Z., Khanin, R., and Rajewsky, N. (2008). Widespread changes in protein synthesis induced by microRNAs. Nature *455*, 58-63.

Seok, H.Y., Tatsuguchi, M., Callis, T.E., He, A., Pu, W.T., and Wang, D.Z. (2011). miR-155 inhibits expression of the MEF2A protein to repress skeletal muscle differentiation. J Biol Chem *286*, 35339-35346.

Shang, Y., Myers, M., and Brown, M. (2002). Formation of the androgen receptor transcription complex. Mol Cell *9*, 601-610.

Sharan, R., Maron-Katz, A., and Shamir, R. (2003). CLICK and EXPANDER: a system for clustering and visualizing gene expression data. Bioinformatics *19*, 1787-1799.

Sharan, R., and Shamir, R. (2000). CLICK: a clustering algorithm with applications to gene expression analysis. Proc Int Conf Intell Syst Mol Biol *8*, 307-316.

Shinin, V., Gayraud-Morel, B., Gomes, D., and Tajbakhsh, S. (2006). Asymmetric division and cosegregation of template DNA strands in adult muscle satellite cells. Nat Cell Biol *8*, 677-687.

Shtutman, M., Zhurinsky, J., Simcha, I., Albanese, C., D'Amico, M., Pestell, R., and Ben-Ze'ev, A. (1999). The cyclin D1 gene is a target of the beta-catenin/LEF-1 pathway. Proc Natl Acad Sci U S A *96*, 5522-5527.

Sledz, C.A., Holko, M., de Veer, M.J., Silverman, R.H., and Williams, B.R. (2003). Activation of the interferon system by short-interfering RNAs. Nat Cell Biol *5*, 834-839.

Smale, S.T., and Baltimore, D. (1989). The "initiator" as a transcription control element. Cell *57*, 103-113.

Smale, S.T., and Kadonaga, J.T. (2003). The RNA polymerase II core promoter. Annu Rev Biochem *72*, 449-479.

Spitz, F., Demignon, J., Porteu, A., Kahn, A., Concordet, J.P., Daegelen, D., and Maire, P. (1998). Expression of myogenin during embryogenesis is controlled by Six/sine oculis homeoproteins through a conserved MEF3 binding site. Proc Natl Acad Sci U S A *95*, 14220-14225.

Stauber, W.T., Fritz, V.K., and Dahlmann, B. (1990). Extracellular matrix changes following blunt trauma to rat skeletal muscles. Exp Mol Pathol *52*, 69-86.

Stuart, J.M., Segal, E., Koller, D., and Kim, S.K. (2003). A gene-coexpression network for global discovery of conserved genetic modules. Science *302*, 249-255.

Subramanian, A., Tamayo, P., Mootha, V.K., Mukherjee, S., Ebert, B.L., Gillette, M.A., Paulovich, A., Pomeroy, S.L., Golub, T.R., Lander, E.S.*, et al.* (2005). Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci U S A *102*, 15545-15550.

Szutorisz, H., Dillon, N., and Tora, L. (2005). The role of enhancers as centres for general transcription factor recruitment. Trends Biochem Sci *30*, 593-599.

Tajbakhsh, S., Borello, U., Vivarelli, E., Kelly, R., Papkoff, J., Duprez, D., Buckingham, M., and Cossu, G. (1998). Differential activation of Myf5 and MyoD by different Wnts in explants of mouse paraxial mesoderm and the later activation of myogenesis in the absence of Myf5. Development *125*, 4155-4162.

Tajbakhsh, S., and Buckingham, M. (2000). The birth of muscle progenitor cells in the mouse: spatiotemporal considerations. Curr Top Dev Biol *48*, 225-268.

Tajbakhsh, S., Rocancourt, D., and Buckingham, M. (1996). Muscle progenitor cells failing to respond to positional cues adopt non-myogenic fates in myf-5 null mice. Nature *384*, 266-270.

Tajbakhsh, S., Rocancourt, D., Cossu, G., and Buckingham, M. (1997). Redefining the genetic hierarchies controlling skeletal myogenesis: Pax-3 and Myf-5 act upstream of MyoD. Cell *89*, 127-138.

Takahashi, Y., Rayman, J.B., and Dynlacht, B.D. (2000). Analysis of promoter binding by the E2F and pRB families in vivo: distinct E2F proteins mediate activation and repression. Genes Dev *14*, 804-816.

Tamayo, P., Slonim, D., Mesirov, J., Zhu, Q., Kitareewan, S., Dmitrovsky, E., Lander, E.S., and Golub, T.R. (1999). Interpreting patterns of gene expression with self-organizing maps: Methods and application to hematopoietic differentiation. Proceedings of the National Academy of Sciences of the United States of America *96*, 2907-2912.

Tapscott, S.J. (2005). The circuitry of a master switch: Myod and the regulation of skeletal muscle gene transcription. Development *132*, 2685-2695.

Tavazoie, S., Hughes, J.D., Campbell, M.J., Cho, R.J., and Church, G.M. (1999). Systematic determination of genetic network architecture. Nature Genet *22*, 281-285.

Tetsu, O., and McCormick, F. (1999). Beta-catenin regulates expression of cyclin D1 in colon carcinoma cells. Nature *398*, 422-426.

Thalamuthu, A., Mukhopadhyay, I., Zheng, X., and Tseng, G.C. (2006). Evaluation and comparison of gene clustering methods in microarray analysis. Bioinformatics *22*, 2405-2412.

Tidball, J.G., and Villalta, S.A. (2010). Regulatory interactions between muscle and the immune system during muscle regeneration. Am J Physiol Regul Integr Comp Physiol *298*, R1173-1187.

Tomczak, K.K., Marinescu, V.D., Ramoni, M.F., Sanoudou, D., Montanaro, F., Han, M., Kunkel, L.M., Kohane, I.S., and Beggs, A.H. (2004). Expression profiling and identification of novel genes involved in myogenic differentiation. FASEB J *18*, 403-405.

Trimarchi, J.M., and Lees, J.A. (2002). Sibling rivalry in the E2F family. Nat Rev Mol Cell Biol *3*, 11-20.

Tuschl, T., Zamore, P.D., Lehmann, R., Bartel, D.P., and Sharp, P.A. (1999). Targeted mRNA degradation by double-stranded RNA in vitro. Genes Dev *13*, 3191-3197.

Ulitsky, I., Maron-Katz, A., Shavit, S., Sagir, D., Linhart, C., Elkon, R., Tanay, A., Sharan, R., Shiloh, Y., and Shamir, R. (2010). Expander: from expression microarrays to networks and functions. Nat Protoc *5*, 303-322.

van Bakel, H., and Holstege, F.C. (2004). In control: systematic assessment of microarray performance. EMBO Rep *5*, 964-969.

van Oevelen, C., Bowman, C., Pellegrino, J., Asp, P., Cheng, J., Parisi, F., Micsinai, M., Kluger, Y., Chu, A., Blais, A.*, et al.* (2010). The mammalian Sin3 proteins are required for muscle development and sarcomere specification. Mol Cell Biol *30*, 5686-5697.

Vermeulen, K., Van Bockstaele, D.R., and Berneman, Z.N. (2003). The cell cycle: a review of regulation, deregulation and therapeutic targets in cancer. Cell Prolif *36*, 131-149.

Vidal, M. (2001). A biological atlas of functional maps. Cell *104*, 333-339.

Visel, A., Blow, M.J., Li, Z., Zhang, T., Akiyama, J.A., Holt, A., Plajzer-Frick, I., Shoukry, M., Wright, C., Chen, F.*, et al.* (2009). ChIP-seq accurately predicts tissue-specific activity of enhancers. Nature *457*, 854-858.

Voit, T., Krogmann, O., Lenard, H.G., Neuen-Jacob, E., Wechsler, W., Goebel, H.H., Rahlf, G., Lindinger, A., and Nienaber, C. (1988). Emery-Dreifuss muscular dystrophy: disease spectrum and differential diagnosis. Neuropediatrics *19*, 62-71.

Wakitani, S., Saito, T., and Caplan, A.I. (1995). Myogenic cells derived from rat bone marrow mesenchymal stem cells exposed to 5-azacytidine. Muscle Nerve *18*, 1417-1426.

Wang, I.C., Chen, Y.J., Hughes, D., Petrovic, V., Major, M.L., Park, H.J., Tan, Y., Ackerson, T., and Costa, R.H. (2005). Forkhead box M1 regulates the transcriptional network of genes essential for mitotic progression and genes encoding the SCF (Skp2-Cks1) ubiquitin ligase. Mol Cell Biol *25*, 10875-10894.

Wang, X., Kiyokawa, H., Dennewitz, M.B., and Costa, R.H. (2002). The Forkhead Box m1b transcription factor is essential for hepatocyte DNA replication and mitosis during mouse liver regeneration. Proc Natl Acad Sci U S A *99*, 16881-16886.

Wang, Y., and Jaenisch, R. (1997). Myogenin can substitute for Myf5 in promoting myogenesis but less efficiently. Development *124*, 2507-2513.

Watanabe, Y., Duprez, D., Monsoro-Burq, A.H., Vincent, C., and Le Douarin, N.M. (1998). Two domains in vertebral development: antagonistic regulation by SHH and BMP4 proteins. Development *125*, 2631-2639.

Weintraub, H., Tapscott, S.J., Davis, R.L., Thayer, M.J., Adam, M.A., Lassar, A.B., and Miller, A.D. (1989). Activation of muscle-specific genes in pigment, nerve, fat, liver, and fibroblast cell lines by forced expression of MyoD. Proc Natl Acad Sci U S A *86*, 5434-5438.

Williams, B.A., and Ordahl, C.P. (1994). Pax-3 expression in segmental mesoderm marks early stages in myogenic cell specification. Development *120*, 785-796.

Wolber, P.K., Collins, P.J., Lucas, A.B., De Witte, A., and Shannon, K.W. (2006). The Agilent in situ-synthesized microarray platform. Methods Enzymol *410*, 28-57.

Wu, Z. (2009). A review of statistical methods for preprocessing oligonucleotide microarrays. Stat Methods Med Res *18*, 533-541.

Xie, Z., Tan, G., Ding, M., Dong, D., Chen, T., Meng, X., Huang, X., and Tan, Y. (2010). Foxm1 transcription factor is required for maintenance of pluripotency of P19 embryonal carcinoma cells. Nucleic Acids Res *38*, 8027-8038.

Yaffe, D., and Saxel, O. (1977). Serial passaging and differentiation of myogenic cells isolated from dystrophic mouse muscle. Nature *270*, 725-727.

Yajima, H., Motohashi, N., Ono, Y., Sato, S., Ikeda, K., Masuda, S., Yada, E., Kanesaki, H., Miyagoe-Suzuki, Y., Takeda, S.*, et al.* (2010). Six family genes control the proliferation and differentiation of muscle satellite cells. Exp Cell Res *316*, 2932-2944.

Yang, S., Zhao, L., Yang, J., Chai, D., Zhang, M., Zhang, J., Ji, X., and Zhu, T. (2007). deltaEF1 represses BMP-2-induced differentiation of C2C12 myoblasts into the osteoblast lineage. J Biomed Sci *14*, 663-679.

Yauk, C.L., and Berndt, M.L. (2007). Review of the literature examining the correlation among DNA microarray technologies. Environ Mol Mutagen *48*, 380-394.

Yauk, C.L., Berndt, M.L., Williams, A., and Douglas, G.R. (2004). Comprehensive comparison of six microarray technologies. Nucleic Acids Res *32*, e124.

Ye, H., Kelly, T.F., Samadani, U., Lim, L., Rubio, S., Overdier, D.G., Roebuck, K.A., and Costa, R.H. (1997). Hepatocyte nuclear factor 3/fork head homolog 11 is expressed in proliferating epithelial and mesenchymal cells of embryonic and adult tissues. Mol Cell Biol *17*, 1626-1641.

Yoon, J.K., Olson, E.N., Arnold, H.H., and Wold, B.J. (1997). Different MRF4 knockout alleles differentially disrupt Myf-5 expression: cis-regulatory interactions at the MRF4/Myf-5 locus. Dev Biol *188*, 349-362.

Zammit, P.S., Golding, J.P., Nagata, Y., Hudon, V., Partridge, T.A., and Beauchamp, J.R. (2004). Muscle satellite cells adopt divergent fates: a mechanism for self-renewal? J Cell Biol *166*, 347-357.

Zammit, P.S., Heslop, L., Hudon, V., Rosenblatt, J.D., Tajbakhsh, S., Buckingham, M.E., Beauchamp, J.R., and Partridge, T.A. (2002). Kinetics of myoblast proliferation show that resident satellite cells are competent to fully regenerate skeletal muscle fibers. Exp Cell Res *281*, 39-49.

Zammit, P.S., Relaix, F., Nagata, Y., Ruiz, A.P., Collins, C.A., Partridge, T.A., and Beauchamp, J.R. (2006). Pax7 and myogenic progression in skeletal muscle satellite cells. J Cell Sci *119*, 1824-1832.

Zamore, P.D., Tuschl, T., Sharp, P.A., and Bartel, D.P. (2000). RNAi: double-stranded RNA directs the ATP-dependent cleavage of mRNA at 21 to 23 nucleotide intervals. Cell *101*, 25-33.

Zegerman, P., Canas, B., Pappin, D., and Kouzarides, T. (2002). Histone H3 lysine 4 methylation disrupts binding of nucleosome remodeling and deacetylase (NuRD) repressor complex. J Biol Chem *277*, 11621-11624.

Zhang, N., Wei, P., Gong, A., Chiu, W.T., Lee, H.T., Colman, H., Huang, H., Xue, J., Liu, M., Wang, Y.*, et al.* (2011). FoxM1 Promotes beta-Catenin Nuclear Localization and Controls Wnt Target-Gene Expression and Glioma Tumorigenesis. Cancer Cell *20*, 427-442.

Zhang, P., Wong, C., Liu, D., Finegold, M., Harper, J.W., and Elledge, S.J. (1999). p21(CIP1) and p57(KIP2) control muscle differentiation at the myogenin step. Genes Dev *13*, 213-224.

Zhao, Y., Samal, E., and Srivastava, D. (2005). Serum response factor regulates a muscle-specific microRNA that targets Hand2 during cardiogenesis. Nature *436*, 214-220.

Zhou, Q., and Wong, W.H. (2004). CisModule: de novo discovery of cis-regulatory modules by hierarchical mixture modeling. Proc Natl Acad Sci U S A *101*, 12114-12119.

Zhu, C.C., Dyer, M.A., Uchikawa, M., Kondoh, H., Lagutin, O.V., and Oliver, G. (2002). Six3-mediated auto repression and eye development requires its interaction with members of the Groucho-related family of co-repressors. Development *129*, 2835-2849.

Zhu, J., and Emerson, S.G. (2002). Hematopoietic cytokines, transcription factors and lineage commitment. Oncogene *21*, 3295-3313.

Zhu, Z., and Miller, J.B. (1997). MRF4 can substitute for myogenin during early stages of myogenesis. Dev Dyn *209*, 233-241.

7. Contributions of Collaborators

Alphonse Chu and Alexandre Blais designed all of the experiments. Iman Chakroun and Yubing Liu performed primary myoblast isolation from mice. Yubing Liu performed the ChIP-on-Chip of Six1 and western blot of C2C12 differentiation probing for expression of Six1, Six4 and Myog. Alphonse Chu performed siRNA transfections, gene transcript profiling, qRT-PCR, western blots, immunofluorescence, BrdU pulse experiments. Alphonse Chu and Alexandre Blais performed the microarray data analysis including, normalization, significance analysis, clustering analysis and the generation of the heat-maps.

8. Appendices

Table S 1 – Oligonucleotide sequences used for siRNA and qRT-PCR.

Gene Name	Strand	Sequence	Usage
Control siRNA	Sense	UUCUCCGAACGUGUCACGUUU	siRNA
Myog siRNA	Sense	GGCUCAAGAAAGUGAAUGAUU	siRNA
Six1 siRNA	Sense	GAAAGGGAGAACACCGAAAUU	siRNA
Six4 siRNA	Sense	GCAAGCAAAUGUACUGUGUUU	siRNA
Six1	Forward	TTTTACGCAAGAGCAAGTGG	qRT-PCR
	Reverse	CTCTCGTTCTTGTGCAGGTG	qRT-PCR
Six4	Forward	CGAGACCCAGTCCAAAAGC	qRT-PCR
	Reverse	GCTAGAGAGGCTGAGGTTGG	qRT-PCR
Rps26	Forward	GCCATCCATAGCAAGGTTGT	qRT-PCR
	Reverse	GCCTCTTTACATGGGCTTTG	qRT-PCR
Tbp	Forward	TCATGGACCAGAACAACAGC	qRT-PCR
	Reverse	GCTGTGGAGTAAGTCCTGTGC	qRT-PCR
Actnb	Forward	CAGCTTCTTTGCAGCTCCTT	qRT-PCR
	Reverse	CACGATGGAGGGGAATACAG	qRT-PCR
Myog	Forward	CAGTGAATGCAACTCCCACA	qRT-PCR
	Reverse	ACCCAGCCTGACAGACAATC	qRT-PCR
Myod	Forward	CCCCGGCGGCAGAATGGCTACG	qRT-PCR
	Reverse	GGTCTGGGTTCCCTGTTCTGTGT	qRT-PCR
Actn3	Forward	GCCTCCTTCAACCACTTTGA	qRT-PCR
	Reverse	CACCATGGTCATGATTCGAG	qRT-PCR
Atp2a1	Forward	TGTCCCTCCACTTCCTCATC	qRT-PCR
	Reverse	TCCGAGCAATGAACTTGAGA	qRT-PCR
Mck	Forward	ACCTGTTCGACTTCGGATGA	qRT-PCR
	Reverse	GATTCTCACTCGCCTTCGTC	qRT-PCR
Chrna1	Forward	ATCACCGTCATCGTCATCAA	qRT-PCR
	Reverse	GGCTTCCCAGAGATGTCAGA	qRT-PCR
Cacng1	Forward	ATCTGCGCATTTCTGTCCTT	qRT-PCR
	Reverse	TCCTCGCTGTCAATCATACG	qRT-PCR
Ldb3	Forward	GATGGCGTCAACACAGACAC	qRT-PCR
	Reverse	GTCGTGGAGATGGGAATAGG	qRT-PCR
Chrng1	Forward	GGACCTGATGCGAAACTACG	qRT-PCR
	Reverse	CCTCCTCTCGTTCATTCAGG	qRT-PCR
Cenpf	Forward	CCAAGAGCTTGAAGGACAGC	qRT-PCR
	Reverse	TCTGCACCCTCAGTCTTTCC	qRT-PCR
Cenpe	Forward	AGACCAGCTCAAGGACAACC	qRT-PCR
	Reverse	GCACACAGTAACCCCTTTCC	qRT-PCR
Bub1b	Forward	GATTATTTACAACCTCTTGACAATCG	qRT-PCR
	Reverse	GAAGTCCATGATCCTCACAGC	qRT-PCR
Ccnd1	Forward	CGGATGAGAACAAGCAGACC	qRT-PCR
	Reverse	GCAGGAGAGGAAGTTGTTGG	qRT-PCR
Ccnb2	Forward	GCTGGTCCAAGTCCATTCC	qRT-PCR
	Reverse	GACCAGCTGCAGTTTCTTCC	qRT-PCR
Smc2	Forward	AAACCAGGCCCTAAATATTGC	qRT-PCR

	Reverse	GCAACCTTGAACTCCAGACC	qRT-PCR
Foxm1	Forward	ACCAATATCCAGTGGCTTGG	qRT-PCR
	Reverse	AGGGCTCCTCAACCTTAACC	qRT-PCR
E2f1	Forward	ATCCCAGTCAATCCCTGTTG	qRT-PCR
	Reverse	TGGTGACAGTTGGTCCTCTTC	qRT-PCR
Ccna2	Forward	TTCACTCATTGCTGGAGCTG	qRT-PCR
	Reverse	TCCAGTCTGTTGTGCCAATG	qRT-PCR
Ccne1	Forward	GACCCACACCAACAGCTTG	qRT-PCR
	Reverse	ACTCGGAGGAGGAGAAATCC	qRT-PCR
Cdc6	Forward	TCAGGAGCCAGACAGTCCTC	qRT-PCR
	Reverse	TTGGGATATGTGAGCAAGACC	qRT-PCR

Table S 2 – Expression of interferon gene in the different siRNA knock-down.
The normalized expression values of genes involved in the interferon pathway are indicated for each of the siRNA treatments: Control, siSix1, siSix4, siSix1&4 and siMyog. Agilent Probe ID, Gene ID, Gene Name are provided with the corresponding normalized expression values. ND – Not Detected.

Gene Symbol	GeneId	Gene Name	Normalized Expression Values				
			Control	siSix1	siSix4	siSix1&4	siMyog
Ifit1	15957	interferon-induced protein with tetratricopeptide repeats 1	ND	ND	ND	ND	ND
Mx1	17857	myxovirus (influenza virus) resistance 1	ND	ND	ND	ND	ND
Irf7	54123	interferon regulatory factor 7	0.1840	-0.0895	-0.2004	0.0000	0.0897
Oas1a	246730	2'-5' oligoadenylate synthetase 1A	ND	ND	ND	ND	ND
Oas1b	23961	2'-5' oligoadenylate synthetase 1B	ND	ND	ND	ND	ND
Oas1c	114643	2'-5' oligoadenylate synthetase 1C	ND	ND	ND	ND	ND
Oas1d	100535	2'-5' oligoadenylate synthetase 1D	ND	ND	ND	ND	ND
Oas1e	231699	2'-5' oligoadenylate synthetase 1E	ND	ND	ND	ND	ND
Oas1f	243262	2'-5' oligoadenylate synthetase 1F	ND	ND	ND	ND	ND
Oas1g	23960	2'-5' oligoadenylate synthetase 1G	ND	ND	ND	ND	ND
Ifi35	70110	interferon-induced protein 35	0.0830	-0.1314	0.0000	-0.0080	0.3715
Isg20	57444	interferon-stimulated protein	0.0591	-0.2114	-0.0534	0.0000	0.4253
Ifi27	76933	interferon, alpha-inducible protein 27	0.0000	-0.6637	0.4134	-0.4325	1.4097
Gbp1	14468	guanylate nucleotide binding protein 1	-0.5826	0.1664	-0.8362	0.0000	0.3399
Ifi44	99899	interferon-induced protein 44	ND	ND	ND	ND	ND
Irf9	16391	interferon regulatory factor 9	0.0148	-0.1539	-0.1816	0.0000	0.5318
Ifitm1	68713	interferon induced transmembrane protein 1	-0.3267	1.1121	-0.1215	0.6053	0.0000
Ifi204	15951	interferon activated gene 204	0.0000	-0.1374	-0.0943	0.0756	0.1992

Table S 3 – Cluster #1 Gene list.

Genes categorized into Cluster #1 are listed in the following table. Agilent Probe ID, Gene ID, Genebank accession number and the fold changes of each individual knock-down (siSix1, siSix4, siSix1&4) against the control siRNA are provided. Fold change values are in log 2 base.

ProbeID	GeneID	Genebank	Gene Symbol	*siSix1 vs. Ctrl*	*siSix4 vs. Ctrl*	*siSix1&4 vs. Ctrl*
A_52_P655842	11733	NM_031158	Ank1	-1.524988	-0.639001	-1.592440
A_51_P102122	17928	NM_031189	Myog	-2.276436	-1.938158	-2.959733
A_52_P131836	53414	NM_016859	Bysl	-0.350141	-0.275408	-0.399876
A_52_P166393	11870	NM_009710	Art1	-2.231942	-1.631071	-2.577022
A_52_P589568	329934	NM_194060	Foxo6	-1.208956	-0.703509	-1.091756
A_51_P263965	15368	NM_010442	Hmox1	-0.508051	-0.452707	-0.518262
A_52_P157673	13400	NM_032418	Dmpk	-0.558174	-0.571407	-0.643661
A_51_P380699	216739	NM_001033599	Acsl6	-1.275149	-0.408328	-1.058836
A_51_P508864	268859	NM_021477	A2bp1	-1.760596	-0.877422	-1.854841
A_51_P125745	27053	NM_012055	Asns	-1.413754	-0.283349	-1.036636
A_52_P118958	19298	NM_023041	Pex19	-0.730495	-0.406828	-0.579760
A_51_P136888	19645	NM_009029	Rb1	-1.061752	-0.571229	-0.916930
A_51_P103586	320862	AK043086	A730054J21Rik	-1.264361	-0.778780	-1.129631
A_51_P177667	14360	U70324	Fyn	-0.605607	-0.191450	-0.394081
A_51_P364788	17879	NM_030679	Myh1	-1.097558	-0.311186	-1.265332
A_51_P134812	69065	NM_026929	Chac1	-1.205984	-0.325477	-1.090001
A_52_P558988	68526	NM_001080707	Gpr155	-0.740693	-0.513080	-0.631214
A_51_P423666	50760	NM_015796	Fbxo17	-0.910526	-0.598423	-1.125516
A_51_P284823	20017	NM_009086	Rpo1-2	-0.446188	-0.475900	-0.659742
A_52_P370534	193385	NM_029679	6330500D04Rik	-0.981945	-0.622134	-0.968284
A_51_P508510	18128	Z11886	Notch1	-0.421332	-0.300942	-0.620342
A_51_P242663	20844	NM_011484	Stam	-0.400185	-0.154811	-0.278289
A_51_P346445	29818	NM_013868	Hspb7	-1.291984	-0.895909	-1.402797
A_52_P1004880	-	-	-	-2.286893	-2.035194	-2.750253
A_52_P949155	-	-	-	-0.317323	-0.180911	-0.313721
A_52_P172441	228859	NM_173397	D930001I22Rik	-0.527027	-0.510411	-0.585549
A_51_P345393	14102	NM_007987	Fas	-0.136733	0.254256	-0.141239
A_52_P367760	75600	NM_138304	Calml4	-1.007410	-0.564568	-1.167773
A_51_P248638	59006	NM_021503	Myoz2	-2.302237	-0.965072	-2.653121
A_51_P363947	12575	NM_007669	Cdkn1a	-1.074367	-0.672686	-1.276153
A_52_P602147	17884	NM_010855	Myh4	-2.360472	-1.335844	-2.648531
A_51_P476829	228859	AK052945	D930001I22Rik	-0.573218	-0.422401	-0.591053
A_52_P50325	14570	NM_008113	Arhgdig	-1.318232	-1.108714	-1.333471
A_52_P196105	78339	NM_175274	Ttyh3	-0.906419	-0.080590	-0.497716
A_51_P450929	11447	NM_021600	Chrnd	-1.128775	-0.652823	-1.182550
A_51_P205907	68794	NM_001081185	Flnc	-0.815629	-0.728374	-1.005062
A_51_P424191	216961	NM_139128	Coro6	-1.099086	-0.792601	-1.100563
A_52_P391983	76952	NM_029810	Nt5c2	-0.491935	-0.261333	-0.506806
A_51_P390628	239706	NM_146247	BC024814	-0.599514	-0.314138	-0.718941
A_51_P103541	12292	NM_001081023	Cacna1s	-2.853807	-1.826816	-3.204171
A_52_P249733	21393	NM_011540	Tcap	-1.949433	-1.631617	-2.468240
A_51_P278034	70240	NM_027356	Ufsp1	-0.630135	-0.604588	-0.820521
A_52_P590205	56551	NM_019913	Txn2	-0.671723	-0.279399	-0.662879
A_51_P390496	99138	NM_139308	Stard7	-0.723595	-0.794781	-1.072514
A_51_P190106	78255	BC033592	Ralgps2	-0.480766	-0.320422	-0.582706
A_52_P49727	70247	NM_027357	Psmd1	-0.439684	-0.258033	-0.564890
A_51_P181922	234365	XM_978251	BC028663	-0.573056	0.081738	-0.612409
A_51_P116755	67252	NM_026056	Cap2	-1.435141	-0.825899	-1.616585
A_52_P467726	18196	NM_010942	Nsg1	-0.876901	-0.326705	-0.698239
A_52_P217807	13175	NM_019978	Dclk1	-1.581741	-1.233551	-2.151950
A_51_P335969	13346	NM_010043	Des	-0.846932	-0.560664	-0.813070
A_51_P491648	215031	NM_153786	Vgll2	-2.355175	-0.469190	-2.081245
A_51_P265348	24131	NM_001039076	Ldb3	-1.600194	-0.930024	-1.714284

A_51_P132170	545428	NM_001025576	2610301F02Rik	-0.591497	-0.258204	-1.021351
A_52_P429534	15568	NM_010485	Elavl1	-0.226107	0.112054	-0.409700
A_52_P297724	56405	NM_019819	Dusp14	-0.403736	-0.199734	-0.276469
A_51_P112223	14860	NM_010357	Gsta4	-0.822125	-0.068798	-0.459307
A_51_P367263	235416	NM_199222	Lman1l	-1.569022	-1.072383	-1.614808
A_51_P458973	19309	NM_011224	Pygm	-2.290389	-1.488826	-2.573228
A_52_P398158	232670	NM_146173	Tspan33	-1.830122	-1.485202	-1.828955
A_51_P480578	18563	NM_008797	Pcx	-0.337113	-0.219798	-0.371975
A_52_P337259	56198	NM_013905	Heyl	-1.755106	0.059180	-1.382660
A_52_P169381	76071	AK021217	Jakmip1	-1.821883	-1.374815	-1.935336
A_51_P147942	16000	NM_010512	Igf1	-1.074605	-0.083836	-0.884852
A_52_P106391	270156	NM_212449	AU019823	-0.461162	0.032323	-0.596697
A_51_P325914	16476	NM_010591	Jun	-0.740732	-0.080571	-0.648815
A_51_P346634	102910	BC030323	Armcx4	-0.619683	-0.406809	-0.742552
A_51_P106527	68241	NM_026633	9530058B02Rik	-0.504572	-0.422378	-0.746867
A_51_P189104	19012	NM_008903	Ppap2a	-0.835411	-0.172524	-0.528004
A_51_P434670	19317	NM_021881	Qk	-0.778110	-0.364957	-0.856524
A_51_P249608	76757	AF223416	Trdn	-1.594711	0.164474	-1.491138
A_51_P435380	140709	NM_024474	Emid2	-1.152488	-0.132445	-0.878505
A_51_P459465	77480	AK083260	C330002I19Rik	-1.962019	-0.946270	-2.086605
A_51_P453818	102153	AK082735	C230098O21Rik	-0.474299	0.117452	-0.511507
A_51_P157406	19215	NM_008963	Ptgds	-0.903532	-0.783268	-1.154737
A_51_P242796	98956	NM_153126	Nat10	-0.045912	0.003047	-0.230204
A_52_P438036	78134	NM_175271	Lpar4	-0.630453	-0.119386	-0.440568
A_51_P317427	107272	NM_177420	Psat1	-0.922646	-0.623187	-0.962466
A_51_P397983	226999	AK077026	Slc9a2	-1.785095	-0.393867	-1.272237
A_51_P191865	16773	U12147	Lama2	-0.769497	0.071255	-0.717048
A_51_P492742	71769	XM_125817	Bbs10	-0.416016	-0.143271	-0.470396
A_51_P163238	271424	NM_173027	Ihpk3	-1.655783	-0.820041	-1.620377
A_52_P572447	52123	NM_026792	Agpat5	-0.617001	-0.114048	-0.508724
A_51_P232355	19184	NM_008950	Psmc5	-0.810048	-0.415419	-0.569707
A_51_P197645	70426	AK014352	Tekt5	-1.137488	-0.987715	-1.267337
A_51_P451758	74760	BC020147	Rab3il1	-1.462239	-0.559984	-1.289742
A_51_P217498	20528	NM_009204	Slc2a4	-0.748754	-0.011560	-0.741168
A_52_P144363	14186	AF127140	Fgfr4	-1.141418	-0.473374	-0.984721
A_52_P227267	98660	NM_178405	Atp1a2	-1.654013	-1.378262	-1.911770
A_52_P623211	545428	AK053175	2610301F02Rik	-0.860670	-0.311195	-1.441989
A_52_P194680	-	BI152977	-	-0.512297	0.011209	-0.313253
A_51_P511081	14773	NM_018869	Grk5	-0.968396	-0.295611	-0.510698
A_51_P341285	56046	NM_018888	Uqcc	-0.606660	-0.654679	-0.683900
A_51_P288010	11449	NM_009604	Chrng	-1.901674	-1.072830	-1.683033
A_52_P383572	17907	NM_016754	Mylpf	-2.776151	-1.619830	-3.198656
A_52_P73475	241303	NM_175511	A130092J06Rik	-2.327114	-1.833953	-2.531039
A_52_P645855	52331	NM_175096	Stbd1	-0.791864	-0.462972	-0.862862
A_51_P421559	192120	NM_138653	Bspry	-1.067970	-0.633598	-1.062710
A_52_P241742	-	AK008077	-	-0.667366	-0.486340	-0.751243
A_51_P506674	227522	NM_001013376	Rpp38	-0.686932	-0.451651	-0.760294
A_51_P173709	64297	NM_022420	Gprc5b	-0.663413	-0.112371	-0.513571
A_51_P322138	69253	NM_024441	Hspb2	-1.056193	-0.940121	-1.285948
A_51_P431785	17930	NM_008664	Myom2	-1.354192	-0.590405	-1.412751
A_52_P93284	16918	NM_008506	Mycl1	-1.799096	-0.368651	-1.616414
A_51_P269084	103172	NM_175329	Ndg2	-0.768449	-0.161090	-0.830415
A_51_P151628	80907	NM_030717	Lactb	-0.491807	-0.246434	-0.259323
A_52_P454815	14166	BC066859	Fgf11	-0.517834	-0.533929	-0.831469
A_51_P464703	20307	NM_021443	Ccl8	-1.679193	-0.732597	-1.334896
A_52_P261858	20192	NM_177652	Ryr3	-0.773482	-0.435561	-0.774910
A_52_P111715	207839	NM_172451	Galnt6	-2.505411	-1.336476	-2.524973
A_51_P408649	219134	NM_145463	Shisa2	-1.781506	-0.856962	-1.774175
A_51_P235088	240892	NM_001033344	Dusp27	-1.166815	-0.509985	-0.990486
A_52_P157450	57742	NM_021304	Abhd1	-0.983678	-0.651776	-0.836903
A_51_P199352	69563	XM_483917	2310015B20Rik	-0.655900	-0.117368	-0.827495
A_52_P64356	13602	NM_010097	Sparcl1	-1.658856	0.122924	-1.269901
A_51_P438514	73738	NM_207104	Uchl5ip	-0.544887	-0.034464	-0.548169

A_51_P232748	140571	NM_019587	Plxnb3	-1.405733	-0.969898	-1.486966
A_51_P483280	19122	NM_011170	Prnp	-0.394995	-0.140403	-0.300320
A_52_P431981	18845	NM_008882	Plxna2	-0.504231	-0.048444	-0.701097
A_51_P351166	17927	NM_010866	Myod1	-1.385000	-0.682412	-1.207588
A_52_P430348	282663	NM_173052	Serpinb1b	-0.550454	-0.020054	-0.255672
A_52_P429106	69961	NM_027278	2810432D09Rik	-0.376712	-0.058934	-0.308955
A_52_P177847	66873	NM_025817	1200009O22Rik	-1.011623	-0.223625	-0.630171
A_51_P262721	-	AK002420	-	-0.474156	-0.294223	-0.596785
A_51_P408100	14186	NM_008011	Fgfr4	-1.182410	-0.327615	-0.995398
A_51_P406557	329828	NM_001085515	AI464131	-1.287032	-0.762574	-1.194494
A_51_P258529	18858	NM_008885	Pmp22	-0.692135	-0.002319	-0.581614
A_51_P432199	60406	NM_021788	Sap30	-0.408691	-0.436397	-0.536905
A_52_P39756	380711	NM_001015046	Garnl4	-1.987840	-1.361559	-2.089490
A_51_P205545	171508	NM_133930	Creld1	-0.535938	-0.386304	-0.578774
A_52_P222624	69253	NM_024441	Hspb2	-0.820000	-0.939855	-1.156875
A_51_P240329	67048	NM_001081356	2610030H06Rik	-0.941701	-0.480543	-0.861990
A_51_P359262	66286	NM_025468	Sec11c	-0.939454	0.172885	-0.796202
A_52_P42231	78688	NM_030152	Nol3	-0.610298	-0.410079	-0.626238
A_51_P317512	70617	AK162420	5730508B09Rik	-1.024209	-0.968539	-0.982468
A_51_P516826	16002	NM_010514	Igf2	-3.076576	-2.005527	-3.232357
A_52_P58066	-	BB180072	-	-0.687608	-0.583680	-0.704625
A_51_P403413	93877	NM_053131	Pcdhb6	-0.116430	-0.177530	-0.371494
A_52_P646783	74165	NM_175206	Fbxl22	-1.179675	-0.740734	-1.291649
A_52_P376214	228410	NM_001037326	Cstf3	-0.328720	-0.510777	-0.588699
A_51_P244856	11937	NM_007504	Atp2a1	-2.064434	-1.917448	-2.233686
A_51_P455157	233199	NM_146189	Mybpc2	-1.757295	-0.346584	-1.817105
A_52_P142170	73158	XM_126172	Larp1	-0.148772	-0.018315	-0.149526
A_52_P655743	78651	NM_030145	Lsm6	-0.475803	-0.085284	-0.585912
A_52_P1034723	-	AK003800	-	-1.766313	-0.714406	-1.929993
A_51_P124741	76142	NM_133485	Ppp1r14c	-0.536148	-0.220489	-0.498657
A_51_P133562	12401	NM_007618	Serpina6	-1.227311	-0.835591	-0.899612
A_51_P207412	105377	NM_134071	Ankrd32	-0.730305	-0.194642	-0.516706
A_52_P920158	666794	NM_001081425	Rbm24	-1.538449	-0.672598	-1.518832
A_51_P395921	93841	NM_033607	Uchl4	-0.530552	-0.228712	-0.434764
A_51_P451759	74760	AK085375	Rab3il1	-1.530019	-0.445761	-1.276373
A_51_P210286	12299	NM_007582	Cacng1	-1.043425	-0.522627	-0.907505
A_51_P516833	16002	NM_010514	Igf2	-2.919533	-2.016147	-3.145259
A_52_P615958	-	-	-	-0.218266	-0.090775	-0.354417
A_51_P270355	13175	NM_019978	Dclk1	-1.124203	-0.389582	-0.866392
A_51_P168945	53332	NM_016985	Mtmr1	-0.599307	-0.315631	-0.590526
A_52_P214612	-	AK009498	-	-0.346691	-0.304457	-0.489106
A_51_P150480	67268	NM_026064	2900073G15Rik	-0.716959	0.092502	-0.549416
A_52_P582732	74100	NM_033264	Arpp21	-2.298992	-1.314765	-2.362915
A_52_P348648	223513	NM_175456	Abra	-2.281532	-1.478647	-2.510763
A_51_P393699	30963	NM_013935	Ptpla	-0.938160	-0.461273	-0.849751
A_51_P242687	18419	U96411	Otog	-1.738570	-1.601700	-2.021930
A_51_P454927	68626	NM_023479	Elac2	-0.824493	-0.404695	-0.685935
A_52_P136275	116940	NM_054089	Tgs1	0.150850	-0.132104	-0.400526
A_51_P334308	52588	NM_145928	Tspan14	-0.676228	-0.451662	-0.693720
A_52_P578043	50720	AK173011	Sacs	-0.530373	-0.160241	-0.573383
A_51_P412846	98363	NM_028889	Efhd1	-1.517883	-0.960645	-1.266589
A_52_P661412	11539	NM_001008533	Adora1	-1.170000	-0.885494	-1.345777
A_52_P65719	-	BC034637	-	-0.487570	-0.138246	-0.427142
A_51_P213334	232232	NM_144919	Hdac11	-1.215901	-0.725820	-1.230360
A_51_P487404	109272	NM_175418	8030451F13Rik	-1.126470	0.071255	-1.227205
A_52_P532456	22634	NM_009538	Plagl1	-2.437789	-2.434721	-2.531913
A_51_P379750	59011	NM_021508	Myoz1	-0.711257	-0.507320	-0.839798
A_52_P654604	-	AK003195	-	-0.200767	-0.138113	-0.334594
A_52_P588633	81799	NM_030888	C1qtnf3	-1.703662	-0.886062	-1.545439
A_51_P462516	66315	NM_001003972	Senp7	-0.105640	-0.136559	-0.272767
A_52_P572765	-	-	-	-1.818670	-0.906697	-1.538453
A_51_P137317	66966	NM_025873	Trit1	-0.405998	-0.163218	-0.658507
A_52_P90684	244923	NM_172925	Klhl31	-1.352954	-1.025839	-1.778032

A_51_P385786	15901	NM_010495	Id1	-0.522910	-0.001965	-0.450678
A_52_P157880	384929	XM_357943	Gm1947	-0.236916	-0.206234	-0.393866
A_52_P339011	18011	NM_021360	Neurl	-1.113114	-0.848505	-1.373681
A_51_P138152	52637	NM_134007	Cisd1	-0.289534	-0.065775	-0.150630
A_51_P409429	234734	NM_146217	Aars	-0.521417	-0.142434	-0.333511
A_52_P423814	12869	NM_007751	Cox8b	-1.647003	-1.078441	-1.953421
A_51_P129803	12301	NM_009786	Cacybp	-0.449284	-0.333263	-0.381582
A_51_P121947	20184	NM_011307	Uimc1	-0.389220	-0.072639	-0.426599
A_52_P250578	77574	AK173013	3321401G04Rik	-0.086900	0.310709	-0.050264
A_51_P275527	20496	NM_009194	Slc12a2	-1.115823	-0.887434	-1.132540
A_52_P253317	74931	XM_001474642	4930481A15Rik	-0.505371	-0.540950	-0.713620
A_52_P651987	106369	NM_023249	Ypel1	-0.335178	0.181845	-0.315982
A_51_P175699	67900	NM_026443	1700020C11Rik	-0.440063	-0.059397	-0.415900
A_51_P133684	13009	NM_013808	Csrp3	-2.989683	-2.629856	-3.280671
A_52_P255406	239336	NM_178717	Rxfp3	-1.239413	-0.105531	-0.834175
A_51_P111757	65247	NM_001039126	Asb1	-0.400898	-0.105856	-0.484147
A_51_P337412	216616	NM_146015	Efemp1	-0.579110	-0.547215	-0.479719
A_51_P511015	14371	NM_010246	Fzd9	-1.748554	-1.653341	-1.932936
A_51_P163778	69953	BC055845	2810025M15Rik	-0.618420	-0.299173	-0.471330
A_51_P164987	330189	NM_001039723	Tmem120b	-1.434676	-1.051987	-1.361726
A_51_P241769	19746	NM_011270	Rhd	-1.474831	-1.123311	-1.545574
A_51_P390387	67187	AK086718	Zmynd19	-0.483303	-0.386716	-0.466448
A_51_P374190	77938	NM_212473	A930008G19Rik	-0.813565	-0.430353	-0.869736
A_52_P480096	73635	NM_026865	1700113I22Rik	-0.590920	-0.398777	-0.554115
A_52_P233801	26903	NM_001077694	Dysf	-1.075302	-0.483223	-1.176552
A_51_P451458	71738	NM_174857	Mamdc2	-0.831611	-0.276611	-0.481606
A_51_P429366	55927	NM_019479	Hes6	-0.795287	-0.552972	-1.116785
A_51_P300817	11928	NM_144900	Atp1a1	-0.851819	-0.613636	-0.866454
A_52_P106482	12955	NM_009964	Cryab	-0.949572	-0.670807	-0.978882
A_51_P117995	18642	NM_021514	Pfkm	-1.012221	-0.210980	-0.825948
A_52_P157402	381511	NM_001098230	Ppm2c	-0.331466	-0.121753	-0.317842
A_52_P252469	101142	NM_133927	Itfg2	-0.537134	-0.240174	-0.420875
A_52_P312102	218877	AK129018	Sema3g	-1.305621	-0.564599	-1.234073
A_52_P384956	30948	NM_009668	Bin1	-1.284049	-0.932923	-1.346574
A_51_P151484	11931	NM_009721	Atp1b1	-0.997067	-0.052093	-1.042557
A_51_P134533	67860	NM_026416	S100a16	-1.608375	-0.353428	-1.367652
A_51_P288419	-	-	-	-1.020154	-0.756039	-0.897418
A_51_P493886	108682	NM_173866	Gpt2	-1.127152	-1.163947	-1.583719
A_52_P549427	17318	NM_183151	Mid1	-0.170048	-0.030607	-0.374338
A_51_P172085	14570	NM_008113	Arhgdig	-1.225205	-1.303834	-1.557911
A_52_P528772	17758	AK019079	Mtap4	-1.209864	-0.596954	-1.624435
A_52_P463235	67434	NM_027496	5730557B15Rik	-0.458961	-0.270096	-0.417742
A_51_P339824	234776	NM_177700	Atmin	-0.382107	0.030697	-0.278477
A_52_P28844	18175	NM_008733	Nrap	-2.354190	-0.851050	-2.753952
A_52_P592909	67800	NM_026384	Dgat2	-0.842430	-0.057500	-0.627036
A_52_P1076740	-	AK082805	-	-0.586458	-0.111865	-0.619428
A_52_P474775	27027	NM_020286	Tspan32	-1.635437	-0.902885	-1.508091
A_51_P478138	12295	NM_031173	Cacnb1	-1.324162	-1.052773	-1.488141
A_51_P463452	14081	NM_007981	Acsl1	-0.615328	-0.593547	-0.593816
A_51_P269728	19173	NM_011186	Psmb5	-0.361660	-0.298359	-0.532534
A_52_P184525	328186	AK046516	ENSMUSG00000071543	-0.312456	-0.079095	-0.353398
A_51_P450296	69876	NM_175152	Thap3	-0.226650	-0.283974	-0.576842
A_52_P263589	-	BB340137	-	-0.566748	-0.177690	-0.401537
A_51_P224311	64082	NM_022318	Popdc2	-1.855640	-1.189311	-2.066917
A_51_P521106	22138	NM_011652	Ttn	-0.893354	-0.505447	-1.249180
A_51_P239386	76142	NM_133485	Ppp1r14c	-1.863319	-0.857887	-1.571774
A_52_P11402	378937	NM_198119	Lrrc24	-0.498267	-0.148593	-0.426973
A_51_P367240	232670	NM_146173	Tspan33	-2.468834	-1.783102	-2.739160
A_51_P311038	235130	NM_001024139	Adamts15	-1.985220	-0.366384	-1.735119
A_51_P217990	231086	NM_145558	Hadhb	-0.240057	-0.285663	-0.338908
A_52_P249672	74604	AK014716	4833417J20Rik	-0.877642	-0.311786	-0.673040
A_52_P499206	27368	NM_013763	Tbl2	-0.654150	-0.457609	-0.483473
A_51_P332917	209558	BC005527	Enpp3	-1.220624	-0.018266	-1.186068

A_52_P517224	433766	NM_001039048	Trim63	-2.169691	-1.132203	-2.505831
A_51_P163906	70549	NM_001081242	Tln2	-0.395060	0.186536	-0.120982
A_51_P114407	-	-	-	-0.867478	-0.397220	-0.745452
A_52_P643375	13400	BC056615	Dmpk	-0.841494	-0.725172	-1.010391
A_52_P127465	106369	NM_023249	Ypel1	-0.512078	0.137354	-0.298688
A_51_P454008	16803	NM_008489	Lbp	-1.648243	-0.584724	-1.573073
A_51_P363914	12461	NM_007636	Cct2	-0.263847	-0.190426	-0.256787
A_52_P218271	320368	AK080508	A730063M14Rik	-0.405500	-0.478280	-0.751567
A_51_P304125	13340	NM_007854	Slc29a2	-0.409389	-0.396556	-0.588174
A_51_P379660	56772	NM_019914	Mllt11	-1.659285	-1.230928	-1.620798
A_51_P207962	19185	NM_008951	Psmd4	-0.388876	-0.242373	-0.494277
A_51_P259774	56857	NM_020258	Slc37a2	-1.125028	-1.130482	-1.047486
A_51_P268068	20471	X80339	Six1	-0.863179	0.456721	-0.919459
A_51_P275123	234353	NM_177698	Psd3	-1.142594	-0.377761	-0.889600
A_51_P269375	11733	NM_031158	Ank1	-1.162814	-0.340685	-1.814877
A_51_P187507	71985	NM_028037	Acad10	-1.300165	-0.580743	-1.199432
A_51_P317272	69556	NM_001024919	2310022M17Rik	-0.233984	-0.089395	-0.302805
A_51_P521304	67455	NM_026167	Klhl13	-1.299355	-0.353860	-1.148993
A_52_P162306	320808	AK122561	Wdr22	-0.536689	-0.310637	-0.572093
A_51_P219970	109901	AK007931	Ela1	-0.841597	-0.167150	-0.487694
A_51_P361448	71145	NM_028903	Scara5	-2.055761	-0.461130	-1.942473
A_51_P433834	229715	NM_146137	Amigo1	-0.239792	-0.183850	-0.367213
A_51_P345985	110880	NM_133199	Scn4a	-0.509063	-0.520632	-0.588263
A_51_P168708	11950	NM_009725	Atp5f1	-0.409594	-0.041072	-0.297563
A_52_P84037	216233	AK033206	Socs2	-0.859988	-0.410793	-0.914572
A_51_P268831	227659	NM_172659	Slc2a6	-2.466095	-2.042430	-2.910246
A_52_P169507	52705	NM_178610	Krr1	-0.399882	-0.092904	-0.424083
A_51_P438527	72017	NM_028057	Cyb5r1	-0.553177	-0.283428	-0.605024
A_51_P101283	-	-	-	-0.153860	-0.142577	-0.257258
A_51_P256246	66109	NM_025359	Tspan13	-1.845353	-0.175771	-1.962533
A_52_P652859	16773	NM_008481	Lama2	-0.818203	-0.027251	-0.736914
A_51_P193036	13175	NM_019978	Dclk1	-1.038470	-0.454840	-0.862141
A_51_P380750	12398	NM_009824	Cbfa2t3	-1.342847	-1.223257	-2.123376
A_51_P179041	21679	NM_011567	Tead4	-1.857521	-1.439266	-2.037127
A_51_P338728	68332	NM_001033140	0610010E21Rik	-0.983682	-0.905446	-1.244335
A_51_P331752	20292	NM_011330	Ccl11	-1.605516	-0.189566	-1.562132
A_52_P490863	66181	NM_025403	Nola3	-0.572856	-0.567290	-0.666667
A_51_P389028	21933	NM_020275	Tnfrsf10b	-0.447244	-0.014485	-0.347410
A_52_P303	83962	NM_146193	Btbd1	-0.593281	-0.085710	-0.343901
A_51_P223132	70549	NM_001081242	Tln2	0.080460	0.393271	-0.286747
A_52_P325265	68493	NM_026742	1110007M04Rik	-0.560675	-0.596424	-0.665653
A_51_P334199	13845	NM_010143	Ephb3	-0.388878	-0.347451	-0.749751
A_51_P388298	67096	NM_025962	Mmachc	-0.691624	-0.353681	-0.445242
A_51_P288558	71860	NM_027963	Wdr16	-1.806628	-1.118241	-1.901748
A_51_P486239	21922	NM_011606	Clec3b	-1.236276	-1.110537	-1.293650
A_51_P375862	56504	NM_019684	Srpk3	-2.926552	-1.935494	-3.383517
A_51_P444543	13166	NM_138942	Dbh	-1.285961	-1.434261	-1.437495
A_52_P144964	77048	AK011135	Ccdc41	-1.065306	-0.916977	-0.949590
A_52_P598309	68949	NM_001081005	1500012F01Rik	-0.574578	-0.828781	-0.851527
A_51_P193395	76457	NM_172428	Ccdc134	-0.958536	-0.783498	-1.316449
A_52_P559061	50850	AK129282	Spast	-0.716358	-0.404091	-0.799449
A_51_P325624	68196	NM_024219	Hsbp1	-0.361065	-0.202706	-0.404601
A_51_P258078	110109	NM_138747	Nol1	-0.566175	-0.347523	-0.534665
A_51_P119749	68760	AK084541	Synpo2l	-1.533770	-0.773222	-1.593466
A_52_P470150	13195	NM_016672	Ddc	-2.535258	-1.357202	-2.520906
A_51_P457187	72049	NM_028075	Tnfrsf13c	-2.096368	-1.862325	-2.080550
A_51_P304200	16660	NM_010659	Krt31	-1.823588	-1.699631	-1.789413
A_52_P381553	72014	NM_028055	1500005I02Rik	-2.369983	-2.064848	-2.566367
A_51_P279437	76574	NM_029662	Mfsd2	-0.926088	-1.079924	-1.053015
A_51_P303217	68527	NM_026754	1110017I16Rik	-1.282877	-1.393755	-1.458685
A_52_P309177	22770	NM_001042438	Zhx1	-0.460671	0.075834	-0.549729
A_51_P369762	69564	NM_027120	Itgb1bp3	-2.216260	-1.634741	-2.621184
A_52_P686383	14186	AF127140	Fgfr4	-1.438667	-0.828286	-1.341786

A_51_P147562	66084	NM_025343	Rmnd1	-0.284297	-0.101777	-0.433192
A_52_P635097	320949	AK028369	D830039M14Rik	-0.817943	-0.468817	-0.797839
A_51_P359315	11811	NM_009694	Apobec2	-1.371409	-0.860454	-1.229598
A_51_P374863	27267	NM_013742	Cars	-0.351239	-0.127764	-0.454754
A_51_P312997	26919	NM_012017	Zfp346	-0.522585	-0.444218	-0.731447
A_51_P508919	110326	NM_031867	Tas1r1	-1.591315	-1.232534	-1.682538
A_52_P379277	209558	NM_134005	Enpp3	-1.250782	-0.007995	-0.960676
A_52_P683315	105782	NM_134089	Scrib	-1.003441	-0.878458	-1.196323
A_52_P10041	11677	NM_009658	Akr1b3	-0.753418	-0.671239	-0.770600
A_51_P210340	67170	AK011997	2610306M01Rik	-0.565539	-0.664519	-0.831565
A_51_P227345	13479	NM_007876	Dpep1	-2.030436	-1.742246	-2.393312
A_52_P508991	14261	NM_010231	Fmo1	-1.337756	-0.347578	-1.138448
A_51_P111562	217835	NM_177620	Rin3	-0.961661	-0.523924	-0.807671
A_51_P268331	66889	NM_023270	Rnf128	-0.937569	-0.712619	-1.071466
A_51_P181691	257632	NM_145857	Nod2	-0.525231	0.124207	-0.542604
A_51_P393934	12521	NM_007656	Cd82	-0.624857	-0.420506	-0.569156
A_52_P269667	20681	NM_011447	Sox8	-0.982819	-0.371384	-1.142740
A_52_P559957	107227	NM_134147	Macrod1	-0.839379	-0.828000	-1.055950
A_52_P154517	21379	NM_134011	Tbrg4	-0.310572	-0.191294	-0.411435
A_52_P237948	107769	NM_145375	Tm6sf1	-1.837590	-1.304760	-1.935867
A_51_P188046	13426	NM_010063	Dync1i1	-0.924689	-0.128700	-0.827979
A_51_P108701	19181	NM_011188	Psmc2	-0.463844	-0.236954	-0.320924
A_51_P185593	233335	NM_201639	Dmn	-0.728497	-0.090926	-0.514342
A_51_P268069	20471	NM_009189	Six1	-0.928975	0.339409	-0.998093
A_52_P590396	20661	NM_019972	Sort1	-1.196618	-0.726514	-1.168032
A_52_P577205	71704	NM_027871	Arhgef3	-0.878968	-0.868356	-0.871696
A_52_P247788	225280	NM_172625	D030070L09Rik	-0.253777	-0.094053	-0.168833
A_52_P55661	330064	NM_177870	Slc5a6	-0.693465	-0.659248	-0.658513
A_52_P251672	234797	AK122299	6430548M08Rik	-0.259522	0.127428	-0.384883
A_52_P168047	27402	NM_175094	Pdhx	-0.494271	-0.199829	-0.450574
A_51_P209193	26446	NM_011971	Psmb3	-0.329373	-0.114908	-0.246084
A_52_P13802	12398	NM_009824	Cbfa2t3	-1.337739	-1.315671	-1.999516
A_51_P142196	14955	AK145379	H19	-2.277888	-1.181807	-2.209489
A_52_P314129	18767	NM_008862	Pkia	-0.943709	-0.825321	-1.067564
A_52_P276840	11981	NM_015731	Atp9a	-1.036574	-1.040412	-1.138740
A_51_P189361	71839	NM_027950	Osgin1	-0.491636	-0.158565	-0.474155
A_51_P481930	12555	NM_007662	Cdh15	-0.777996	-0.366510	-0.805225
A_51_P214083	67101	NM_025966	2310039H08Rik	-0.543911	-0.741681	-0.811024
A_51_P144160	239447	NM_173422	Colec10	-2.752619	-1.238890	-2.023750
A_52_P657844	56490	AK038731	Zbtb20	-0.205558	-0.206879	-0.465501
A_51_P359237	66241	NM_025439	Tmem9	-0.358860	-0.344479	-0.314735
A_51_P249286	19734	NM_011267	Rgs16	-1.440545	-0.744106	-1.495982
A_51_P297068	21916	S76831	Tmod1	-2.285109	-1.671723	-2.690534
A_51_P225793	72446	NM_175181	2600010E01Rik	-0.822632	-0.199384	-0.536768
A_51_P141071	67088	NM_025958	Cand2	-1.360507	-1.162814	-1.415005
A_51_P181297	66222	NM_025429	Serpinb1a	-0.891717	-0.348563	-0.716520
A_52_P317393	14766	NM_018882	Gpr56	-2.445600	-2.097932	-2.580802
A_51_P414548	12369	NM_007611	Casp7	-0.831300	-0.329224	-0.642538
A_52_P244064	22213	NM_019803	Ube2g2	-1.126158	-0.149399	-1.129372
A_51_P461319	67092	NM_025961	Gatm	-1.588089	-0.916170	-1.971312
A_51_P209736	71093	NM_153778	Atoh8	-0.699289	-0.012144	-0.231572
A_51_P317321	320678	NM_001039669	A930037G23Rik	-1.533853	-1.236172	-1.734646
A_51_P385974	19400	NM_009023	Rapsn	-0.497037	-0.442544	-0.548605
A_52_P290926	68460	NM_001013013	Dhrs7c	-2.840333	-1.662218	-2.968929
A_51_P104077	72607	NM_001013024	Usp13	-1.038407	-0.978742	-0.967431
A_52_P417825	209212	NM_145950	Osgin2	-0.479112	-0.287250	-0.432502
A_51_P504588	69742	NM_027194	Tm2d2	-1.068544	-0.095178	-0.897822
A_52_P168567	12606	NM_007678	Cebpa	-0.865320	-0.154802	-0.650954
A_52_P533707	11435	NM_007389	Chrna1	-0.620125	-0.280033	-0.648351
A_51_P236755	69454	NM_027085	Clic3	-0.840391	-0.831112	-0.982373
A_51_P165914	228491	NM_175466	Zfp770	-0.214093	0.077855	-0.220391
A_51_P272283	69574	NM_181588	Cmbl	-0.654122	0.185245	-0.525137
A_51_P142744	240725	NM_172294	Sulf1	-0.516363	0.136484	-0.288521

A_52_P199554	107328	NM_153597	Trpt1	-0.510621	-0.294934	-0.645865
A_51_P234359	20287	NM_011328	Sct	-1.287296	-0.949945	-1.396496
A_51_P246924	67971	NM_026481	Tppp3	-0.756184	-0.662870	-0.696791
A_52_P2706	68802	AK220521	Mypn	-1.184641	-0.963572	-1.364612
A_52_P243152	68897	NM_026866	Disp1	-0.400384	-0.100223	-0.316387
A_51_P260098	270076	U18992	Gcdh	-0.523865	-0.310789	-0.490834
A_52_P1197913	17873	NM_008655	Gadd45b	-0.553247	-0.471728	-0.807448
A_52_P166846	227331	AK034426	Gigyf2	-0.671632	-0.463790	-0.560835
A_51_P207622	14264	NM_021355	Fmod	-0.925532	-0.926722	-0.932815
A_52_P85292	77938	NM_212473	A930008G19Rik	-1.264990	-0.398038	-1.128211
A_52_P324848	83922	AK019270	Tsga14	-0.474475	-0.277716	-0.493093
A_51_P264495	56012	NM_018870	Pgam2	-2.444136	-1.188136	-2.736102
A_51_P200561	71653	NM_175172	4930506M07Rik	-0.355145	0.043611	-0.293772
A_52_P70856	666060	NM_001081172	Frmpd1	-0.710571	-0.603547	-0.972482
A_52_P409142	-	-	-	-0.606084	-0.525376	-0.546628
A_52_P449718	230971	XM_620148	Megf6	-1.578502	-0.940048	-1.517947
A_51_P333780	53599	NM_016898	Cd164	-0.833579	-0.538637	-0.698691
A_51_P381683	11302	NM_007377	Aatk	-2.793350	-2.120345	-2.822852
A_51_P355943	192156	NM_138656	Mvd	-0.529890	-0.159935	-0.374878
A_51_P315466	69794	BY705763	1600027J07Rik	-0.913885	-0.474109	-1.018396
A_51_P170178	232984	NM_146184	B3gnt8	-0.944059	-0.762751	-0.792377
A_51_P292332	-	-	-	-0.072054	0.323610	0.001188
A_51_P188993	68460	AK003267	Dhrs7c	-1.890090	-1.057523	-2.134920
A_51_P211334	117592	NM_080445	B3galt6	-0.416936	-0.451501	-0.429998
A_52_P172145	97484	NM_139229	Cog8	-0.314018	-0.402206	-0.361671
A_51_P212057	380839	NM_173051	Serpinb1c	-1.000544	-0.566593	-0.727269
A_51_P301435	93961	NM_033149	B3galt5	-2.682758	-1.968733	-2.630900
A_52_P176333	12695	NM_001005787	Inadl	-1.676638	-0.894397	-1.364925
A_52_P195107	16420	NM_021359	Itgb6	-1.857310	-0.784177	-2.064071
A_52_P191633	432552	NM_001013783	OTTMUSG00000005491	-1.348645	-1.193783	-1.327619
A_52_P189707	12322	NM_177407	Camk2a	-1.590034	-0.651131	-1.858858
A_51_P450888	65079	NM_022982	Rtn4r	-1.145243	-0.557863	-1.216889
A_51_P424079	13667	NM_010122	Eif2b4	-0.282571	-0.319416	-0.368131
A_51_P363187	14825	NM_008176	Cxcl1	-0.755207	0.207223	-0.620427
A_51_P115715	65256	NM_023049	Asb2	-1.779164	-0.955741	-1.895192
A_52_P532227	13609	NM_007901	Edg1	-1.479598	-0.270049	-1.408267
A_51_P264695	12971	NM_016669	Crym	-1.681047	-1.140301	-1.620267
A_51_P159835	93886	NM_053140	Pcdhb15	-0.627209	-0.546234	-0.676799
A_52_P295432	20311	NM_009141	Cxcl5	-1.880683	-0.443609	-1.765358
A_52_P562817	11733	NM_031158	Ank1	-1.416171	-0.604136	-1.648065
A_51_P126817	57314	NM_020580	Th1l	-0.249407	-0.010642	-0.339255
A_51_P261931	68162	AK153879	A930003A15Rik	-2.023168	-1.199029	-2.114362
A_52_P619880	72341	NM_001081381	Tmem103	-0.111371	-0.035552	-0.168019
A_52_P379337	68585	NM_194054	Rtn4	-0.699200	-0.657569	-0.744723
A_52_P358093	269529	XM_194139	Fbxo10	-0.144464	-0.171921	-0.276492
A_51_P206665	67281	NM_026069	Rpl37	-0.327115	-0.565114	-0.484504
A_51_P519420	94045	NM_033321	P2rx5	-0.370569	-0.229335	-0.452994
A_51_P436669	26875	NM_001110796	Pclo	-1.071558	-0.591562	-1.149312
A_51_P114049	68539	NM_134142	Tmem109	-0.928784	-0.541460	-0.929388
A_51_P142896	12509	NM_007652	Cd59a	-0.477173	0.211064	-0.366494
A_51_P119776	118449	AK004418	Synpo2	-1.075000	-0.140060	-1.037823
A_51_P375453	216795	NM_139298	Wnt9a	-0.442465	-0.580650	-0.610363
A_51_P119239	22437	AF051945	Xirp1	-1.964978	-1.431348	-2.432929
A_52_P484691	18120	NM_026246	Mrpl49	-0.748071	-0.541848	-0.601895
A_51_P238448	12445	NM_007632	Ccnd3	-0.961936	-0.550506	-1.019473
A_51_P176352	29811	NM_013864	Ndrg2	-1.129156	-0.699063	-1.179598
A_51_P388042	30928	NM_013915	Zfp238	-0.920911	-0.758289	-1.248300
A_51_P148597	67229	NM_026045	Prpf18	-0.625463	0.113822	-0.371513
A_51_P191262	71912	NM_028001	Jsrp1	-1.888108	-0.697804	-2.061352
A_51_P249118	18670	NM_008830	Abcb4	-0.360384	-0.063485	-0.318380
A_51_P192042	17879	NM_030679	Myh1	-2.747373	-0.715985	-3.289089
A_52_P485551	66789	AK017354	Alg14	-0.489217	0.130054	-0.276978
A_51_P349495	218121	NM_153546	Mboat1	-0.725944	-0.235042	-0.528218

A_52_P450918	74760	NM_144538	Rab3il1	-1.577154	-0.598184	-1.502699
A_52_P594568	74316	BC022589	Isca2	-0.835949	-0.221874	-0.697485
A_51_P321374	227721	NM_145521	Ppapdc3	-2.326284	-1.766450	-2.493800
A_52_P262967	19325	NM_016676	Rab10	-0.425768	-0.076863	-0.258726
A_52_P431872	69956	NM_027275	Ptcd3	-0.592570	-0.365486	-0.588449
A_52_P663742	76051	NM_172672	Ganc	-0.219221	-0.094198	-0.246565
A_52_P215539	223665	NM_176828	C030006K11Rik	-0.490588	-0.398126	-0.513518
A_52_P390448	-	-	-	-0.347912	-0.086946	-0.383826
A_51_P303725	74166	NM_144534	Tmem38a	-1.710408	-1.537737	-2.085316
A_51_P443344	75572	NM_029344	Acyp2	-0.586975	-0.062999	-0.625065
A_52_P355709	664837	XM_973329	LOC664837	-0.991004	-1.023718	-0.941487
A_52_P137415	18798	NM_013829	Plcb4	-0.500312	-0.212704	-0.955448
A_51_P502701	16800	NM_008487	Arhgef2	-0.184183	-0.348150	-0.561016
A_52_P153019	19220	NM_008966	Ptgfr	-0.512773	-0.225835	-0.546206
A_51_P425071	81799	NM_030888	C1qtnf3	-1.984363	-1.246410	-1.734242
A_52_P338110	67064	NM_024190	Chmp1b	-0.395392	-0.390670	-0.582240
A_51_P268953	100201	AK170682	Tmem64	-1.042542	-0.144085	-0.972919
A_52_P249687	67679	AK002941	0710001D07Rik	-0.693067	-0.660013	-0.673672
A_52_P389874	24131	NM_011918	Ldb3	-1.068984	-0.615916	-1.742997
A_52_P434729	74166	NM_144534	Tmem38a	-1.960497	-1.647852	-2.210054
A_52_P507310	64660	NM_026080	Mrps24	-0.258665	0.004167	-0.266169
A_51_P157677	53318	NM_016798	Pdlim3	-1.809839	-1.423257	-2.474108
A_51_P164256	-	-	-	-1.400318	-1.259524	-1.811972
A_52_P319326	64817	NM_022814	Svep1	-0.948494	0.131588	-0.690563
A_51_P221651	67426	NM_023341	Cabc1	-1.767790	-1.140184	-1.774821
A_51_P313337	381820	AK077261	2700089E24Rik	-0.717067	-0.686827	-0.865647
A_51_P300602	-	AV124335	-	-0.237266	-0.031283	-0.197839
A_51_P133252	26568	NM_011988	Slc27a3	-0.467192	0.037588	-0.433155
A_52_P23308	70617	NM_027482	5730508B09Rik	-0.750605	-0.864242	-0.779853
A_51_P398525	63828	NM_022014	Fn3k	-2.260961	-1.353642	-2.463796
A_51_P307589	74691	AB362563	Tdrd9	-1.247484	-1.067056	-1.282965
A_51_P508781	24115	NM_011913	Best1	-2.302196	-1.339392	-2.182976
A_51_P384243	27389	NM_001007268	Dusp13	-1.689809	-0.649548	-1.471420
A_52_P87900	268709	NM_183187	BC055107	-1.424293	-0.558365	-1.265733
A_52_P564724	-	CB587035	-	-0.559543	-0.687163	-0.633061
A_52_P376337	71238	NM_001077713	Acn9	-0.469505	-0.527925	-0.607996
A_52_P671700	-	XM_488540	-	-0.697600	-0.571936	-0.719293
A_51_P337944	140780	NM_080708	Bmp2k	-0.080565	-0.258826	-0.252570
A_52_P199084	13136	AK030285	Cd55	-1.595720	-1.344191	-1.583839
A_51_P427964	320502	NM_001081157	Lmod3	-2.220305	-1.327795	-2.588768
A_51_P333831	76511	AK008102	2010004M13Rik	-1.026565	-0.919747	-1.129444
A_52_P1019369	73761	AK014704	4833415N18Rik	-1.489680	-0.436277	-1.141448
A_51_P296905	231327	XM_896000	Ppat	-0.474731	-0.103152	-0.396842
A_52_P150675	67281	NM_026069	Rpl37	-0.720767	0.010312	-0.608545
A_51_P112355	16145	NM_018738	Igtp	-1.442907	-0.781856	-1.222372
A_52_P476775	17240	NM_010783	Mdfi	-1.097434	-0.846624	-1.148359
A_51_P452875	66268	NM_001082532	Pigyl	-0.153920	-0.038524	-0.233998
A_52_P297207	67923	AK039746	Tceb1	-0.760016	-0.486903	-0.693754
A_51_P416074	20190	NM_009109	Ryr1	-2.635049	-1.275398	-2.764946
A_52_P610541	76927	NM_029801	1700021C14Rik	-0.391417	-0.333908	-0.656296
A_51_P218535	17996	NM_010889	Neb	-2.015435	-1.285153	-2.583722
A_51_P393862	66701	NM_025716	Spryd4	-1.034431	-0.522911	-0.828551
A_52_P474266	71306	AK122332	Mfap3l	-0.522205	0.162812	-0.199169
A_52_P486175	240892	NM_001033344	Dusp27	-1.366768	-0.754904	-1.042026
A_52_P537050	50933	NM_016723	Uchl3	-0.721748	-0.359998	-0.464234
A_52_P350554	16500	NM_008420	Kcnb1	-1.348043	-0.916521	-1.546115
A_51_P185757	69743	AK052616	Casz1	-1.665132	-1.329467	-1.961963
A_51_P171883	114301	NM_023245	Palmd	-1.253101	-0.340975	-1.108118
A_51_P328202	76441	NM_001008231	Daam2	-0.545564	-0.331300	-0.245466
A_52_P197215	74136	NM_028777	Sec14l1	-1.018547	-0.810831	-0.955773
A_52_P296338	233733	NM_173739	Galntl4	-1.133582	-0.176687	-0.766827
A_51_P511546	104174	NM_138595	Gldc	-1.706907	-1.246321	-1.505558
A_52_P58257	121022	NM_080456	Mrps6	-0.389135	-0.288540	-0.306601

A_51_P148509	22138	NM_011652	Ttn	-2.118462	-1.309949	-2.600615
A_52_P371129	12569	NM_009871	Cdk5r1	-0.498983	-0.535115	-0.786693
A_51_P392350	232288	NM_145148	Frmd4b	-1.402807	-1.474243	-1.578799
A_52_P662961	68087	AK013898	Dcakd	-1.170365	-1.018096	-1.176102
A_51_P287617	211945	XM_126961	Plekhh1	-1.398931	-1.068476	-1.328606
A_52_P520397	30057	NM_013897	Timm8b	-0.476346	-0.753958	-0.788088
A_51_P122425	13019	NM_007795	Ctf1	-0.434274	-0.580459	-0.694557
A_51_P252199	236792	NM_146234	Tmem32	-0.616185	-0.302359	-0.460210
A_51_P144926	108679	NM_133805	Cops8	-0.777924	-0.029741	-0.357076
A_52_P453833	19027	NM_013635	Sypl	-0.478019	-0.362131	-0.510650
A_51_P342082	56530	NM_019953	Cnpy2	-0.449274	-0.082530	-0.424669
A_51_P302149	12632	NM_007688	Cfl2	-0.846811	-0.399807	-0.926602
A_51_P206824	69585	NM_027126	Hfe2	-3.308157	-1.884423	-3.301802
A_52_P563123	21954	NM_009406	Tnni3	-1.108197	0.026518	-1.128806
A_52_P358505	70598	NM_001081243	Filip1	-1.348135	-0.934945	-1.680918
A_51_P352924	14155	NM_010193	Fem1b	-0.765730	-0.551796	-0.731115
A_51_P266168	26987	NM_001039169	Eif4e2	-0.773782	-0.241417	-0.553376
A_51_P389421	68636	NM_023480	Fahd1	-0.437249	-0.611014	-0.698100
A_52_P139569	-	BE650457	-	-0.559610	0.145074	-0.415648
A_52_P609448	93871	NM_145125	Brwd1	0.134283	-0.157001	-0.550588
A_51_P288009	11449	NM_009604	Chrng	-1.786617	-1.028267	-1.839942
A_51_P431316	66799	NM_025773	Ube2w	-0.618829	-0.292779	-0.519496
A_51_P307232	16579	NM_010629	Kifap3	-0.498944	-0.179922	-0.556203
A_51_P413785	12238	NM_147778	Commd3	-0.101634	0.002052	-0.282909
A_52_P346706	11677	NM_009658	Akr1b3	-1.241983	-1.085361	-1.232675
A_51_P423859	12335	NM_007601	Capn3	-0.591631	-0.101906	-0.417500
A_51_P307624	102093	NM_199446	Phkb	-0.399184	-0.347111	-0.384330
A_51_P442990	72416	NM_028233	Lrpprc	-0.429687	-0.221822	-0.550636
A_51_P309988	70355	NM_147217	Gprc5c	-2.224701	-1.400239	-2.215357
A_51_P208603	13371	NM_010050	Dio2	-2.077338	-0.927585	-2.315596
A_51_P437559	21853	NM_011589	Timeless	-0.349966	0.116415	-0.281370
A_51_P396003	67800	NM_026384	Dgat2	-0.766326	0.060531	-0.508182
A_51_P219312	22138	AK010153	Ttn	-1.739090	-0.179532	-1.626932
A_51_P161342	74104	NM_023732	Abcb6	-1.071366	-0.641602	-1.051429
A_51_P416858	17901	NM_021285	Myl1	-2.439735	-0.744762	-2.432315
A_51_P109144	66790	NM_025768	Grtp1	-0.481944	-0.364057	-0.552566
A_51_P354913	384061	NM_027402	Fndc5	-2.181371	-2.035293	-2.503033
A_52_P439502	211945	XM_126961	Plekhh1	-1.021864	-1.205676	-1.299219
A_52_P165657	13421	NM_007870	Dnase1l3	-1.230396	-1.005980	-1.318390
A_51_P480311	14061	NM_010168	F2	-1.399844	-0.428494	-1.447938
A_52_P385606	12709	NM_021273	Ckb	-0.875308	-0.767416	-0.793126
A_51_P131653	76927	NM_029801	1700021C14Rik	-0.347280	-0.230545	-0.381922
A_51_P297069	21916	NM_021883	Tmod1	-2.581255	-1.631170	-2.680351
A_51_P303371	69216	NM_024462	Ccdc23	-0.777415	-0.918040	-1.032904
A_51_P118885	17117	NM_008537	Amacr	-0.742708	-0.307775	-0.763525
A_51_P438083	14412	NM_144512	Slc6a13	-0.988615	-0.876440	-1.147065
A_51_P393897	110119	NM_025837	Mpi	-0.848245	-0.785322	-0.754171
A_51_P393223	78321	BC023221	Ankrd23	-2.464521	-1.576108	-2.738753
A_52_P444235	-	-	-	-0.657320	-0.552180	-0.582053
A_52_P518182	99470	NM_133853	Magi3	-0.379015	-0.221930	-0.625168
A_52_P299934	24115	NM_011913	Best1	-1.752512	-1.114375	-1.353154
A_51_P497993	68526	BC026382	Gpr155	-0.577163	-0.086386	-0.385975
A_52_P322141	78317	NM_001081291	Ccdc88b	-1.285373	-1.130059	-1.442519
A_51_P164630	68680	NM_026808	1110028A07Rik	-1.325888	-1.000606	-1.330807
A_52_P568792	68263	NM_024221	Pdhb	-0.501058	-0.293393	-0.470152
A_51_P322941	231134	NM_172708	Dok7	-1.716697	-1.024262	-1.662033
A_51_P464300	14559	NM_008107	Gdf1	-1.422516	-0.956018	-1.578354
A_51_P487832	666794	NM_001081425	Rbm24	-0.892893	-0.327347	-1.374443
A_51_P243808	215243	NM_153137	Traf3ip3	-1.211517	-1.011872	-1.350562
A_51_P283516	20666	NM_009234	Sox11	-0.634161	-0.544384	-0.896753
A_51_P211786	71797	XM_978459	Chst13	-1.729233	-1.554766	-1.730335
A_51_P298570	13682	NM_013506	Eif4a2	-0.623190	-0.609663	-0.651205
A_52_P313813	52530	NM_026631	Nola2	-0.416844	-0.383513	-0.492498

A_51_P368012	17240	NM_010783	Mdfi	-0.967141	-0.926403	-1.031221
A_52_P151362	67216	NM_026037	Mboat2	-0.894746	-0.893226	-0.961140
A_51_P265351	24131	NM_001039076	Ldb3	-1.600928	-1.218657	-1.777144
A_51_P115953	629147	AK042789	Ctxn3	-2.310253	-1.920467	-2.371765
A_51_P323081	70237	NM_198161	Bhlhb9	-0.674615	-0.555531	-0.534685
A_52_P512301	55927	NM_019479	Hes6	-1.268192	-0.851906	-1.230411
A_52_P257812	16956	NM_008509	Lpl	-1.551516	0.398327	-1.173935
A_51_P172663	53376	NM_198092	Usp2	-1.823384	-1.489301	-2.009555
A_52_P38208	74100	NM_028755	Arpp21	-2.476615	-2.348665	-2.958655
A_52_P564005	71591	AK018584	Zfp251	-0.681966	0.287275	-0.312818
A_51_P298266	66643	NM_025681	Lix1	-0.661756	0.055866	-0.493494
A_52_P153685	329152	AK129325	Hecw2	-0.839877	-0.385969	-0.641365
A_52_P89099	70769	NM_001039352	Nolc1	-0.719276	-0.317258	-0.578530
A_51_P480290	69743	NM_027195	Casz1	-2.099258	-1.606427	-2.147401
A_52_P161237	109697	NM_025350	Cpa1	-1.187920	-1.273495	-1.434179
A_52_P2659	-	-	-	-0.889177	-0.780568	-0.815740
A_51_P219266	71753	NM_027902	Tmprss6	-0.926196	-0.484942	-1.006562
A_51_P177210	17897	NM_010859	Myl3	-2.630115	-0.800569	-2.624063
A_51_P445841	97998	NM_145470	Depdc6	-1.026395	-0.121589	-0.560707
A_51_P194273	24053	NM_011892	Sgcg	-2.246076	-1.081122	-2.295311
A_52_P387564	21917	NM_001080129	Tmpo	-0.574910	0.305183	-0.281260
A_52_P443330	70465	NM_027432	Wdr77	-0.576288	-0.443500	-0.570537
A_52_P220090	53599	NM_016898	Cd164	-0.697860	-0.678421	-0.706016
A_51_P315391	378431	NM_138628	Txlnb	-1.533091	-0.510812	-1.298001
A_52_P222981	241062	AK033485	Pgap1	-0.457659	-0.553809	-0.589257
A_52_P454295	22138	NM_011652	Ttn	-1.192833	-0.586624	-1.657597
A_52_P482941	22791	NM_009584	Dnajc2	0.117378	-0.021483	-0.292512
A_51_P265016	106369	NM_023249	Ypel1	-0.214518	0.333950	-0.248224
A_52_P153700	208177	NM_153412	Phldb2	-0.435138	0.041719	-0.211826
A_51_P179258	269152	NM_177757	Kif26b	-1.051106	-0.478906	-1.019684
A_51_P352720	93677	NM_053098	Lmod2	-1.468748	-1.197104	-1.885540
A_52_P262883	18938	NM_008889	Ppp1r14b	-0.396027	-0.156924	-0.289569
A_52_P672532	66273	NM_183251	1810020D17Rik	-1.298167	-0.847695	-1.342517
A_51_P384441	230917	NM_177672	Tmem201	-0.857402	-0.830422	-0.932746
A_52_P284495	67434	NM_026153	5730557B15Rik	-0.735634	-0.537994	-0.845163
A_52_P110002	66488	NM_025591	2010309E21Rik	-0.441508	-0.397890	-0.445148
A_52_P295104	66528	BC032884	2210020M01Rik	-2.769060	-2.464994	-3.141548
A_51_P413687	68033	BI905520	Cox19	-0.957639	-0.113261	-0.841612
A_52_P561650	246228	NM_147776	Vwa1	-0.873359	-0.224777	-0.731336
A_51_P401792	22138	NM_028004	Ttn	-1.397080	-0.535483	-2.099657
A_52_P656699	11474	NM_013456	Actn3	-1.672950	-1.085425	-1.909024
A_51_P309208	71802	AK005186	1500010C09Rik	-1.346288	-0.846867	-1.631922
A_51_P289836	63959	NM_022880	Slc29a1	-0.918781	-0.517866	-1.074550
A_51_P200667	94040	NM_053155	Clmn	-1.063897	-0.480795	-1.117968
A_51_P209401	-	BC062811	-	-1.541751	-1.097971	-1.601139
A_51_P420037	54563	NM_018815	Nup210	-1.185146	-0.983389	-1.446607
A_51_P339540	12577	NM_009876	Cdkn1c	-1.809116	-2.182526	-2.257298
A_52_P141488	14773	NM_018869	Grk5	-0.784831	-0.213662	-0.495053
A_52_P380887	78321	NM_153502	Ankrd23	-2.015613	-1.467662	-2.374443
A_51_P375187	67239	NM_001042556	Bxdc1	-0.460280	-0.102460	-0.433888
A_51_P500135	234593	NM_145602	Ndrg4	-1.055555	-0.634898	-0.746166
A_52_P391110	20190	NM_009109	Ryr1	-2.588949	-1.172871	-2.717910
A_52_P330289	234515	NM_001024617	Inpp4b	-1.095057	-0.694426	-1.291509
A_52_P134141	66459	NM_025574	Pigy	-0.708169	-0.730304	-0.692965
A_51_P229602	18845	NM_008882	Plxna2	-0.913389	-0.140042	-0.556035
A_52_P594410	17681	NM_010827	Msc	-1.244569	-0.433587	-1.255888
A_51_P282630	80708	NM_030880	Pacsin3	-0.491116	-0.253928	-0.537710
A_51_P503433	67916	NM_080555	Ppap2b	-0.738873	-0.034497	-0.864318
A_52_P563340	-	AK007443	-	-0.822385	-0.577958	-1.027373
A_51_P438905	12695	NM_172696	Inadl	-1.747547	-1.219214	-1.711654
A_51_P168123	107045	NM_134137	Lars	-0.292228	-0.256620	-0.449560
A_51_P433141	56191	NM_001002272	Tro	-1.128402	-1.212733	-1.494512
A_52_P586679	20471	NM_009189	Six1	-1.384835	0.197528	-1.157776

A_52_P236883	19291	AK139376	Purb	-0.749966	-0.277010	-0.704427
A_52_P351116	16000	NM_010512	Igf1	-1.462175	-0.453277	-1.415190
A_51_P267751	15464	NM_010473	Hrc	-2.479048	-1.393804	-3.040073
A_51_P502888	17977	NM_010881	Ncoa1	-0.556030	-0.360538	-0.673790
A_52_P63709	223601	NM_144846	0910001A06Rik	-0.903303	-0.593810	-1.081522
A_52_P230947	108121	NM_024187	U2af1	-0.458714	-0.110653	-0.286053
A_51_P184300	13527	NM_010087	Dtna	-1.135639	-0.776748	-1.215292
A_52_P141136	75578	NM_001113412	Fggy	-0.994133	-0.380827	-0.795645
A_51_P335491	20391	NM_009161	Sgca	-1.395491	-0.645470	-1.198698
A_52_P517762	70129	NM_023557	Slc44a4	-2.218528	-1.663627	-2.398668
A_51_P457989	52187	NM_027491	Rragd	-0.693376	0.112518	-0.689215
A_52_P498193	216188	NM_153543	Aldh1l2	-0.865698	0.154510	-0.681986
A_51_P224505	12017	NM_009736	Bag1	-0.624837	-0.228343	-0.455119
A_51_P504478	98766	NM_133835	Ubac1	-0.640296	-0.408475	-0.670216
A_52_P493867	69742	NM_027194	Tm2d2	-1.264353	-0.089371	-0.722031
A_52_P218100	54673	AK086477	Sh3glb1	-0.636063	-0.592515	-0.480617
A_51_P388412	13136	NM_010016	Cd55	-1.773198	-1.294956	-1.445927
A_51_P160870	68585	NM_194054	Rtn4	-0.649799	-0.621183	-0.731431
A_52_P559877	227835	NM_172662	Gtdc1	-0.290225	-0.263517	-0.324134
A_51_P161054	16420	NM_021359	Itgb6	-1.602431	-0.710293	-1.895035
A_51_P425490	28169	NM_053014	Agpat3	-0.625355	-0.299160	-0.571242
A_51_P207940	320709	NM_178789	Tmem117	-0.887416	-0.668588	-1.185494
A_51_P434859	17356	NM_010806	Mllt4	-0.561874	0.165371	-0.566773
A_52_P524700	22138	NM_028004	Ttn	-1.934562	-1.325515	-2.396254
A_51_P142813	66695	NM_025711	Aspn	-1.463884	0.059379	-1.218560
A_51_P316311	68507	XM_129443	Ppfia4	-2.118430	-1.351244	-2.508065
A_51_P181175	58522	NM_021447	Trim54	-2.439911	-1.332152	-2.709323
A_51_P122141	74490	AK017382	5430432N15Rik	-1.425806	-0.954944	-1.758690
A_52_P331043	72575	AK021262	C430049B03Rik	-0.709629	-0.473842	-0.895579
A_51_P210510	13602	NM_010097	Sparcl1	-1.174028	0.006121	-1.350220
A_51_P107652	70129	NM_023557	Slc44a4	-1.809590	-1.500593	-2.128713
A_52_P468537	56046	NM_018888	Uqcc	-0.351672	-0.325939	-0.471865
A_51_P520894	20536	NM_009208	Slc4a3	-0.621717	-0.460856	-0.875412
A_52_P184149	17768	NM_008638	Mthfd2	-1.080479	-0.440576	-1.106949
A_51_P159833	93886	NM_053140	Pcdhb15	-0.659449	-0.444666	-0.803009
A_51_P211305	116904	NM_054085	Alpk3	-1.937498	-1.006141	-2.343049
A_51_P412966	16979	NM_008516	Lrrn1	-0.415872	0.033058	-0.361005
A_51_P177261	237611	NM_177707	Stac3	-2.097912	-1.249408	-2.409540
A_51_P213030	107227	NM_134147	Macrod1	-0.809035	-0.727638	-0.943035
A_52_P817257	432995	DV055726	EG432995	-1.245368	-0.935812	-1.289108
A_51_P207411	105377	BC009101	Ankrd32	-0.402362	-0.166951	-0.541108
A_51_P497090	-	XM_001473388	-	-0.491905	0.058806	-0.447791
A_51_P158120	97884	NM_178640	B3galnt2	-0.640822	-0.098583	-0.645748
A_52_P405460	70355	NM_001110337	Gprc5c	-1.459782	-0.962294	-1.613270
A_51_P368543	64296	NM_022419	Abhd8	-0.591918	-0.079960	-0.481331
A_52_P268563	76757	XM_483890	Trdn	-1.016766	0.201943	-1.327505
A_51_P500984	17873	NM_008655	Gadd45b	-0.345047	-0.108785	-0.384761
A_51_P401683	107769	NM_145375	Tm6sf1	-1.950687	-1.486794	-2.173888
A_51_P104670	12181	NM_013481	Bop1	-0.468481	-0.475268	-0.450430
A_52_P273120	17901	NM_001113387	Myl1	-2.805207	-0.941290	-2.780981
A_52_P299231	18674	AK028313	Slc25a3	-2.320125	-1.157469	-2.083949
A_52_P514488	277978	NM_177788	Exoc3l	-1.437064	-1.057777	-1.872740
A_51_P372418	68036	NM_026521	Zfp706	-0.800665	-0.595346	-0.614301
A_51_P133229	72043	NM_028072	Sulf2	-0.472873	0.013788	-0.338771
A_51_P261351	18120	NM_026246	Mrpl49	-0.443617	-0.099082	-0.271793
A_52_P26991	13175	NM_001111053	Dclk1	-1.542093	-0.918747	-1.696810
A_52_P533809	15901	NM_010495	Id1	-0.424732	0.157751	-0.285856
A_51_P319392	208366	AK048810	Rpp40	-0.442051	-0.064409	-0.427658
A_52_P300533	71930	AK009785	2310043M15Rik	-1.700814	-1.408604	-1.713488
A_51_P433281	67216	NM_026037	Mboat2	-0.994937	-0.988426	-1.216259
A_51_P353232	21925	NM_009394	Tnnc2	-2.730774	-1.299901	-3.178287
A_52_P658320	243197	NM_172883	Mfsd7	-1.410559	-1.079422	-1.703905
A_51_P362627	21955	NM_011618	Tnnt1	-1.318493	-0.475367	-1.620889

A_52_P579531	53318	NM_016798	Pdlim3	-1.812541	-1.171951	-2.229054
A_51_P508341	52250	NM_178608	Reep1	-2.454023	-1.978055	-2.708039
A_51_P302377	14181	NM_008009	Fgfbp1	-1.254170	-0.531704	-1.475843
A_52_P175157	17258	AK136368	Mef2a	-0.587490	-0.573395	-0.631034

Table S 4 – Cluster #3 Gene list.

Genes categorized into Cluster #3 are listed in the following table. Agilent Probe ID, Gene ID, Genebank accession number and the fold changes of each individual knock-down (siSix1, siSix4, siSix1&4) against the control siRNA are provided. Fold change values are in log 2 base.

ProbeID	GeneID	Genebank	Gene Symbol	*siSix1 vs. Ctrl*	*siSix4 vs. Ctrl*	*siSix1&4 vs. Ctrl*
A_51_P256202	64138	NM_022325	Ctsz	-0.051241	0.160681	0.086149
A_51_P165435	12857	NM_009941	Cox4i1	0.027799	0.000746	-0.081114
A_52_P47006	17535	AK161013	Mre11a	0.198810	0.376134	-0.135202
A_52_P75348	70385	NM_027411	Ccdc99	0.165420	0.577960	0.337885
A_52_P84814	18812	NM_011118	Prl2c3	-0.004689	1.099005	1.076897
A_52_P272915	226182	NM_177342	Taf5	0.368982	0.574077	0.436181
A_51_P215190	78767	NM_030172	2610021K21Rik	-0.109717	0.268143	0.192788
A_51_P415059	20877	NM_011496	Aurkb	0.118884	0.740311	0.425755
A_52_P267824	72361	NM_197999	2210023G05Rik	0.030987	1.118361	0.187300
A_51_P509651	74107	NM_028760	Cep55	0.036642	0.796815	0.265576
A_51_P263407	73338	AK028913	1700041B20Rik	0.103478	0.370489	0.152415
A_52_P362997	70454	NM_027429	Cenpl	-0.200099	0.419944	0.142532
A_51_P431737	107869	NM_145953	Cth	-0.069573	0.177851	-0.024396
A_51_P465331	14178	NM_008008	Fgf7	0.143141	0.757828	0.275916
A_51_P174192	229841	NM_173762	Cenpe	-0.126790	0.438988	0.065899
A_51_P462249	66197	NM_025415	Cks2	0.234376	0.641046	0.445380
A_51_P109258	12879	NM_138686	Cys1	0.404422	0.820646	0.453489
A_51_P359303	66497	NM_025599	2610528E23Rik	0.083774	0.371782	0.035999
A_51_P464822	50709	NM_015787	Hist1h1e	0.526677	0.891306	0.225124
A_51_P396351	18538	NM_011045	Pcna	-0.137180	0.297855	0.157486
A_51_P120636	13555	NM_007891	E2f1	-0.020784	0.454195	0.219697
A_51_P249989	211550	NM_145133	Tifa	0.312212	0.597729	0.447704
A_52_P533065	65105	NM_144509	Arl6ip4	0.199115	0.393853	0.196282
A_51_P246119	108000	NM_001081363	Cenpf	0.330943	0.970936	0.480117
A_52_P87839	20348	NM_013657	Sema3c	-0.291847	0.478685	-0.179098
A_51_P112627	20446	NM_009180	St6galnac2	0.121858	0.632923	0.244794
A_51_P512541	654795	NM_001082975	2310014G06Rik	0.073996	0.875042	0.483713
A_51_P145220	18040	NM_008691	Nefm	-0.378521	0.971542	0.259894
A_51_P313896	68743	NM_028390	Anln	0.357996	0.709560	0.367736
A_51_P481920	12428	NM_009828	Ccna2	-0.106470	0.560492	0.033825
A_52_P228667	-	BB277834	-	0.471499	1.037684	0.817277
A_52_P488092	208618	NM_001081006	Etl4	0.087595	0.462722	0.166029
A_52_P468268	51788	NM_016750	H2afz	-0.247664	0.286484	-0.057831
A_51_P408059	68051	NM_026532	Nutf2	-0.094609	0.199834	-0.071704
A_52_P657800	99100	NM_001081091	Cep152	-0.036389	0.477753	0.218103
A_51_P431018	20342	NM_019414	Selenbp2	-0.299716	0.602849	-0.041866
A_51_P193832	170799	NM_001081346	Rtkn2	-0.081159	0.438766	0.078895
A_52_P28792	229504	NM_177663	Isg20l2	0.496292	0.721319	0.481311
A_51_P442284	30045	NM_013888	Dnajc12	0.353528	0.715978	0.639312
A_51_P215627	211623	NM_207229	Plac9	0.565000	1.108295	0.544028
A_52_P211223	108912	NM_175384	Cdca2	0.052421	0.650227	0.281803
A_52_P485417	97165	NM_008252	Hmgb2	0.311067	0.703056	0.364949
A_52_P124279	22343	NM_011699	Lin7c	-0.356041	0.513645	0.054234
A_51_P410451	71924	NM_028006	Tube1	-0.106103	0.459065	0.024253
A_52_P683038	77733	NM_029965	Rnf170	0.171025	0.432821	-0.181168
A_52_P528592	93686	NM_175387	Rbm9	0.692593	0.830122	0.550646
A_52_P414464	53901	NM_207649	Rcan2	-0.028254	0.761217	0.161354
A_52_P68033	75767	NM_001080813	Rab11fip1	0.003441	0.563070	0.228361
A_51_P345139	67144	NM_024194	Lrrc40	-0.047855	0.455922	0.119533
A_52_P681659	213582	NM_001081230	Mtap9	0.205410	0.600847	0.370462
A_51_P495012	12412	NM_007622	Cbx1	0.012275	0.204810	-0.227393
A_51_P517157	74122	NM_028766	Tmem43	0.171426	0.590503	0.411708
A_52_P624168	70472	NM_027435	Atad2	0.164982	0.707325	0.243767
A_51_P340355	272396	NM_172310	Tarsl2	-0.149472	0.315294	0.062957
A_52_P205282	59026	NM_021523	Huwe1	0.720864	0.434531	-0.009032
A_51_P107362	216233	NM_007706	Socs2	-0.156494	0.385902	-0.193632
A_51_P357195	232286	NM_001081111	Tmf1	-0.089668	0.425621	0.167549
A_51_P490509	12236	NM_009773	Bub1b	-0.266356	0.455026	-0.018234
A_52_P581100	70005	NM_027285	1700029I01Rik	0.235393	0.428374	0.385013
A_51_P381381	22152	NM_023279	Tubb3	-0.131143	0.222056	0.018731

A_51_P144264	16598	NM_008452	Klf2	-0.611074	0.170293	-0.224301
A_51_P264064	16319	AK034764	Incenp	0.173396	0.789671	0.419180
A_51_P100174	17427	NM_008613	Mns1	0.677393	1.008800	0.765630
A_51_P430678	68876	AK004405	Xrcc6bp1	0.155908	0.355418	0.257932
A_52_P139650	66468	NM_025581	2810433K01Rik	-0.079441	0.677849	0.176972
A_51_P369224	-	AK019785	-	0.258963	0.654101	0.584140
A_52_P179011	321008	NM_177312	6330408A02Rik	-0.225510	0.321333	0.090931
A_52_P351882	242291	NM_177730	Impad1	0.563342	0.520348	0.417253
A_51_P405766	21915	NM_023136	Dtymk	-0.052128	0.371665	0.078922
A_51_P328652	70025	NM_133348	Acot7	0.175848	0.361621	0.369603
A_51_P133138	19348	Y09632	Kif20a	-0.152288	0.692266	0.315494
A_51_P452533	27877	BC035277	D1Ertd471e	0.096465	0.325725	0.343016
A_51_P465327	320727	NM_001081113	Ipo8	0.213439	0.282815	0.201167
A_51_P195153	29870	NM_013882	Gtse1	0.308120	0.877375	0.649109
A_52_P20727	105193	NM_175340	Nhlrc1	-0.266787	0.097350	-0.110434
A_52_P167249	27360	NM_013758	Add3	-0.327966	0.193736	-0.315602
A_52_P134146	66274	NM_001076681	1810012P15Rik	0.004206	0.470189	0.231621
A_52_P359621	269593	NM_024452	Luzp1	0.718152	0.509100	0.336080
A_52_P163795	22154	NM_011655	Tubb5	-0.081970	0.236708	-0.034586
A_51_P130015	13605	NM_007900	Ect2	0.152760	0.764918	0.354170
A_51_P468249	18675	NM_011077	Phex	-0.294832	1.027601	-0.123635
A_52_P406819	67121	BC086483	Mastl	-0.135799	0.305825	0.006442
A_52_P9656	15366	NM_013552	Hmmr	0.250235	0.433695	0.178404
A_51_P472630	77605	XM_001476068	H2afv	0.076800	0.427833	0.288275
A_51_P209921	212448	AK034141	9330159F19Rik	-0.611511	0.385367	-0.380232
A_51_P502082	20133	NM_009103	Rrm1	-0.145340	0.211792	-0.182745
A_52_P361551	50497	NM_015765	Hspa14	0.094634	0.292717	0.212047
A_52_P763938	100043189	BQ960954	LOC100043189	0.117507	0.291708	0.149673
A_52_P588483	14114	NM_010180	Fbln1	0.457655	0.692675	0.774566
A_52_P276607	19244	NM_008974	Ptp4a2	0.029567	0.116192	-0.098499
A_51_P122321	77125	NM_133775	Il33	-0.169602	1.178404	0.020095
A_52_P244682	226421	NM_145509	5430435G22Rik	0.722744	1.001140	0.728139
A_51_P385043	72723	NM_178384	Zfp74	-0.092837	0.273687	-0.063973
A_52_P647260	118449	AK035258	Synpo2	-0.438006	0.380324	-0.297543
A_51_P230103	11799	NM_001012273	Birc5	0.109383	0.722279	0.316891
A_51_P145433	22343	NM_011699	Lin7c	-0.198696	0.573542	0.014405
A_51_P102782	12040	NM_199195	Bckdhb	-0.110139	0.289971	0.004413
A_51_P350048	14872	NM_010361	Gstt2	-0.077280	0.166682	-0.234791
A_52_P685539	70458	NM_183287	2610318N02Rik	0.043054	0.316519	0.223994
A_51_P444934	20963	NM_011518	Syk	-0.089130	0.707312	0.222800
A_51_P484537	14211	NM_008017	Smc2	0.352481	0.776470	0.098055
A_52_P589591	50776	NM_015810	Polg2	-0.014424	0.299992	-0.125257
A_52_P424826	234723	NM_175646	Txnl4b	0.053971	0.260528	0.014392
A_51_P164014	229841	NM_173762	Cenpe	0.158704	1.026961	0.704244
A_51_P488308	245944	NM_139061	Vps54	0.497890	0.663767	0.500319
A_51_P399853	170753	NM_133218	Zfp704	-0.127418	0.281400	0.057559
A_52_P13448	68026	NM_026515	2810417H13Rik	-0.158120	0.617063	0.078684
A_52_P292651	22171	NM_021288	Tyms	-0.089693	0.415902	0.215983
A_52_P246293	-	CF584628	-	0.129301	0.328717	0.243683
A_52_P144285	108900	NM_175382	2700049P18Rik	-0.299037	0.157204	-0.288116
A_52_P65077	75424	NM_029281	Zfp820	0.462471	0.752423	0.342537
A_51_P487999	72415	NM_028232	Sgol1	0.023518	0.644564	0.261458
A_51_P165330	66845	NM_025796	Mrpl33	-0.478836	0.091295	-0.356476
A_51_P410022	68294	NM_026660	Mfsd10	0.338734	0.525074	0.401849
A_51_P329370	78372	NM_030093	3300001G02Rik	0.056380	0.428540	0.063161
A_51_P198775	381306	NM_201364	BC055324	-0.196100	0.434944	0.098890
A_51_P361022	107995	NM_023223	Cdc20	0.161818	0.744889	0.479738
A_51_P475049	22223	NM_011670	Uchl1	-0.003423	0.715744	0.200428
A_51_P214209	67629	NM_026282	Spc24	-0.216852	0.491829	-0.123768
A_51_P398366	22171	NM_021288	Tyms	-0.086992	0.375547	0.204364
A_51_P398723	14254	NM_010228	Flt1	0.330776	1.022510	0.428526
A_51_P122356	12581	NM_009878	Cdkn2d	0.022473	0.328872	0.097775
A_52_P381430	210789	NM_001081278	Tbc1d4	-0.126144	0.256140	-0.248128
A_51_P383489	60530	NM_021891	Fignl1	-0.486623	0.278605	-0.351556
A_52_P612803	12450	NM_009831	Ccng1	-0.321128	0.272454	-0.109958
A_51_P208355	66634	NM_025676	Mcm8	-0.264899	0.322983	-0.108763
A_51_P312779	18812	NM_011118	Prl2c3	-0.139243	0.865979	0.935153
A_51_P166258	108013	NM_133195	Brunol4	0.985520	1.194323	0.850669
A_51_P220222	21335	NM_001040435	Tacc3	0.165534	0.540364	0.032018
A_51_P246317	17750	NM_008630	Mt2	0.248227	0.725306	0.406048
A_52_P646034	18020	NM_010900	Nfatc2ip	-0.199528	0.382048	0.067864

A_51_P241645	210035	NM_173732	Tmem194	-0.161766	0.340178	0.234965
A_51_P317214	242642	NM_146256	Hpdl	-0.454307	0.160833	-0.302204
A_52_P416879	12040	NM_199195	Bckdhb	-0.207368	0.179763	0.035283
A_51_P422893	75712	NM_029398	Tmem14a	0.778027	1.230935	0.989814
A_51_P480796	75317	NM_029249	4930547N16Rik	-0.017223	0.348003	0.068992
A_51_P409694	56742	NM_019976	Psrc1	0.759753	1.339078	0.884604
A_51_P326555	67358	AK007052	1700093K21Rik	0.036200	0.638140	0.040398
A_51_P328432	638892	XM_915073	LOC638892	0.043155	0.200991	0.070478
A_52_P480360	110074	NM_023595	Dut	0.134780	0.418340	0.099433
A_51_P305052	243958	NM_172900	Siglecg	-0.120133	0.650965	0.059956
A_52_P76931	218973	NM_172598	Wdhd1	-0.254343	0.283271	-0.092610
A_51_P378079	14265	NM_008031	Fmr1	-0.484920	0.349106	-0.308334
A_51_P288447	234865	NM_172288	Nup133	-0.158137	0.086498	-0.044736
A_52_P295712	15465	NM_008285	Hrh1	0.675706	0.940492	0.551285
A_51_P201035	102920	NM_145924	Cenpi	-0.010050	0.458602	0.105453
A_51_P414837	14084	NM_007983	Faf1	0.129865	0.194054	0.050775
A_51_P212491	170768	NM_133232	Pfkfb3	-0.132403	0.445604	0.063764
A_51_P415220	22401	NM_009517	Zmat3	-0.042650	0.985722	0.219492
A_51_P492830	26886	NM_021886	Cenph	-0.133223	0.506231	0.101512
A_51_P204402	20419	NM_011369	Shcbp1	-0.113121	0.586481	0.158701
A_51_P210474	67681	NM_026310	Mrpl18	-0.224645	0.169257	-0.088805
A_52_P130490	-	XM_978341	-	-0.031531	0.255355	-0.132256
A_51_P130830	66403	NM_025541	Asf1a	-0.331590	0.056794	-0.141788
A_51_P327632	78908	NM_207205	Igsf3	0.520298	0.748433	0.392448
A_51_P371190	30939	NM_013917	Pttg1	0.299734	1.069832	0.545071
A_52_P238027	17105	AK159276	Lyz2	0.605838	1.386349	1.037699
A_52_P153929	237911	NM_178309	Brip1	-0.351284	0.529276	0.160021
A_51_P312780	18812	NM_011118	Prl2c3	0.082523	0.958655	0.979082
A_51_P164686	81910	NM_133626	Rrbp1	0.617101	1.070045	0.897379
A_51_P216702	101351	NM_175313	A130022J15Rik	0.345143	0.621220	0.416627
A_51_P104418	63953	NM_022019	Dusp10	0.075571	0.361371	0.135986
A_51_P469449	22256	NM_001040691	Ung	-0.048513	0.206777	-0.060628
A_51_P408071	208628	NM_001042421	Kntc1	-0.449029	0.435407	-0.020943
A_51_P121236	57913	NM_022654	Lrdd	-0.169619	0.393933	-0.046520
A_51_P489285	76464	NM_029617	Casc5	0.257565	0.720519	0.068925
A_51_P491011	320655	NM_001033537	Perld1	-0.220458	0.132452	-0.168606
A_52_P378157	207213	NM_148949	Tdpoz1	-0.133089	0.292015	0.141911
A_51_P137433	100042970	XM_001479207	LOC100042970	-0.308512	0.677229	0.064854
A_51_P495641	16765	NM_019641	Stmn1	-0.090264	0.504739	0.110960
A_51_P343689	108052	NM_028122	Slc14a1	-0.210496	0.512368	-0.095138
A_52_P532982	23886	NM_011819	Gdf15	-0.777296	0.855181	-0.315163
A_51_P376445	18617	NM_008818	Rhox5	0.347839	0.599799	0.331479
A_51_P205106	17349	NM_010801	Mlf1	0.256506	0.650010	0.482854
A_51_P484998	15234	NM_010427	Hgf	0.325410	0.685276	0.645293
A_51_P204442	74016	NM_028716	Phf19	0.166550	0.639988	0.350013
A_51_P365656	76547	NM_029649	Tmem101	-0.276807	0.124609	-0.264206
A_52_P195809	69772	NM_027208	Bdh2	-0.027831	0.599078	0.283077
A_52_P509886	20732	NM_016907	Spint1	-0.145705	0.647227	0.007653
A_51_P489268	54392	NM_019438	Ncapg	-0.113449	0.534187	0.232952
A_51_P108767	19366	NM_009015	Rad54l	-0.090053	0.368480	0.062624
A_52_P75668	-	BB653684	-	-0.149556	0.763410	0.128539
A_51_P467006	67283	NM_026071	Slc25a19	0.118543	0.174309	0.118546
A_52_P268549	72440	NM_197981	5930416I19Rik	0.281892	0.425069	0.315484
A_52_P583438	-	AK137451	-	0.278739	0.353755	0.224677
A_51_P311905	72341	NM_001081381	Tmem103	-0.021635	0.180252	-0.055281
A_51_P239984	26909	NM_012012	Exo1	-0.004580	0.665801	0.300367
A_52_P265745	67120	NM_027619	Ttc14	0.187390	0.541705	0.250618
A_52_P204864	-	AK078008	-	0.010110	0.638690	0.234831
A_51_P493467	110033	NM_145588	Kif22	0.197316	0.664512	0.406106
A_51_P329332	116914	NM_054087	Slc19a2	-0.110755	0.703382	0.033826
A_52_P254095	17470	NM_010818	Cd200	-0.443440	0.518611	-0.201235
A_51_P357592	14942	NM_010373	Gzme	-0.114433	0.892433	0.406476
A_52_P497553	13361	NM_010049	Dhfr	-0.150914	0.118442	-0.271838
A_51_P348624	30059	NM_013899	Timm10	-0.106886	0.075062	-0.170038
A_51_P118535	76967	BC043104	2700049A03Rik	-0.071885	0.262808	0.043573
A_52_P205951	209334	AK031077	Gen1	-0.209617	0.374039	-0.102326
A_51_P482265	-	AK163132	-	0.680218	0.603781	0.491484
A_52_P18775	230582	NM_175471	2810410C14Rik	0.050333	0.302498	0.077896
A_51_P134468	110956	L78788	D17H6S56E-5	0.089833	1.030285	0.478255
A_52_P143671	230098	NM_001013377	E130306D19Rik	-0.249982	0.179470	-0.120679
A_51_P451588	27276	NM_013746	Plekhb1	0.231172	0.427923	0.274630

A_52_P1197950	-	AI323028	-	-0.021233	0.383088	0.139927
A_51_P499940	97165	NM_008252	Hmgb2	0.397634	0.607375	0.315363
A_52_P116704	-	-	-	0.440695	0.723766	0.485778
A_52_P8922	20650	NM_009229	Sntb2	0.375736	0.790692	0.398503
A_52_P371108	23834	NM_011799	Cdc6	-0.170966	0.253911	0.039286
A_51_P457528	12442	NM_007630	Ccnb2	0.050230	0.739718	0.361083
A_52_P214479	57344	NM_020577	As3mt	0.161841	0.873179	0.338298
A_52_P348560	-	-	-	0.315299	0.455041	0.359498
A_51_P501018	18005	NM_010892	Nek2	0.004980	0.546750	0.265946
A_52_P227746	-	-	-	0.247077	0.662175	0.545102
A_52_P587611	13555	NM_007891	E2f1	0.145158	0.421695	0.174005
A_51_P392529	76799	NM_029748	2510006D16Rik	0.072900	0.174491	-0.003114
A_51_P307721	12404	NM_019626	Cbln1	-0.506512	0.241309	-0.260045
A_51_P125135	67849	NM_026410	Cdca5	-0.162523	0.417349	0.152414
A_51_P159042	-	-	-	0.022654	0.964962	0.509672
A_52_P27878	230098	NM_001013377	E130306D19Rik	0.040228	0.610442	0.016145
A_51_P345159	67196	NM_026024	Ube2t	0.130844	0.680889	0.346140
A_52_P654565	78412	NM_199145	3110062M04Rik	-0.044926	0.287369	-0.183545
A_51_P440584	54608	NM_018811	Abhd2	-0.095930	0.108723	0.031614
A_52_P529570	381318	NM_198654	Nsl1	-0.129167	0.619369	0.201981
A_51_P310298	231549	NM_178701	Lrrc8d	-0.228157	0.467859	-0.079533
A_52_P118310	108699	NM_175752	Chn1	-0.546045	0.328827	-0.267105
A_51_P314652	320302	NM_177130	Glt28d2	-0.270026	0.234924	-0.089214
A_51_P227004	54124	NM_016904	Cks1b	0.356814	0.582868	0.483254
A_51_P251069	19212	NM_008961	Pter	0.224511	0.642015	0.395095
A_51_P314418	12144	NM_007550	Blm	-0.202142	0.259035	-0.133806
A_52_P403157	234214	NM_172752	Sorbs2	-0.292693	0.386923	-0.188343
A_51_P242201	67111	NM_025972	Naaa	-0.021837	0.474808	0.193703
A_51_P229633	70470	NM_027434	2610304G08Rik	0.394443	0.166534	-0.008610
A_52_P660242	57905	NM_133934	Isy1	0.166414	0.279563	0.202045
A_51_P438149	212307	NM_153058	Mapre2	0.331638	1.003413	0.767484
A_51_P337089	69270	XM_130416	Gins1	-0.316953	0.204544	-0.108955
A_52_P335354	237436	NM_001079876	Gas2l3	-0.033872	0.759019	0.302686
A_51_P419768	19735	NM_009061	Rgs2	0.683862	1.368485	0.925070
A_52_P584374	224171	NM_172616	C330027C09Rik	0.003166	0.420486	0.058133
A_52_P297457	68364	XM_132614	0610030E20Rik	0.024638	0.381790	0.140079
A_51_P454666	65105	NM_144509	Arl6ip4	0.153023	0.354540	0.186137
A_51_P451335	18779	NM_008867	Pla2r1	0.192324	0.634197	0.398950
A_51_P444437	67103	NM_025968	Ltb4dh	0.335745	0.607914	0.456978
A_51_P362104	83965	NM_032003	Enpp5	0.370482	0.793749	0.427786
A_52_P322364	74570	NM_133906	Zkscan1	-0.035150	0.183163	-0.015347
A_51_P139320	13180	NM_025273	Pcbd1	-0.200402	0.086388	-0.215137
A_51_P153224	66220	NM_025428	Zdhhc12	0.138466	0.452530	0.286790
A_51_P212682	70024	NM_027290	Mcm10	-0.078179	0.592166	0.206239
A_51_P324450	76400	NM_029595	Pbp2	0.374501	1.157207	0.740847
A_51_P255805	229474	NM_001033301	Fhdc1	0.286773	0.492496	0.335424
A_51_P415905	18968	NM_008892	Pola1	-0.117663	0.397258	0.227759
A_51_P511899	234814	NM_172761	Mthfsd	0.619979	0.797473	0.676737
A_52_P302395	240028	NM_172827	Lnpep	0.385161	0.723210	0.582364
A_51_P327983	235345	NM_177702	4833427G06Rik	0.165120	1.244541	0.282270
A_51_P279575	12532	NM_009860	Cdc25c	-0.317459	0.500863	0.116133
A_51_P173022	67711	NM_026330	Nsmce1	-0.037039	0.005489	-0.003872
A_52_P255034	75317	NM_029249	4930547N16Rik	0.073917	0.409684	0.192170
A_52_P502771	623474	NM_001039556	E130016E03Rik	-0.149878	0.572139	0.019663
A_51_P324934	17215	AK088142	Mcm3	-0.045123	0.404860	0.038583
A_51_P127615	20492	NM_009193	Slbp	0.109243	0.393380	0.233620
A_51_P273979	12615	NM_007681	Cenpa	0.394648	1.095618	0.886135
A_52_P220370	234396	AK083552	Ankle1	-0.053902	0.348479	0.099619
A_52_P518142	75991	NM_153567	Slain2	0.216126	0.557750	0.557720
A_51_P514700	66442	NM_025565	Spc25	-0.091059	0.503110	0.151728
A_51_P221291	20650	NM_009229	Sntb2	0.480032	0.416800	0.008983
A_52_P190339	66578	NM_025642	2610039C10Rik	0.206195	0.433384	0.315199
A_52_P28806	14235	NM_008021	Foxm1	-0.059791	0.622294	0.302194
A_52_P148553	60530	NM_021891	Fignl1	-0.413830	0.278540	-0.287818
A_51_P489192	50706	NM_015784	Postn	-0.524319	1.152738	-0.305993
A_51_P186178	216152	NM_181681	BC005764	-0.306493	0.858056	0.389762
A_52_P199375	19175	NM_008946	Psmb6	-0.227689	0.079160	-0.133743
A_51_P243168	381259	NM_001037812	Als2cr4	-0.100652	0.331748	0.040292
A_51_P499233	224171	NM_172616	C330027C09Rik	-0.133919	0.594401	0.168184
A_51_P164998	20257	NM_025285	Stmn2	0.059277	0.887235	0.578087
A_51_P444696	108912	NM_175384	Cdca2	0.351332	0.653219	0.204683

A_51_P327322	60411	NM_021790	Cenpk	-0.124406	0.595410	0.013251
A_51_P487791	11555	NM_007420	Adrb2	0.253942	0.497384	0.157773
A_51_P472217	72080	NM_001081085	2010317E24Rik	0.267582	0.660752	0.407043
A_51_P430263	100678	NM_133900	Psph	-0.608011	0.242107	-0.241737
A_52_P92302	56709	AK089427	Dnajb12	-0.238234	0.186622	0.001972
A_51_P151362	72667	NM_028316	Zfp444	0.454121	0.355403	0.327551
A_51_P133137	19348	NM_009004	Kif20a	-0.098233	0.618721	0.276021
A_52_P796682	12447	NM_007633	Ccne1	-0.421218	0.184894	-0.215249
A_51_P113003	83815	NM_031863	Cenpq	-0.126658	0.300085	0.025194
A_51_P166155	57441	NM_020567	Gmnn	0.147340	0.707069	0.403128
A_52_P582814	68395	BC025170	0610037M15Rik	-0.146924	0.233709	0.188125
A_52_P569375	14176	NM_010203	Fgf5	-0.740737	0.872093	-0.006007
A_52_P364970	66433	NM_181391	Chchd7	0.040135	0.256706	0.089717
A_51_P288277	-	AK003987	-	-0.134504	0.415931	0.050009
A_51_P217778	12316	NM_009791	Aspm	0.083589	0.711558	0.478818
A_52_P284441	76380	NM_029606	Ccdc46	0.147538	0.539588	0.507464
A_52_P4666	60411	NM_021790	Cenpk	-0.218524	0.575136	0.061049
A_52_P223495	16401	NM_010576	Itga4	-0.152992	0.584258	-0.091074
A_51_P401263	268465	NM_177752	Eme1	0.552314	0.697232	0.335814
A_51_P378789	55985	NM_018866	Cxcl13	-0.758386	1.841365	-0.461668
A_51_P463303	67553	AK033409	Gstcd	-0.096554	0.108793	-0.197655
A_52_P293222	170826	AK042378	Ppargc1b	0.053737	0.586857	-0.079735
A_51_P155142	52276	NM_026560	Cdca8	0.223512	0.658721	0.399570
A_52_P619248	107849	NM_181852	Prl2c5	-0.030242	0.942662	1.006110
A_51_P161612	12450	NM_009831	Ccng1	0.422577	0.779186	0.637600
A_51_P367720	12759	NM_013492	Clu	-0.837949	1.035270	-0.428392
A_52_P279845	81840	AK047881	Sorcs2	-0.051847	0.526255	0.212759
A_51_P424810	76044	NM_133762	Ncapg2	-0.137344	0.403043	0.076258
A_51_P111762	433375	NM_011804	Creg1	0.361343	0.289417	0.315978
A_52_P302012	66531	NM_026844	2310061C15Rik	0.397302	0.630733	0.258818
A_51_P295206	237911	NM_178309	Brip1	0.108832	0.554459	0.156309
A_52_P641185	19247	NM_011202	Ptpn11	0.432887	0.378635	0.094551
A_51_P507234	68190	XM_992800	5330426P16Rik	-0.064430	0.317703	0.047389
A_52_P402960	-	AI846867	-	0.152534	0.550105	0.181475
A_52_P187855	68729	NM_197987	Trim37	0.735874	0.467343	0.099948
A_51_P505172	106795	NM_025674	Tcf19	-0.155669	0.725718	0.329251
A_51_P452876	11636	NM_021515	Ak1	-0.096694	0.266849	0.132023
A_51_P162471	108671	NM_134081	Dnajc9	-0.164234	0.411398	0.099176
A_52_P396774	72657	NM_001037279	2700094K13Rik	0.274239	0.626717	0.410971
A_51_P169576	-	-	-	0.282359	0.297950	0.101354
A_51_P253808	17345	NM_001081117	Mki67	0.181059	0.755915	0.163743
A_52_P550734	16580	NM_053173	Kifc1	-0.162154	0.803675	0.206443
A_51_P253803	17345	NM_001081117	Mki67	0.227054	0.863258	0.300787
A_51_P242560	66531	NM_026844	2310061C15Rik	0.425334	0.743768	0.326982
A_52_P560792	68066	NM_026542	Slc25a39	0.312467	0.412130	0.372090
A_51_P254912	19718	NM_020022	Rfc2	0.044031	0.109632	0.012674
A_51_P420276	67448	NM_026162	Plxdc2	-1.193398	0.334436	-0.692837
A_51_P330213	66929	NM_024184	Asf1b	-0.320917	0.221222	0.066809
A_52_P683306	225358	NM_146084	2610024E20Rik	0.220044	0.609093	0.329167
A_52_P629112	-	-	-	0.407782	0.711434	0.568283
A_52_P16657	72477	NM_028248	Tmem87b	0.342432	0.444014	0.147324
A_52_P392544	52276	NM_026560	Cdca8	0.080916	0.474219	0.224477
A_51_P480233	67824	NM_026393	Nmral1	-0.290221	0.090467	-0.136440
A_52_P636199	14211	NM_008017	Smc2	0.378703	0.964919	0.298429
A_51_P408946	12447	NM_007633	Ccne1	-0.421488	0.329969	-0.028272
A_51_P256093	26399	NM_011943	Map2k6	0.155541	0.732816	0.261450
A_52_P1126526	98415	NM_175294	Nucks1	0.396605	0.479994	0.108680
A_51_P254045	22036	NM_011634	Traip	0.062088	0.644438	0.311650
A_51_P366277	70930	NM_001081350	Nol8	0.195117	0.252491	-0.262818
A_51_P351363	22172	M30774	Tyms-ps	-0.046470	0.451076	0.218296
A_52_P364776	17215	NM_008563	Mcm3	-0.329826	0.216662	-0.097169
A_51_P195808	70454	NM_027429	Cenpl	0.103678	0.517592	0.236972
A_51_P366931	233406	NM_145150	Prc1	-0.202864	0.603561	0.400100
A_51_P410900	171167	NM_001012517	Fut10	0.269640	0.571518	0.403451
A_52_P8216	67693	NM_026318	2310003F16Rik	-0.118402	-0.084951	-0.234232
A_52_P479539	12704	AK047894	Cit	0.525670	0.912836	0.568810
A_51_P138341	66408	NM_025545	Aptx	0.552004	0.551453	0.390558
A_52_P329197	13361	NM_010049	Dhfr	-0.007109	0.335040	0.022049
A_51_P178435	56330	NM_019746	Pdcd5	-0.186546	0.094314	-0.195577
A_51_P458538	66414	NM_025551	Ndufa12	-0.406199	-0.025363	-0.287204
A_51_P513682	66977	NM_023284	Nuf2	-0.142184	0.701524	0.319550

A_51_P315530	17121	NM_016662	Mxd3	-0.221441	0.543168	0.148785
A_51_P142113	14533	NM_015740	Bloc1s1	-0.198623	0.008213	-0.060877
A_51_P464911	76467	NM_029619	Msrb2	-0.058701	0.748268	0.211267
A_51_P295192	18035	NM_010907	Nfkbia	0.582841	0.681694	0.559881
A_51_P124315	-	BC034076	-	0.600265	0.921190	0.573711
A_51_P158388	78334	NM_198164	Cdc2l6	0.508631	0.780866	0.633347
A_52_P658437	105988	NM_001014976	Espl1	-0.118167	0.507419	0.284147
A_51_P264769	13406	NM_016779	Dmp1	-0.549385	1.202153	-0.289359

Table S 5 – Cluster #1 Gene Ontology results.
Genes from Cluster #1 were analyzed by GO. The results from the analysis are presented in the following table. The enrichment of each category, calculated from the Count, List Total, Pop Hits and Pop Total, are indicated. The corresponding False Discovery Rate (FDR) and p-values corrected by Benjamini-Hochberg multiple hypothesis are also indicated.

Gene Ontology ID and Term	Count	List Total	Pop Hits	Pop Total	Enrichment	Benjamini	FDR
GO:0003012~muscle system process	18	371	92	14312	7.548	1.499E-07	2.669E-07
GO:0006936~muscle contraction	18	348	92	13294	7.474	3.561E-07	3.320E-07
GO:0007517~muscle development	23	348	185	13294	4.749	2.635E-06	4.914E-06
GO:0006941~striated muscle contraction	11	325	32	12587	13.313	5.659E-06	7.692E-06
GO:0007517~muscle development	23	325	185	12587	4.815	5.671E-06	3.855E-06
GO:0014706~striated muscle development	18	325	138	12587	5.052	7.831E-05	1.597E-04
GO:0048747~muscle fiber development	13	371	71	14312	7.063	1.054E-04	3.754E-04
GO:0048747~muscle fiber development	13	325	71	12587	7.091	1.477E-04	4.016E-04
GO:0048741~skeletal muscle fiber development	13	348	71	13294	6.995	1.613E-04	4.513E-04
GO:0045445~myoblast differentiation	9	325	42	12587	8.299	5.060E-03	1.724E-02
GO:0042692~muscle cell differentiation	11	348	73	13294	5.756	8.439E-03	3.160E-02
GO:0055001~muscle cell development	6	371	18	14312	12.859	2.098E-02	1.132E-01
GO:0055001~muscle cell development	6	325	18	12587	12.910	3.049E-02	1.262E-01
GO:0048628~myoblast maturation	6	371	24	14312	9.644	5.403E-02	4.932E-01
GO:0048627~myoblast development	6	371	25	14312	9.258	5.508E-02	6.033E-01
GO:0006937~regulation of muscle contraction	7	371	36	14312	7.501	6.119E-02	4.486E-01
GO:0048628~myoblast maturation	6	348	24	13294	9.550	9.612E-02	5.638E-01
GO:0006937~regulation of muscle contraction	7	325	36	12587	7.531	9.968E-02	4.983E-01
GO:0006937~regulation of muscle contraction	7	348	36	13294	7.428	1.051E-01	5.162E-01
GO:0048627~myoblast development	6	325	25	12587	9.295	1.168E-01	6.731E-01
GO:0048469~cell maturation	9	371	86	14312	4.037	1.689E-01	2.601
GO:0009057~macromolecule catabolic process	20	371	351	14312	2.198	1.745E-01	3.025
GO:0048731~system development	66	371	1749	14312	1.456	1.752E-01	2.372
GO:0030239~myofibril assembly	5	348	17	13294	11.236	1.944E-01	1.401
GO:0055002~striated muscle cell development	5	348	17	13294	11.236	1.944E-01	1.401
GO:0030239~myofibril assembly	5	325	17	12587	11.391	2.039E-01	1.385
GO:0044248~cellular catabolic process	24	371	484	14312	1.913	2.808E-01	5.698
GO:0048878~chemical homeostasis	16	348	242	13294	2.526	3.005E-01	2.955
GO:0051146~striated muscle cell differentiation	6	325	33	12587	7.042	3.129E-01	2.518
GO:0048469~cell maturation	9	325	86	12587	4.053	3.226E-01	2.870
GO:0051239~regulation of multicellular organismal process	17	371	309	14312	2.122	3.234E-01	9.906
GO:0055082~cellular chemical homeostasis	12	371	178	14312	2.601	3.392E-01	9.811
GO:0009308~amine metabolic process	20	371	388	14312	1.988	3.524E-01	8.863
GO:0006082~organic acid metabolic process	24	371	507	14312	1.826	3.568E-01	9.709
GO:0030154~cell differentiation	66	371	1847	14312	1.378	3.598E-01	8.362
GO:0006942~regulation of striated muscle contraction	3	348	4	13294	28.651	3.810E-01	6.489
GO:0048468~cell development	49	348	1233	13294	1.518	3.849E-01	5.293
GO:0007010~cytoskeleton organization and biogenesis	26	348	537	13294	1.850	3.911E-01	6.272
GO:0042592~homeostatic process	19	371	375	14312	1.955	3.912E-01	13.181
GO:0043285~biopolymer catabolic process	17	348	279	13294	2.328	3.941E-01	4.565
GO:0044262~cellular carbohydrate metabolic process	17	348	282	13294	2.303	3.986E-01	5.082
GO:0030163~protein catabolic process	14	348	211	13294	2.535	4.020E-01	6.042
GO:0045214~sarcomere organization	4	325	11	12587	14.083	4.057E-01	4.155
GO:0030163~protein catabolic process	14	325	211	12587	2.570	4.295E-01	5.561
GO:0008366~axon ensheathment	5	371	34	14312	5.673	4.345E-01	15.844
GO:0007026~negative regulation of microtubule depolymerization	4	325	12	12587	12.910	4.421E-01	5.402
GO:0031114~regulation of microtubule depolymerization	4	325	12	12587	12.910	4.421E-01	5.402
GO:0006811~ion transport	32	371	784	14312	1.575	4.508E-01	17.471
GO:0006942~regulation of striated muscle contraction	3	325	4	12587	29.047	4.650E-01	6.575
GO:0006066~alcohol metabolic process	15	371	282	14312	2.052	4.682E-01	21.030

GO:0009894~regulation of catabolic process	4	371	21	14312	7.348	4.752E-01	22.315
GO:0005975~carbohydrate metabolic process	20	371	425	14312	1.815	4.777E-01	20.645
GO:0048513~organ development	51	371	1416	14312	1.389	4.793E-01	19.808
GO:0022008~neurogenesis	18	371	373	14312	1.862	4.841E-01	23.736
GO:0019752~carboxylic acid metabolic process	24	348	506	13294	1.812	5.348E-01	11.424
GO:0031111~negative regulation of microtubule polymerization or depolymerization	4	325	14	12587	11.065	5.352E-01	8.471
GO:0008016~regulation of heart contraction	5	371	41	14312	4.704	5.370E-01	28.033
GO:0031110~regulation of microtubule polymerization or depolymerization	4	348	16	13294	9.550	5.416E-01	12.273
GO:0006873~cellular ion homeostasis	12	348	178	13294	2.575	5.549E-01	11.377
GO:0050801~ion homeostasis	13	325	200	12587	2.517	5.601E-01	9.559
GO:0019538~protein metabolic process	102	371	3268	14312	1.204	5.692E-01	31.258
GO:0055082~cellular chemical homeostasis	12	325	178	12587	2.611	5.845E-01	10.721
GO:0042552~myelination	5	348	32	13294	5.969	5.888E-01	14.567
GO:0001508~regulation of action potential	5	371	44	14312	4.384	5.934E-01	34.065
GO:0031110~regulation of microtubule polymerization or depolymerization	4	325	16	12587	9.682	6.192E-01	12.301
GO:0009892~negative regulation of metabolic process	18	371	398	14312	1.745	6.311E-01	38.084
GO:0007272~ensheathment of neurons	5	348	34	13294	5.618	6.488E-01	17.729
GO:0044260~cellular macromolecule metabolic process	98	371	3165	14312	1.194	6.489E-01	40.653
GO:0006811~ion transport	32	348	784	13294	1.559	6.860E-01	21.148
GO:0044265~cellular macromolecule catabolic process	15	348	275	13294	2.084	6.894E-01	20.464
GO:0048519~negative regulation of biological process	37	371	1029	14312	1.387	6.955E-01	45.875
GO:0048513~organ development	51	348	1416	13294	1.376	7.226E-01	24.944
GO:0009894~regulation of catabolic process	4	348	21	13294	7.276	7.337E-01	24.705
GO:0008366~axon ensheathment	5	325	34	12587	5.695	7.428E-01	17.619
GO:0007272~ensheathment of neurons	5	325	34	12587	5.695	7.428E-01	17.619
GO:0008016~regulation of heart contraction	5	348	41	13294	4.659	7.970E-01	31.048
GO:0043412~biopolymer modification	59	348	1717	13294	1.313	7.973E-01	32.083
GO:0043632~modification-dependent macromolecule catabolic process	10	325	152	12587	2.548	8.416E-01	25.957
GO:0008033~tRNA processing	5	371	58	14312	3.326	8.478E-01	63.412
GO:0022008~neurogenesis	18	325	373	12587	1.869	8.480E-01	25.510
GO:0008610~lipid biosynthetic process	12	371	256	14312	1.808	8.575E-01	67.043
GO:0051246~regulation of protein metabolic process	13	371	286	14312	1.753	8.628E-01	66.583
GO:0044257~cellular protein catabolic process	10	325	157	12587	2.467	8.824E-01	30.492
GO:0009892~negative regulation of metabolic process	18	348	398	13294	1.728	8.902E-01	42.662
GO:0006519~amino acid and derivative metabolic process	14	371	329	14312	1.642	8.990E-01	75.036
GO:0006996~organelle organization and biogenesis	41	371	1241	14312	1.274	9.049E-01	74.903
GO:0015931~nucleobase, nucleoside, nucleotide and nucleic acid transport	5	371	66	14312	2.922	9.075E-01	77.304
GO:0003015~heart process	5	348	49	13294	3.898	9.155E-01	48.730
GO:0048699~generation of neurons	16	348	345	13294	1.772	9.166E-01	47.721
GO:0006464~protein modification process	56	325	1648	12587	1.316	9.195E-01	35.927
GO:0001508~regulation of action potential	5	325	44	12587	4.401	9.221E-01	37.401
GO:0019725~cellular homeostasis	13	348	261	13294	1.903	9.241E-01	51.380
GO:0044267~cellular protein metabolic process	96	348	3123	13294	1.174	9.499E-01	57.920
GO:0006399~tRNA metabolic process	8	325	118	12587	2.626	9.549E-01	44.549
GO:0002026~regulation of the force of heart contraction	3	348	15	13294	7.640	9.648E-01	63.162
GO:0060047~heart contraction	5	325	49	12587	3.952	9.659E-01	48.620
GO:0006357~regulation of transcription from RNA polymerase II promoter	18	325	415	12587	1.680	9.687E-01	51.791
GO:0006366~transcription from RNA polymerase II promoter	19	348	464	13294	1.564	9.698E-01	65.936
GO:0030003~cellular cation homeostasis	8	325	123	12587	2.519	9.702E-01	51.161
GO:0051246~regulation of protein metabolic process	13	348	286	13294	1.736	9.782E-01	72.316
GO:0019226~transmission of nerve impulse	12	348	254	13294	1.805	9.794E-01	70.801

GO:0051248~negative regulation of protein metabolic process	6	348	88	13294	2.605	9.799E-01	75.828
GO:0008610~lipid biosynthetic process	12	348	256	13294	1.791	9.803E-01	72.244
GO:0042176~regulation of protein catabolic process	3	348	18	13294	6.367	9.807E-01	75.325
GO:0007399~nervous system development	25	348	674	13294	1.417	9.815E-01	74.750
GO:0051129~negative regulation of cellular component organization and biogenesis	4	348	40	13294	3.820	9.833E-01	78.275
GO:0045988~negative regulation of striated muscle contraction	2	325	2	12587	38.729	9.857E-01	60.329
GO:0007507~heart development	9	348	178	13294	1.932	9.868E-01	81.663
GO:0015931~nucleobase, nucleoside, nucleotide and nucleic acid transport	5	348	66	13294	2.894	9.873E-01	81.182
GO:0002026~regulation of the force of heart contraction	3	325	15	12587	7.746	9.891E-01	63.739
GO:0051261~protein depolymerization	4	325	37	12587	4.187	9.952E-01	71.881
GO:0008610~lipid biosynthetic process	12	325	256	12587	1.815	9.952E-01	70.938
GO:0042176~regulation of protein catabolic process	3	325	18	12587	6.455	9.953E-01	75.880
GO:0032787~monocarboxylic acid metabolic process	11	325	228	12587	1.869	9.954E-01	73.157
GO:0051248~negative regulation of protein metabolic process	6	325	88	12587	2.641	9.957E-01	75.582
GO:0045989~positive regulation of striated muscle contraction	2	325	3	12587	25.819	9.960E-01	75.015
GO:0051129~negative regulation of cellular component organization and biogenesis	4	325	40	12587	3.873	9.965E-01	78.521
GO:0019318~hexose metabolic process	8	325	149	12587	2.079	9.968E-01	81.318
GO:0030029~actin filament-based process	10	325	207	12587	1.871	9.970E-01	80.154
GO:0007507~heart development	9	325	178	12587	1.958	9.970E-01	80.942
GO:0030182~neuron differentiation	13	325	304	12587	1.656	9.972E-01	82.824
GO:0015858~nucleoside transport	2	325	4	12587	19.365	9.973E-01	84.264
GO:0005996~monosaccharide metabolic process	8	325	152	12587	2.038	9.974E-01	83.861

Table S 6 – Cluster #3 Gene Ontology results.
Genes from Cluster #3 were analyzed by GO. The results from the analysis are presented in the following table. The enrichment of each category, calculated from the Count, List Total, Pop Hits and Pop Total, are indicated. The corresponding False Discovery Rate (FDR) and p-values corrected by Benjamini-Hochberg multiple hypothesis are also indicated.

Gene Ontology ID and Term	Count	List Total	Pop Hits	Pop Total	Enrichment	Benjamini	FDR
GO:0000279~M phase	40	194	256	13294	10.7072	9.650E-26	8.999E-26
GO:0022402~cell cycle process	54	203	596	14312	6.3878	3.559E-25	6.335E-25
GO:0000279~M phase	40	192	256	12587	10.2433	6.637E-25	4.511E-25
GO:0022403~cell cycle phase	41	203	306	14312	9.4464	7.564E-25	2.693E-24
GO:0022403~cell cycle phase	41	194	306	13294	9.1816	3.775E-24	7.040E-24
GO:0000087~M phase of mitotic cell cycle	34	194	193	13294	12.0719	1.358E-23	5.066E-23
GO:0007067~mitosis	34	194	192	13294	12.1348	1.520E-23	4.251E-23
GO:0000087~M phase of mitotic cell cycle	34	192	193	12587	11.5490	9.928E-23	2.024E-22
GO:0007067~mitosis	34	192	192	12587	11.6091	1.251E-22	1.700E-22
GO:0000278~mitotic cell cycle	36	203	260	14312	9.7619	4.679E-22	2.499E-21
GO:0006259~DNA metabolic process	36	194	682	13294	3.6172	1.826E-08	8.512E-08
GO:0006974~response to DNA damage stimulus	19	203	264	14312	5.0740	8.068E-06	5.745E-05
GO:0006281~DNA repair	17	194	211	13294	5.5210	2.201E-05	1.232E-04
GO:0006260~DNA replication	16	192	170	12587	6.1701	2.904E-05	7.895E-05
GO:0006281~DNA repair	17	192	211	12587	5.2819	6.819E-05	2.317E-04
GO:0051726~regulation of cell cycle	18	203	374	14312	3.3932	3.343E-03	3.576E-02
GO:0051321~meiotic cell cycle	9	203	82	14312	7.7381	3.479E-03	3.102E-02
GO:0051276~chromosome organization and biogenesis	19	194	402	13294	3.2388	4.616E-03	3.882E-02
GO:0051327~M phase of meiotic cell cycle	9	194	81	13294	7.6140	5.047E-03	3.774E-02
GO:0007126~meiosis	9	194	81	13294	7.6140	5.047E-03	3.774E-02
GO:0000074~regulation of progression through cell cycle	18	194	371	13294	3.3247	5.308E-03	4.961E-02
GO:0051726~regulation of cell cycle	18	194	374	13294	3.2980	5.322E-03	5.472E-02
GO:0000910~cytokinesis	6	203	33	14312	12.8186	1.136E-02	1.423E-01
GO:0007126~meiosis	9	192	81	12587	7.2841	1.309E-02	5.371E-02
GO:0051327~M phase of meiotic cell cycle	9	192	81	12587	7.2841	1.309E-02	5.371E-02
GO:0006996~organelle organization and biogenesis	35	203	1241	14312	1.9884	1.402E-02	2.009E-01
GO:0000074~regulation of progression through cell cycle	18	192	371	12587	3.1807	1.618E-02	8.864E-02
GO:0000070~mitotic sister chromatid segregation	5	203	23	14312	15.3266	2.643E-02	4.282E-01
GO:0000819~sister chromatid segregation	5	203	24	14312	14.6880	2.820E-02	5.080E-01
GO:0006139~nucleobase, nucleoside, nucleotide and nucleic acid metabolic process	72	203	3502	14312	1.4495	3.142E-02	6.231E-01
GO:0051641~cellular localization	27	203	913	14312	2.0850	3.502E-02	7.586E-01
GO:0007017~microtubule-based process	13	192	220	12587	3.8738	3.794E-02	2.363E-01
GO:0000070~mitotic sister chromatid segregation	5	194	23	13294	14.8969	4.562E-02	5.211E-01
GO:0051649~establishment of cellular localization	26	203	894	14312	2.0504	5.370E-02	1.269
GO:0000070~mitotic sister chromatid segregation	5	192	23	12587	14.2516	9.033E-02	6.414E-01
GO:0007018~microtubule-based movement	9	192	119	12587	4.9581	9.378E-02	7.999E-01
GO:0000819~sister chromatid segregation	5	192	24	12587	13.6578	9.703E-02	7.601E-01
GO:0030705~cytoskeleton-dependent intracellular transport	9	194	135	13294	4.5684	1.044E-01	1.328
GO:0051649~establishment of cellular localization	26	194	894	13294	1.9929	1.445E-01	2.017
GO:0051303~establishment of chromosome localization	3	203	6	14312	35.2512	1.636E-01	4.355
GO:0046907~intracellular transport	21	203	724	14312	2.0450	1.689E-01	4.820
GO:0030705~cytoskeleton-dependent intracellular transport	9	192	135	12587	4.3705	1.765E-01	1.831
GO:0006334~nucleosome assembly	7	192	76	12587	6.0382	1.856E-01	1.797
GO:0045132~meiotic chromosome segregation	3	203	7	14312	30.2153	1.950E-01	5.993
GO:0031497~chromatin assembly	7	194	90	13294	5.3298	2.126E-01	3.289
GO:0043283~biopolymer metabolic process	86	203	4747	14312	1.2773	2.167E-01	7.124
GO:0050000~chromosome localization	3	194	6	13294	34.2629	2.922E-01	5.025

GO:0051303~establishment of chromosome localization	3	194	6	13294	34.2629	2.922E-01	5.025
GO:0046907~intracellular transport	21	194	724	13294	1.9876	3.535E-01	7.059
GO:0006270~DNA replication initiation	4	192	19	12587	13.8015	3.844E-01	4.826
GO:0022607~cellular component assembly	14	203	444	14312	2.2231	3.988E-01	15.044
GO:0051303~establishment of chromosome localization	3	192	6	12587	32.7786	4.166E-01	5.692
GO:0050000~chromosome localization	3	192	6	12587	32.7786	4.166E-01	5.692
GO:0017038~protein import	6	194	80	13294	5.1394	4.282E-01	9.901
GO:0065004~protein-DNA complex assembly	7	194	112	13294	4.2829	4.285E-01	9.437
GO:0051651~maintenance of cellular localization	3	203	13	14312	16.2698	4.754E-01	19.603
GO:0045132~meiotic chromosome segregation	3	192	7	12587	28.0960	4.854E-01	7.807
GO:0051640~organelle localization	4	203	37	14312	7.6219	4.856E-01	21.076
GO:0051169~nuclear transport	7	194	119	13294	4.0309	4.916E-01	12.408
GO:0048519~negative regulation of biological process	24	203	1029	14312	1.6444	5.028E-01	25.808
GO:0065003~macromolecular complex assembly	12	203	380	14312	2.2264	5.157E-01	25.687
GO:0048523~negative regulation of cellular process	23	203	963	14312	1.6839	5.162E-01	23.771
GO:0007076~mitotic chromosome condensation	3	203	15	14312	14.1005	5.219E-01	25.099
GO:0017038~protein import	6	192	80	12587	4.9168	5.988E-01	12.221
GO:0065004~protein-DNA complex assembly	7	192	112	12587	4.0973	6.079E-01	11.950
GO:0046907~intracellular transport	21	192	724	12587	1.9015	6.138E-01	11.562
GO:0008608~attachment of spindle microtubules to kinetochore	2	203	2	14312	70.5025	6.290E-01	35.683
GO:0051169~nuclear transport	7	192	119	12587	3.8563	6.462E-01	15.592
GO:0006310~DNA recombination	6	192	85	12587	4.6276	6.577E-01	15.431
GO:0006913~nucleocytoplasmic transport	7	192	118	12587	3.8890	6.637E-01	15.035
GO:0051651~maintenance of cellular localization	3	194	13	13294	15.8136	7.076E-01	22.295
GO:0051640~organelle localization	4	194	37	13294	7.4082	7.283E-01	24.380
GO:0015031~protein transport	17	203	708	14312	1.6929	7.581E-01	49.445
GO:0045143~homologous chromosome segregation	2	203	3	14312	47.0016	7.608E-01	48.420
GO:0007076~mitotic chromosome condensation	3	194	15	13294	13.7052	7.752E-01	28.393
GO:0065003~macromolecular complex assembly	12	194	380	13294	2.1640	8.034E-01	32.589
GO:0048523~negative regulation of cellular process	23	194	963	13294	1.6366	8.148E-01	32.505
GO:0007010~cytoskeleton organization and biogenesis	15	194	537	13294	1.9141	8.176E-01	34.841
GO:0051310~metaphase plate congression	2	203	4	14312	35.2512	8.191E-01	58.636
GO:0051220~cytoplasmic sequestering of protein	2	203	4	14312	35.2512	8.191E-01	58.636
GO:0051656~establishment of organelle localization	3	203	27	14312	7.8336	8.296E-01	58.609
GO:0008104~protein localization	18	203	811	14312	1.5648	8.389E-01	65.785
GO:0051049~regulation of transport	5	203	110	14312	3.2047	8.458E-01	67.752
GO:0006606~protein import into nucleus	5	192	66	12587	4.9665	8.470E-01	27.310
GO:0051098~regulation of binding	3	203	30	14312	7.0502	8.472E-01	65.707
GO:0045184~establishment of protein localization	17	203	750	14312	1.5981	8.508E-01	65.009
GO:0051223~regulation of protein transport	3	194	19	13294	10.8199	8.670E-01	40.943
GO:0007076~mitotic chromosome condensation	3	192	15	12587	13.1115	8.825E-01	31.508
GO:0006520~amino acid metabolic process	9	194	258	13294	2.3904	8.902E-01	44.974
GO:0046822~regulation of nucleocytoplasmic transport	3	194	21	13294	9.7894	8.973E-01	47.099
GO:0006519~amino acid and derivative metabolic process	9	203	329	14312	1.9286	9.150E-01	78.471
GO:0045143~homologous chromosome segregation	2	194	3	13294	45.6838	9.243E-01	52.586
GO:0000075~cell cycle checkpoint	4	192	43	12587	6.0984	9.305E-01	38.692
GO:0051383~kinetochore organization and biogenesis	2	192	2	12587	65.5573	9.421E-01	41.848
GO:0051656~establishment of organelle localization	3	194	27	13294	7.6140	9.434E-01	63.842
GO:0051329~interphase of mitotic cell cycle	4	194	62	13294	4.4210	9.476E-01	65.790
GO:0051325~interphase	4	194	62	13294	4.4210	9.476E-01	65.790
GO:0002520~immune system development	9	194	282	13294	2.1870	9.480E-01	60.823

GO:0045786~negative regulation of progression through cell cycle	6	194	138	13294	2.9794	9.485E-01	58.733
GO:0015031~protein transport	17	194	708	13294	1.6454	9.499E-01	60.192
GO:0051223~regulation of protein transport	3	192	19	12587	10.3512	9.514E-01	44.903
GO:0051310~metaphase plate congression	2	194	4	13294	34.2629	9.526E-01	63.029
GO:0051220~cytoplasmic sequestering of protein	2	194	4	13294	34.2629	9.526E-01	63.029
GO:0042994~cytoplasmic sequestering of transcription factor	2	194	4	13294	34.2629	9.526E-01	63.029
GO:0009069~serine family amino acid metabolic process	3	192	20	12587	9.8336	9.601E-01	48.156
GO:0051258~protein polymerization	4	192	50	12587	5.2446	9.605E-01	51.552
GO:0045143~homologous chromosome segregation	2	192	3	12587	43.7049	9.606E-01	55.656
GO:0006266~DNA ligation	2	192	3	12587	43.7049	9.606E-01	55.656
GO:0006284~base-excision repair	3	192	22	12587	8.9396	9.633E-01	54.432
GO:0046822~regulation of nucleocytoplasmic transport	3	192	21	12587	9.3653	9.635E-01	51.337
GO:0030261~chromosome condensation	3	192	21	12587	9.3653	9.635E-01	51.337
GO:0006520~amino acid metabolic process	9	192	258	12587	2.2869	9.662E-01	54.289
GO:0007127~meiosis I	3	194	31	13294	6.6315	9.674E-01	73.002
GO:0051049~regulation of transport	5	194	110	13294	3.1148	9.680E-01	74.014
GO:0045184~establishment of protein localization	17	194	750	13294	1.5533	9.698E-01	75.408
GO:0006302~double-strand break repair	3	192	24	12587	8.1947	9.721E-01	60.324
GO:0051310~metaphase plate congression	2	192	4	12587	32.7786	9.776E-01	66.186
GO:0051220~cytoplasmic sequestering of protein	2	192	4	12587	32.7786	9.776E-01	66.186
GO:0042994~cytoplasmic sequestering of transcription factor	2	192	4	12587	32.7786	9.776E-01	66.186
GO:0051656~establishment of organelle localization	3	192	27	12587	7.2841	9.785E-01	68.299
GO:0006323~DNA packaging	9	192	281	12587	2.0997	9.792E-01	69.432
GO:0006325~establishment and/or maintenance of chromatin architecture	9	192	274	12587	2.1533	9.796E-01	64.363
GO:0045786~negative regulation of progression through cell cycle	6	192	138	12587	2.8503	9.811E-01	66.014
GO:0015031~protein transport	17	192	708	12587	1.5741	9.813E-01	73.450
GO:0051329~interphase of mitotic cell cycle	4	192	62	12587	4.2295	9.814E-01	71.236
GO:0051325~interphase	4	192	62	12587	4.2295	9.814E-01	71.236
GO:0009263~deoxyribonucleotide biosynthetic process	2	192	5	12587	26.2229	9.815E-01	74.216
GO:0051052~regulation of DNA metabolic process	3	192	29	12587	6.7818	9.819E-01	72.992
GO:0007127~meiosis I	3	192	31	12587	6.3443	9.859E-01	77.175
GO:0006605~protein targeting	8	192	258	12587	2.0328	9.944E-01	84.036

Table S 7 – GSEA Results from the Control vs. Six1 siRNA.
The entire gene expression data set was analyzed by GSEA using the Molecular Signature Database version 3.7, comparing expression values from the Control siRNA treatment with the Six1 siRNA. The number of genes within a gene set as well as the Normalized Enrichment Score (NES) is indicated. The corresponding p-values (Nominal p-value), the False Discovery Rate (FDR) and Familywise Error Rate (FWER) are also provided. Only gene sets with a FDR ≤ 0.05 were retained in this table. **Positive NES** indicate average expression of genes within a gene set is higher in the knock-down than the control treatment. **Negative NES** values indicate average expression of genes within a gene set is lower in the knock-down than the control treatment.

Gene Set	Size	NES	Nominal p-Value	FDR	FWER
KUNINGER_IGF1_VS_PDGFB_TARGETS_UP	41	3.020	0.000	0.000	0.000
REN_ALVEOLAR_RHABDOMYOSARCOMA_UP	83	2.868	0.000	0.000	0.000
GNF2_MYL2	30	2.786	0.000	0.000	0.000
RICKMAN_HEAD_AND_NECK_CANCER_F	45	2.769	0.000	0.000	0.000
GNF2_MYL3	29	2.730	0.000	0.000	0.000
EBAUER_MYOGENIC_TARGETS_OF_PAX3_FOXO1_FUSION	46	2.707	0.000	0.000	0.000
REACTOME_STRIATED_MUSCLE_CONTRACTION	29	2.659	0.000	0.000	0.000
GNF2_TTN	24	2.657	0.000	0.000	0.000
REACTOME_MUSCLE_CONTRACTION	46	2.492	0.000	0.000	0.000
MODULE_329	47	2.477	0.000	0.000	0.000
MODULE_201	45	2.465	0.000	0.000	0.000
EBAUER_TARGETS_OF_PAX3_FOXO1_FUSION_UP	156	2.414	0.000	0.000	0.000
STRUCTURAL_CONSTITUENT_OF_MUSCLE	24	2.398	0.000	0.000	0.000
MODULE_387	45	2.392	0.000	0.000	0.000
GNF2_PA2G4	67	2.359	0.000	0.000	0.000
MODULE_512	35	2.356	0.000	0.000	0.000
MODULE_202	26	2.345	0.000	0.000	0.000
DAIRKEE_CANCER_PRONE_RESPONSE_BPA_E2	104	2.340	0.000	0.000	0.000
GNF2_RAN	70	2.334	0.000	0.000	0.000
MODULE_330	24	2.333	0.000	0.000	0.000
DAIRKEE_CANCER_PRONE_RESPONSE_E2	22	2.315	0.000	0.000	0.000
CONTRACTILE_FIBER_PART	19	2.306	0.000	0.000	0.000
GNF2_NS	34	2.294	0.000	0.000	0.000
KEGG_CARDIAC_MUSCLE_CONTRACTION	65	2.288	0.000	0.000	0.000
GNF2_MSH6	28	2.288	0.000	0.000	0.000
MOOTHA_MITOCHONDRIA	379	2.266	0.000	0.000	0.001
MITOCHONDRIAL_PART	127	2.262	0.000	0.000	0.001
MITOCHONDRIAL_ENVELOPE	85	2.254	0.000	0.000	0.001
CONTRACTILE_FIBER	20	2.250	0.000	0.000	0.001
MOOTHA_VOXPHOS	74	2.240	0.000	0.000	0.001
MOOTHA_HUMAN_MITODB_6_2002	376	2.228	0.000	0.000	0.003
REACTOME_CYTOSOLIC_TRNA_AMINOACYLATION	17	2.215	0.000	0.000	0.004
CTAWWWATA_V$RSRFC4_Q2	256	2.213	0.000	0.000	0.004
MODULE_62	82	2.207	0.000	0.000	0.004
DAIRKEE_CANCER_PRONE_RESPONSE_BPA	40	2.204	0.000	0.000	0.005
KEGG_CITRATE_CYCLE_TCA_CYCLE	27	2.203	0.000	0.000	0.005
KAAB_HEART_ATRIUM_VS_VENTRICLE_DN	201	2.197	0.000	0.000	0.005
REACTOME_CITRIC_ACID_CYCLE	18	2.192	0.000	0.000	0.005
MODULE_152	113	2.175	0.000	0.000	0.005
KEGG_OXIDATIVE_PHOSPHORYLATION	101	2.155	0.000	0.000	0.009
MYOFIBRIL	16	2.153	0.000	0.000	0.009
REACTOME_ORC1_REMOVAL_FROM_CHROMATIN	60	2.150	0.000	0.000	0.010
MITOCHONDRION	271	2.146	0.000	0.000	0.010
MITOCHONDRIAL_INNER_MEMBRANE	60	2.144	0.000	0.000	0.012
REACTOME_ELECTRON_TRANSPORT_CHAIN	56	2.144	0.000	0.000	0.012
MORF_EIF3S2	196	2.140	0.000	0.000	0.016
REACTOME_G1_S_TRANSITION	93	2.117	0.000	0.001	0.030
MORF_FBL	112	2.112	0.000	0.001	0.034
REACTOME_TRNA_AMINOACYLATION	23	2.111	0.000	0.001	0.035
REACTOME_CYCLIN_E_ASSOCIATED_EVENTS_DURING_G1_S_TRANSITION_	55	2.108	0.000	0.001	0.035
REACTOME_DNA_REPLICATION_PRE_INITIATION	69	2.103	0.000	0.001	0.040
KEGG_DILATED_CARDIOMYOPATHY	85	2.103	0.000	0.001	0.041

ORGANELLE_INNER_MEMBRANE	67	2.093	0.000	0.001	0.047
KEGG_HYPERTROPHIC_CARDIOMYOPATHY_HCM	79	2.088	0.000	0.001	0.053
REACTOME_M_G1_TRANSITION	58	2.088	0.000	0.001	0.053
KEGG_PARKINSONS_DISEASE	99	2.085	0.000	0.001	0.055
REACTOME_SYNTHESIS_OF_DNA	80	2.084	0.000	0.001	0.055
MITOCHONDRIAL_MEMBRANE	76	2.081	0.000	0.001	0.060
MORF_SOD1	228	2.079	0.000	0.001	0.064
KEGG_AMINOACYL_TRNA_BIOSYNTHESIS	23	2.079	0.000	0.001	0.064
KEGG_GLYCINE_SERINE_AND_THREONINE_METABOLISM	27	2.077	0.000	0.001	0.065
MODULE_28	30	2.072	0.000	0.001	0.069
REACTOME_CDT1_ASSOCIATION_WITH_THE_CDC6_ORC_ORIGIN_COMPLEX	49	2.058	0.000	0.001	0.095
REACTOME_CELL_CYCLE_CHECKPOINTS	104	2.054	0.000	0.001	0.104
REACTOME_SCF_SKP2_MEDIATED_DEGRADATION_OF_P27_P21	49	2.053	0.000	0.001	0.105
MORF_DEAF1	47	2.049	0.000	0.002	0.115
WONG_MITOCHONDRIA_GENE_MODULE	192	2.046	0.000	0.002	0.117
MODULE_91	35	2.032	0.000	0.002	0.146
NAKAYAMA_SOFT_TISSUE_TUMORS_PCA2_DN	59	2.028	0.000	0.002	0.160
MORF_PRDX3	74	2.024	0.000	0.002	0.167
MUSCLE_DEVELOPMENT	85	2.022	0.000	0.002	0.177
REACTOME_S_PHASE	93	2.016	0.000	0.003	0.199
MORF_BUB3	236	2.010	0.000	0.003	0.224
ORGANELLE_ENVELOPE	145	2.009	0.000	0.003	0.231
MITOCHONDRIAL_MEMBRANE_PART	46	2.007	0.000	0.003	0.236
ENVELOPE	145	2.000	0.000	0.003	0.261
REACTOME_VIF_MEDIATED_DEGRADATION_OF_APOBEC3G	44	1.996	0.000	0.004	0.279
MOOTHA_GLYCOGEN_METABOLISM	16	1.983	0.002	0.004	0.321
MORF_AATF	160	1.972	0.000	0.005	0.376
18310505-TABLE1	49	1.971	0.000	0.005	0.377
MORF_DAP3	162	1.965	0.000	0.005	0.404
MODULE_307	24	1.965	0.002	0.005	0.405
REACTOME_P53_INDEPENDENT_DNA_DAMAGE_RESPONSE	41	1.963	0.002	0.005	0.413
YAUCH_HEDGEHOG_SIGNALING_PARACRINE_DN	223	1.963	0.000	0.006	0.413
KEGG_VIRAL_MYOCARDITIS	45	1.961	0.000	0.006	0.425
GNF2_XRCC5	57	1.956	0.000	0.006	0.450
REACTOME_METABOLISM_OF_AMINO_ACIDS	150	1.953	0.000	0.006	0.460
MORF_MAP2K2	117	1.947	0.000	0.007	0.491
KEGG_PROTEASOME	40	1.943	0.002	0.007	0.515
ENERGY_DERIVATION_BY_OXIDATION_OF_ORGANIC_COMPOUNDS	35	1.939	0.000	0.008	0.541
MORF_RRM1	80	1.932	0.000	0.008	0.583
REACTOME_SCF_BETA_TRCP_MEDIATED_DEGRADATION_OF_EMI1	45	1.930	0.000	0.008	0.597
REACTOME_CDC20_PHOSPHO_APC_MEDIATED_DEGRADATION_OF_CYCLIN_A	60	1.915	0.000	0.010	0.676
CELLULAR_RESPIRATION	19	1.914	0.002	0.010	0.679
SMID_BREAST_CANCER_LUMINAL_A_UP	71	1.914	0.000	0.010	0.680
REACTOME_STABILIZATION_OF_P53	44	1.912	0.000	0.010	0.689
MITOCHONDRIAL_RESPIRATORY_CHAIN	21	1.906	0.002	0.011	0.720
MORF_GMPS	45	1.904	0.000	0.011	0.727
REACTOME_AUTODEGRADATION_OF_CDH1_BY_CDH1_APC	55	1.894	0.000	0.012	0.780
V$E2A_Q2	156	1.893	0.000	0.012	0.785
REACTOME_ACTIVATION_OF_THE_PRE_REPLICATIVE_COMPLEX	25	1.891	0.002	0.013	0.799
CHR2Q33	49	1.890	0.000	0.012	0.801
REACTOME_REGULATION_OF_APC_ACTIVATORS_BETWEEN_G1_S_AND_EARLY_ANAPHASE	67	1.890	0.000	0.013	0.800
MODULE_440	17	1.885	0.000	0.013	0.812
TARTE_PLASMA_CELL_VS_PLASMABLAST_DN	256	1.876	0.000	0.014	0.838
V$MEF2_02	156	1.876	0.000	0.014	0.838
V$MYOD_Q6	158	1.875	0.000	0.014	0.839
MORF_CSNK2B	237	1.873	0.000	0.015	0.849
REACTOME_SIGNALING_BY_WNT	54	1.872	0.000	0.015	0.851
MORF_RAD23A	286	1.862	0.000	0.017	0.891
MORF_HDAC2	226	1.858	0.000	0.017	0.903
MORF_PCNA	69	1.857	0.000	0.018	0.907
16597596-TABLE S1.4	36	1.849	0.000	0.019	0.929
BASAL_LAMINA	17	1.848	0.002	0.019	0.932
MODULE_355	24	1.847	0.000	0.020	0.935
V$E12_Q6	174	1.843	0.000	0.020	0.937
KEGG_ALZHEIMERS_DISEASE	139	1.843	0.000	0.020	0.937
REACTOME_REGULATION_OF_ORNITHINE_DECARBOXYLASE	44	1.840	0.000	0.020	0.942
MORF_G22P1	134	1.831	0.000	0.023	0.967
SHARMA_PILOCYTIC_ASTROCYTOMA_LOCATION_UP	16	1.830	0.006	0.023	0.968
MANALO_HYPOXIA_DN	200	1.827	0.000	0.023	0.971

V$RSRFC4_01	168	1.826	0.000	0.024	0.972
MORF_FEN1	56	1.823	0.000	0.024	0.977
V$RSRFC4_Q2	152	1.821	0.000	0.025	0.979
MORF_ERH	97	1.820	0.000	0.025	0.980
VOLTAGE_GATED_CALCIUM_CHANNEL_COMPLEX	15	1.814	0.004	0.027	0.983
ROME_INSULIN_TARGETS_IN_MUSCLE_UP	58	1.813	0.000	0.027	0.983
MORF_UNG	61	1.811	0.000	0.027	0.985
WEST_ADRENOCORTICAL_TUMOR_MARKERS_UP	18	1.811	0.004	0.027	0.986
MORF_GSPT1	36	1.808	0.000	0.028	0.988
MODULE_77	24	1.800	0.000	0.030	0.992
V$MEF2_Q6_01	168	1.800	0.000	0.030	0.992
16280042-GENELIST	37	1.798	0.002	0.031	0.993
BERTUCCI_INVASIVE_CARCINOMA_DUCTAL_VS_LOBULAR_DN	37	1.791	0.000	0.033	0.995
AEROBIC_RESPIRATION	15	1.790	0.013	0.033	0.995
SCHLOSSER_MYC_TARGETS_REPRESSED_BY_SERUM	106	1.789	0.000	0.033	0.996
MORF_DNMT1	94	1.783	0.000	0.035	0.998
MODULE_299	29	1.782	0.000	0.035	0.998
GCM_PPP1CC	45	1.782	0.002	0.036	0.998
GNF2_NPM1	49	1.780	0.004	0.036	0.998
CALCIUM_CHANNEL_ACTIVITY	31	1.778	0.006	0.036	0.999
STRUCTURAL_MOLECULE_ACTIVITY	190	1.777	0.000	0.037	0.999
BASOLATERAL_PLASMA_MEMBRANE	26	1.773	0.006	0.038	0.999
TSENG_ADIPOGENIC_POTENTIAL_DN	40	1.771	0.006	0.038	0.999
CARBOXYLIC_ACID_METABOLIC_PROCESS	142	1.770	0.000	0.038	0.999
IZADPANAH_STEM_CELL_ADIPOSE_VS_BONE_DN	81	1.770	0.000	0.039	0.999
VOLTAGE_GATED_CALCIUM_CHANNEL_ACTIVITY	18	1.769	0.006	0.038	0.999
REACTOME_SNRNP_ASSEMBLY	40	1.769	0.000	0.038	0.999
MODULE_42	23	1.769	0.006	0.039	0.999
REACTOME_GLUCOSE_REGULATION_OF_INSULIN_SECRETION	131	1.768	0.000	0.038	0.999
ORGANIC_ACID_METABOLIC_PROCESS	144	1.768	0.000	0.038	0.999
MORF_BUB1	48	1.765	0.000	0.039	0.999
REACTOME_PYRUVATE_METABOLISM_AND_TCA_CYCLE	31	1.762	0.004	0.040	0.999
YAO_TEMPORAL_RESPONSE_TO_PROGESTERONE_CLUSTER_11	76	1.757	0.002	0.042	0.999
KEGG_ARRHYTHMOGENIC_RIGHT_VENTRICULAR_CARDIOMYOPATHY_ARVC	67	1.756	0.000	0.042	0.999
PENG_GLUTAMINE_DEPRIVATION_DN	60	1.755	0.004	0.043	0.999
18398820-TABLE1	30	1.753	0.006	0.043	1.000
YAO_TEMPORAL_RESPONSE_TO_PROGESTERONE_CLUSTER_17	146	1.753	0.000	0.043	1.000
TRANSFERASE_ACTIVITY__TRANSFERRING_PENTOSYL_GROUPS	18	1.753	0.010	0.043	1.000
KEGG_HUNTINGTONS_DISEASE	147	1.752	0.000	0.043	1.000
MORF_HAT1	141	1.752	0.000	0.043	1.000
ADHERENS_JUNCTION	16	1.747	0.004	0.044	1.000
17205517-TOP100GOODPROGNOSISGENES	58	1.746	0.004	0.044	1.000
MOOTHA_PGC	298	1.746	0.000	0.044	1.000
SYSTEM_PROCESS	475	1.744	0.000	0.045	1.000
ORGANELLAR_RIBOSOME	20	1.742	0.014	0.046	1.000
V$HMEF2_Q6	92	1.741	0.000	0.046	1.000
LEE_LIVER_CANCER_CIPROFIBRATE_DN	58	1.741	0.000	0.046	1.000
REGULATION_OF_MUSCLE_CONTRACTION	17	1.740	0.004	0.046	1.000
KRIGE_AMINO_ACID_DEPRIVATION	23	1.738	0.006	0.047	1.000
MORF_RAN	226	1.737	0.000	0.047	1.000
MITOCHONDRIAL_RIBOSOME	20	1.733	0.008	0.048	1.000
17053208-TABLE1B	17	1.733	0.012	0.049	1.000
GENERATION_OF_PRECURSOR_METABOLITES_AND_ENERGY	101	1.732	0.002	0.049	1.000
MORF_PRKDC	158	1.731	0.000	0.049	1.000
PROTEASOME_COMPLEX	21	1.728	0.006	0.050	1.000
STOSSI_RESPONSE_TO_ESTRADIOL	30	1.727	0.004	0.050	1.000
GARGALOVIC_RESPONSE_TO_OXIDIZED_PHOSPHOLIPIDS_MAGENTA_UP	19	-1.807	0.000	0.049	0.968
NAKAYAMA_SOFT_TISSUE_TUMORS_PCA2_UP	71	-1.809	0.000	0.049	0.966
14517214-TABLE2	20	-1.810	0.006	0.049	0.964
GTPASE_ACTIVATOR_ACTIVITY	46	-1.810	0.000	0.050	0.964
AAAGGGA,MIR-204,MIR-211	161	-1.813	0.000	0.049	0.961
V$NFKB_C	193	-1.813	0.000	0.049	0.960
WU_HBX_TARGETS_1_DN	21	-1.815	0.008	0.049	0.959
KYNG_DNA_DAMAGE_UP	89	-1.815	0.000	0.049	0.958
CHARAFE_BREAST_CANCER_LUMINAL_VS_BASAL_DN	322	-1.822	0.000	0.046	0.949
MORI_EMU_MYC_LYMPHOMA_BY_ONSET_TIME_DN	17	-1.824	0.002	0.046	0.944
HOFFMANN_PRE_BI_TO_LARGE_PRE_BII_LYMPHOCYTE_DN	52	-1.827	0.000	0.044	0.934
RODRIGUES_THYROID_CARCINOMA_DN	50	-1.830	0.000	0.044	0.928
LINDSTEDT_DENDRITIC_CELL_MATURATION_C	49	-1.832	0.000	0.043	0.921
AMIT_EGF_RESPONSE_60_HELA	35	-1.833	0.000	0.043	0.921
FRASOR_RESPONSE_TO_ESTRADIOL_DN	52	-1.834	0.002	0.043	0.919

MCLACHLAN_DENTAL_CARIES_UP	151	-1.834	0.000	0.044	0.915
GTGTGAG,MIR-342	49	-1.835	0.000	0.044	0.911
TUBE_MORPHOGENESIS	15	-1.837	0.002	0.043	0.906
15817677-TABLES1B	33	-1.839	0.002	0.043	0.905
MODULE_408	18	-1.841	0.002	0.042	0.897
18593951-TABLE1	36	-1.847	0.002	0.041	0.886
WYAAANNRNNNGCG_UNKNOWN	36	-1.849	0.002	0.040	0.878
MODULE_79	82	-1.853	0.000	0.039	0.864
HOMEOSTASIS_OF_NUMBER_OF_CELLS	17	-1.855	0.004	0.039	0.861
17785550-SUPPTABLES2	49	-1.856	0.000	0.039	0.858
GNF2_PTPN6	33	-1.856	0.000	0.039	0.857
DORN_ADENOVIRUS_INFECTION_48HR_DN	31	-1.862	0.008	0.037	0.833
FAELT_B_CLL_WITH_VH3_21_UP	35	-1.863	0.002	0.037	0.827
BILD_HRAS_ONCOGENIC_SIGNATURE	177	-1.869	0.000	0.035	0.807
DE_YY1_TARGETS_DN	81	-1.870	0.000	0.034	0.798
15240595-TABLE2A	40	-1.870	0.000	0.035	0.798
KEGG_AXON_GUIDANCE	119	-1.871	0.002	0.035	0.796
HUTTMANN_B_CLL_POOR_SURVIVAL_DN	48	-1.872	0.000	0.036	0.795
HUPER_BREAST_BASAL_VS_LUMINAL_DN	41	-1.872	0.000	0.036	0.794
NEGATIVE_REGULATION_OF_CELL_DIFFERENTIATION	23	-1.873	0.002	0.037	0.790
RHEIN_ALL_GLUCOCORTICOID_THERAPY_UP	53	-1.873	0.000	0.037	0.788
LANDIS_BREAST_CANCER_PROGRESSION_UP	40	-1.874	0.000	0.037	0.784
ZHANG_TARGETS_OF_EWSR1_FLI1_FUSION	65	-1.876	0.000	0.037	0.775
GGCACAT,MIR-455	39	-1.888	0.000	0.032	0.716
DORN_ADENOVIRUS_INFECTION_32HR_DN	31	-1.893	0.000	0.031	0.689
PARK_TRETINOIN_RESPONSE_AND_PML_RARA_FUSION	22	-1.895	0.000	0.031	0.682
MILI_PSEUDOPODIA_CHEMOTAXIS_DN	365	-1.897	0.000	0.031	0.662
MODULE_170	82	-1.897	0.000	0.031	0.662
14755241-TABLE2B	24	-1.898	0.002	0.032	0.657
MODULE_128	80	-1.904	0.000	0.030	0.632
16273092-TABLES1C	178	-1.906	0.000	0.030	0.624
KINSEY_TARGETS_OF_EWSR1_FLII_FUSION_DN	206	-1.909	0.000	0.030	0.609
MYELOID_CELL_DIFFERENTIATION	30	-1.909	0.000	0.030	0.605
TGCCTTA,MIR-124A	381	-1.914	0.000	0.029	0.577
TONKS_TARGETS_OF_RUNX1_RUNX1T1_FUSION_SUSTAINED_IN_MONOCYTE_UP	19	-1.918	0.000	0.029	0.562
AMIT_SERUM_RESPONSE_60_MCF10A	46	-1.927	0.000	0.026	0.508
FOSTER_INFLAMMATORY_RESPONSE_LPS_DN	311	-1.935	0.000	0.024	0.470
DAZARD_RESPONSE_TO_UV_NHEK_DN	181	-1.936	0.000	0.024	0.464
MODULE_493	44	-1.940	0.000	0.023	0.441
GGGCATT,MIR-365	88	-1.945	0.000	0.023	0.420
ZHANG_ANTIVIRAL_RESPONSE_TO_RIBAVIRIN_DN	37	-1.954	0.000	0.020	0.378
LUI_THYROID_CANCER_PAX8_PPARG_UP	31	-1.955	0.000	0.020	0.368
DELYS_THYROID_CANCER_UP	361	-1.970	0.000	0.016	0.295
KEGG_BASAL_CELL_CARCINOMA	52	-1.980	0.000	0.014	0.253
ACATTCC,MIR-1,MIR-206	212	-1.982	0.000	0.014	0.248
CHR12Q15	18	-1.987	0.000	0.013	0.225
KOBAYASHI_EGFR_SIGNALING_6HR_DN	16	-1.989	0.000	0.013	0.219
TTCCGTT,MIR-191	24	-2.008	0.000	0.010	0.153
NAGASHIMA_NRG1_SIGNALING_UP	142	-2.011	0.000	0.010	0.148
GOLDRATH_IMMUNE_MEMORY	60	-2.016	0.000	0.009	0.136
17053208-TABLE1A	18	-2.018	0.000	0.010	0.134
CHARAFE_BREAST_CANCER_LUMINAL_VS_MESENCHYMAL_DN	321	-2.019	0.000	0.010	0.130
HYDROLASE_ACTIVITY__HYDROLYZING_O_GLYCOSYL_COMPOUNDS	31	-2.025	0.000	0.009	0.118
REACTOME_CLASS_B2_SECRETIN_FAMILY_RECEPTORS	80	-2.034	0.000	0.008	0.101
HEMOPOIESIS	65	-2.040	0.000	0.008	0.093
BIOCARTA_RHO_PATHWAY	31	-2.050	0.000	0.007	0.070
BEGUM_TARGETS_OF_PAX3_FOXO1_FUSION_DN	34	-2.055	0.000	0.007	0.063
IMMUNE_SYSTEM_DEVELOPMENT	71	-2.066	0.000	0.006	0.049
VERRECCHIA_DELAYED_RESPONSE_TO_TGFB1	32	-2.078	0.000	0.005	0.037
HEMOPOIETIC_OR_LYMPHOID_ORGAN_DEVELOPMENT	67	-2.083	0.000	0.004	0.032
ROSS_LEUKEMIA_WITH_MLL_FUSIONS	62	-2.100	0.000	0.004	0.028
BONE_REMODELING	25	-2.102	0.000	0.005	0.027
NAGASHIMA_EGF_SIGNALING_UP	50	-2.105	0.000	0.005	0.027
TISSUE_REMODELING	26	-2.105	0.000	0.006	0.027
KIM_WT1_TARGETS_UP	168	-2.145	0.000	0.004	0.013
YAGI_AML_RELAPSE_PROGNOSIS	27	-2.148	0.000	0.005	0.012
KUNINGER_IGF1_VS_PDGFB_TARGETS_DN	17	-2.391	0.000	0.000	0.000
REN_ALVEOLAR_RHABDOMYOSARCOMA_DN	347	-2.643	0.000	0.000	0.000

Table S 8 – GSEA Results from the Control vs. Six4 siRNA.
The entire gene expression data set was analyzed by GSEA using the Molecular Signature Database version 3.7, comparing expression values from the Control siRNA treatment with the Six4 siRNA. The number of genes within a gene set as well as the Normalized Enrichment Score (NES) is indicated. The corresponding p-values (Nominal p-value), the False Discovery Rate (FDR) and Familywise Error Rate (FWER) are also provided. Only gene sets with a FDR ≤ 0.05 were retained in this table. **Positive NES** indicate average expression of genes within a gene set is higher in the knock-down than the control treatment. **Negative NES** values indicate average expression of genes within a gene set is lower in the knock-down than the control treatment.

Gene Set	Size	NES	Nominal p-Value	FDR	FWER
KUNINGER_IGF1_VS_PDGFB_TARGETS_UP	41	3.200	0.000	0.000	0.000
GNF2_MYL2	30	2.750	0.000	0.000	0.000
REN_ALVEOLAR_RHABDOMYOSARCOMA_UP	83	2.661	0.000	0.000	0.000
GNF2_MYL3	29	2.547	0.000	0.000	0.000
EBAUER_MYOGENIC_TARGETS_OF_PAX3_FOXO1_FUSION	46	2.537	0.000	0.000	0.000
MODULE_329	47	2.524	0.000	0.000	0.000
REACTOME_STRIATED_MUSCLE_CONTRACTION	29	2.519	0.000	0.000	0.000
GNF2_TTN	24	2.509	0.000	0.000	0.001
RICKMAN_HEAD_AND_NECK_CANCER_F	45	2.508	0.000	0.000	0.001
MODULE_387	45	2.441	0.000	0.000	0.001
MODULE_512	35	2.409	0.000	0.000	0.001
MODULE_201	45	2.389	0.000	0.000	0.001
REACTOME_MUSCLE_CONTRACTION	46	2.353	0.000	0.000	0.003
DAIRKEE_CANCER_PRONE_RESPONSE_BPA	40	2.334	0.000	0.000	0.003
STRUCTURAL_CONSTITUENT_OF_MUSCLE	24	2.321	0.000	0.000	0.005
DAIRKEE_CANCER_PRONE_RESPONSE_E2	22	2.285	0.000	0.000	0.005
MITOCHONDRIAL_ENVELOPE	85	2.261	0.000	0.000	0.006
DAIRKEE_CANCER_PRONE_RESPONSE_BPA_E2	104	2.212	0.000	0.001	0.012
MITOCHONDRIAL_INNER_MEMBRANE	60	2.143	0.000	0.002	0.037
KEGG_CARDIAC_MUSCLE_CONTRACTION	65	2.128	0.000	0.003	0.057
MITOCHONDRIAL_MEMBRANE_PART	46	2.124	0.000	0.003	0.061
MITOCHONDRIAL_MEMBRANE	76	2.120	0.000	0.003	0.064
CONTRACTILE_FIBER	20	2.085	0.000	0.005	0.117
SYNAPSE_ORGANIZATION_AND_BIOGENESIS	21	2.053	0.000	0.008	0.192
EBAUER_TARGETS_OF_PAX3_FOXO1_FUSION_UP	156	2.046	0.000	0.008	0.203
MODULE_202	26	2.035	0.000	0.010	0.243
MODULE_152	113	2.030	0.000	0.010	0.257
KEGG_CITRATE_CYCLE_TCA_CYCLE	27	2.021	0.000	0.011	0.296
ORGANELLE_INNER_MEMBRANE	67	2.003	0.000	0.014	0.370
KEGG_AMINOACYL_TRNA_BIOSYNTHESIS	23	2.001	0.000	0.015	0.384
MITOCHONDRIAL_PART	127	1.988	0.000	0.017	0.453
REACTOME_CITRIC_ACID_CYCLE	18	1.984	0.002	0.018	0.470
MODULE_233	19	1.970	0.002	0.020	0.530
CTAWWWATA_V$RSRFC4_Q2	256	1.969	0.000	0.020	0.535
VOLTAGE_GATED_CALCIUM_CHANNEL_ACTIVITY	18	1.967	0.002	0.020	0.548
CONTRACTILE_FIBER_PART	19	1.966	0.000	0.020	0.551
SKELETAL_MUSCLE_DEVELOPMENT	29	1.953	0.000	0.023	0.615
REACTOME_CYTOSOLIC_TRNA_AMINOACYLATION	17	1.953	0.000	0.023	0.617
MOOTHA_MITOCHONDRIA	379	1.952	0.000	0.023	0.622
REACTOME_NCAM1_INTERACTIONS	40	1.943	0.000	0.024	0.658
17999412-GENELIST	15	1.939	0.000	0.025	0.678
REACTOME_SIGNALING_BY_WNT	54	1.929	0.000	0.027	0.731
PENG_RAPAMYCIN_RESPONSE_DN	52	1.924	0.000	0.028	0.752
ROME_INSULIN_TARGETS_IN_MUSCLE_UP	58	1.922	0.000	0.028	0.755
V$E12_Q6	174	1.921	0.000	0.028	0.761
REACTOME_TRANSLATION	98	1.915	0.000	0.029	0.789
RORIE_TARGETS_OF_EWSR1_FLI1_FUSION_DN	19	1.913	0.000	0.029	0.800
KAAB_HEART_ATRIUM_VS_VENTRICLE_DN	201	1.911	0.000	0.029	0.803
VOLTAGE_GATED_CALCIUM_CHANNEL_COMPLEX	15	1.904	0.004	0.031	0.822
REACTOME_VIF_MEDIATED_DEGRADATION_OF_APOBEC3G	44	1.885	0.000	0.037	0.875
TRANSLATION_FACTOR_ACTIVITY__NUCLEIC_ACID_BINDING	27	1.885	0.002	0.038	0.875
MOOTHA_FFA_OXYDATION	18	1.881	0.004	0.038	0.889

REACTOME_TRNA_AMINOACYLATION	23	1.880	0.000	0.037	0.892
EXTRACELLULAR_STRUCTURE_ORGANIZATION_AND_BIOGENESIS	29	1.877	0.000	0.038	0.899
REACTOME_PYRUVATE_METABOLISM_AND_TCA_CYCLE	31	1.874	0.000	0.038	0.906
REACTOME_P53_INDEPENDENT_DNA_DAMAGE_RESPONSE	41	1.874	0.000	0.038	0.906
MUSCLE_DEVELOPMENT	85	1.868	0.000	0.039	0.915
CALCIUM_CHANNEL_ACTIVITY	31	1.868	0.005	0.039	0.915
STRUCTURAL_MOLECULE_ACTIVITY	190	1.865	0.000	0.039	0.924
KEGG_VIBRIO_CHOLERAE_INFECTION	47	1.864	0.000	0.039	0.924
ATPASE_ACTIVITY__COUPLED_TO_TRANSMEMBRANE_MOVEMENT_OF_IONS	22	1.861	0.004	0.040	0.928
SWEET_KRAS_TARGETS_DN	20	1.850	0.004	0.044	0.944
MOOTHA_HUMAN_MITODB_6_2002	376	1.847	0.000	0.044	0.948
BIOCARTA_MCALPAIN_PATHWAY	19	1.846	0.000	0.044	0.951
CHR17P11	32	1.840	0.002	0.046	0.959
KINASE_INHIBITOR_ACTIVITY	20	1.837	0.011	0.048	0.970
MODULE_28	30	1.834	0.002	0.048	0.973
MITOCHONDRIAL_RESPIRATORY_CHAIN	21	1.832	0.006	0.049	0.975
MODULE_355	24	1.829	0.000	0.050	0.976
ATPASE_ACTIVITY__COUPLED_TO_TRANSMEMBRANE_MOVEMENT_OF_IONS__PHOSPHORYLATIVE_MECHANISM	19	1.828	0.000	0.049	0.979
MODULE_421	19	-1.684	0.015	0.050	1.000
CHIARADONNA_NEOPLASTIC_TRANSFORMATION_KRAS_UP	112	-1.685	0.000	0.050	1.000
RESPONSE_TO_DNA_DAMAGE_STIMULUS	130	-1.685	0.000	0.050	1.000
MODULE_241	62	-1.685	0.004	0.050	1.000
CHARAFE_BREAST_CANCER_LUMINAL_VS_BASAL_DN	322	-1.686	0.000	0.050	1.000
BENPORATH_ES_CORE_NINE_CORRELATED	88	-1.688	0.000	0.049	1.000
REACTOME_CENTROSOME_MATURATION	40	-1.688	0.005	0.049	1.000
12173052-TABLE1	21	-1.692	0.006	0.047	1.000
18483624-TABLE1	27	-1.693	0.007	0.047	1.000
CONCANNON_APOPTOSIS_BY_EPOXOMICIN_DN	127	-1.693	0.000	0.047	1.000
SPINDLE_MICROTUBULE	15	-1.693	0.009	0.048	1.000
KORKOLA_TERATOMA	29	-1.697	0.006	0.046	1.000
UZONYI_RESPONSE_TO_LEUKOTRIENE_AND_THROMBIN	32	-1.701	0.002	0.044	1.000
REACTOME_RNA_POLYMERASE_I_PROMOTER_OPENING	24	-1.702	0.006	0.044	1.000
DORN_ADENOVIRUS_INFECTION_24HR_DN	36	-1.704	0.004	0.043	1.000
BOYLAN_MULTIPLE_MYELOMA_C_CLUSTER_UP	26	-1.705	0.004	0.043	1.000
MICROTUBULE_BASED_PROCESS	62	-1.709	0.000	0.041	1.000
SAMOLS_TARGETS_OF_KHSV_MIRNAS_DN	42	-1.710	0.008	0.041	1.000
PYEON_CANCER_HEAD_AND_NECK_VS_CERVICAL_UP	105	-1.710	0.000	0.041	1.000
16948840-TABLE1	18	-1.714	0.004	0.039	1.000
BIOCARTA_TOLL_PATHWAY	32	-1.716	0.004	0.039	1.000
CENTROSOME	38	-1.717	0.006	0.039	1.000
VERNELL_RETINOBLASTOMA_PATHWAY_UP	38	-1.718	0.002	0.038	1.000
GTTATAT,MIR-410	75	-1.719	0.002	0.038	1.000
MODULE_320	16	-1.719	0.009	0.038	1.000
OUYANG_PROSTATE_CANCER_PROGRESSION_UP	18	-1.719	0.009	0.038	1.000
GARGALOVIC_RESPONSE_TO_OXIDIZED_PHOSPHOLIPIDS_MAGENTA_UP	19	-1.720	0.009	0.038	1.000
WIELAND_UP_BY_HBV_INFECTION	73	-1.725	0.000	0.036	1.000
FOURNIER_ACINAR_DEVELOPMENT_LATE_2	209	-1.725	0.000	0.036	1.000
MODULE_52	356	-1.726	0.000	0.036	1.000
MICROTUBULE_MOTOR_ACTIVITY	15	-1.729	0.009	0.035	1.000
MODULE_451	23	-1.730	0.005	0.035	1.000
19224407-TABLE1	17	-1.730	0.011	0.035	1.000
MODULE_312	41	-1.730	0.000	0.035	1.000
17894856-SUPPLIST2	196	-1.730	0.000	0.035	1.000
MODULE_397	90	-1.730	0.000	0.035	1.000
16195754-TABLE2	57	-1.731	0.000	0.035	1.000
DNA_REPAIR	108	-1.731	0.000	0.035	1.000
17471573-TABLE2	86	-1.736	0.000	0.034	1.000
RUIZ_TNC_TARGETS_UP	115	-1.737	0.002	0.034	1.000
16273092-TABLES1C	178	-1.739	0.000	0.033	1.000
MODULE_222	16	-1.739	0.004	0.033	1.000
REGULATION_OF_MITOSIS	35	-1.740	0.002	0.033	1.000
JI_METASTASIS_REPRESSED_BY_STK11	20	-1.742	0.006	0.032	1.000
TONKS_TARGETS_OF_RUNX1_RUNX1T1_FUSION_SUSTAINDED_IN_ERYTHROCYTE_UP	35	-1.744	0.000	0.032	1.000
SIMBULAN_UV_RESPONSE_NORMAL_DN	23	-1.745	0.004	0.031	1.000
17430594-TABLE4	92	-1.746	0.000	0.031	1.000
CHROMATIN_BINDING	28	-1.746	0.004	0.031	1.000
REN_ALVEOLAR_RHABDOMYOSARCOMA_DN	347	-1.748	0.000	0.031	1.000
MODULE_168	17	-1.749	0.003	0.031	1.000
MODULE_53	320	-1.749	0.000	0.031	1.000

RICKMAN_HEAD_AND_NECK_CANCER_A	67	-1.752	0.000	0.030	1.000
REACTOME_MITOTIC_M_M_G1_PHASES	122	-1.755	0.000	0.029	1.000
MOTOR_ACTIVITY	24	-1.757	0.008	0.029	1.000
ALCALAY_AML_BY_NPM1_LOCALIZATION_DN	152	-1.760	0.000	0.028	1.000
15246160-TABLE4	17	-1.763	0.002	0.027	1.000
MODULE_291	54	-1.763	0.000	0.027	1.000
YAO_TEMPORAL_RESPONSE_TO_PROGESTERONE_CLUSTER_15	32	-1.765	0.000	0.027	1.000
REACTOME_E2F_TRANSCRIPTIONAL_TARGETS_AT_G1_S	19	-1.766	0.000	0.026	1.000
JAEGER_METASTASIS_UP	33	-1.767	0.000	0.026	1.000
MORI_EMU_MYC_LYMPHOMA_BY_ONSET_TIME_DN	17	-1.768	0.007	0.026	1.000
MODULE_438	48	-1.768	0.004	0.026	1.000
MODULE_439	17	-1.775	0.002	0.024	1.000
LEE_LIVER_CANCER_ACOX1_UP	53	-1.776	0.004	0.024	1.000
CHR7Q21	47	-1.776	0.000	0.024	1.000
XU_HGF_SIGNALING_NOT_VIA_AKT1_48HR_DN	15	-1.777	0.009	0.024	1.000
AMIT_SERUM_RESPONSE_40_MCF10A	27	-1.778	0.006	0.024	0.998
CHR4Q13	20	-1.779	0.002	0.024	0.998
CHR12Q15	18	-1.779	0.002	0.024	0.998
ORGANELLE_LOCALIZATION	18	-1.780	0.004	0.024	0.996
KEGG_DNA_REPLICATION	31	-1.782	0.002	0.023	0.996
DORN_ADENOVIRUS_INFECTION_12HR_DN	27	-1.782	0.009	0.023	0.996
MORF_BUB1B	59	-1.785	0.000	0.022	0.996
17760805-SUPPTABLE3A	67	-1.787	0.000	0.022	0.996
GEORGES_CELL_CYCLE_MIR192_TARGETS	51	-1.788	0.002	0.022	0.994
MICROTUBULE_ORGANIZING_CENTER	46	-1.793	0.000	0.021	0.992
REACTOME_PACKAGING_OF_TELOMERE_ENDS	23	-1.795	0.002	0.020	0.992
CHROMOSOME_SEGREGATION	20	-1.798	0.008	0.020	0.991
KEGG_AXON_GUIDANCE	119	-1.799	0.000	0.020	0.990
PUJANA_BREAST_CANCER_WITH_BRCA1_MUTATED_UP	47	-1.799	0.000	0.020	0.990
BROWNE_HCMV_INFECTION_2HR_DN	44	-1.803	0.000	0.019	0.982
18798265-TABLES2	114	-1.804	0.000	0.019	0.982
KAUFFMANN_DNA_REPAIR_GENES	185	-1.805	0.000	0.019	0.981
NEWMAN_ERCC6_TARGETS_UP	16	-1.805	0.000	0.019	0.980
MODULE_124	82	-1.806	0.002	0.019	0.980
MODULE_98	325	-1.806	0.000	0.019	0.980
MICROTUBULE_CYTOSKELETON_ORGANIZATION_AND_BIOGENESIS	26	-1.806	0.005	0.019	0.980
CYTOKINESIS	17	-1.808	0.002	0.018	0.980
18347175-TABLES4.1	15	-1.811	0.007	0.018	0.979
KIM_WT1_TARGETS_DN	325	-1.811	0.000	0.018	0.979
CERVERA_SDHB_TARGETS_2	81	-1.812	0.000	0.018	0.978
VANTVEER_BREAST_CANCER_POOR_PROGNOSIS	40	-1.814	0.004	0.017	0.976
GRAHAM_NORMAL_QUIESCENT_VS_NORMAL_DIVIDING_UP	52	-1.815	0.000	0.017	0.976
BECKER_TAMOXIFEN_RESISTANCE_DN	39	-1.815	0.000	0.017	0.976
DAZARD_UV_RESPONSE_CLUSTER_G6	89	-1.817	0.000	0.017	0.974
MOLENAAR_TARGETS_OF_CCND1_AND_CDK4_UP	30	-1.819	0.000	0.017	0.970
CHROMOSOME	96	-1.823	0.000	0.016	0.965
MODULE_76	70	-1.823	0.000	0.016	0.965
CELL_DIVISION	19	-1.827	0.002	0.016	0.961
BIOCARTA_GSK3_PATHWAY	26	-1.827	0.000	0.016	0.961
KRCTCNNNNMANAGC_UNKNOWN	30	-1.830	0.000	0.015	0.958
AMIT_EGF_RESPONSE_240_MCF10A	16	-1.830	0.000	0.015	0.958
12414654-TABLE2	19	-1.831	0.004	0.015	0.957
NAKAYAMA_FRA2_TARGETS	30	-1.834	0.002	0.015	0.946
MASSARWEH_RESPONSE_TO_ESTRADIOL	45	-1.836	0.007	0.014	0.943
NADERI_BREAST_CANCER_PROGNOSIS_UP	31	-1.836	0.000	0.014	0.943
LINDGREN_BLADDER_CANCER_CLUSTER_1_DN	292	-1.841	0.000	0.014	0.933
15831697-TABLES3	103	-1.846	0.000	0.013	0.918
V$WHN_B	167	-1.847	0.000	0.013	0.914
KEGG_MISMATCH_REPAIR	22	-1.851	0.004	0.012	0.902
DACOSTA_UV_RESPONSE_VIA_ERCC3_TTD_DN	50	-1.851	0.000	0.012	0.901
MCBRYAN_PUBERTAL_BREAST_6_7WK_DN	68	-1.853	0.000	0.012	0.892
CELL_CYCLE_GO_0007049	251	-1.853	0.000	0.012	0.892
CHR6P22	42	-1.867	0.000	0.010	0.835
19112514-TABLES1	31	-1.874	0.002	0.009	0.817
16404635-TABLE2	16	-1.877	0.002	0.009	0.810
SARRIO_EPITHELIAL_MESENCHYMAL_TRANSITION_UP	15	-1.880	0.002	0.009	0.792
SCIBETTA_KDM5B_TARGETS_DN	54	-1.881	0.000	0.009	0.784
REACTOME_G2_M_TRANSITION	52	-1.882	0.002	0.009	0.779
BREDEMEYER_RAG_SIGNALING_VIA_ATM_NOT_VIA_NFKB_UP	34	-1.885	0.000	0.008	0.768
16626501-SUPPL2	42	-1.886	0.000	0.008	0.765
11823860-FIGURE2	42	-1.889	0.000	0.008	0.754

OLSSON_E2F3_TARGETS_DN	22	-1.890	0.000	0.008	0.749
NIKOLSKY_BREAST_CANCER_1Q21_AMPLICON	27	-1.892	0.000	0.008	0.737
MODULE_118	334	-1.892	0.000	0.008	0.737
TOYOTA_TARGETS_OF_MIR34B_AND_MIR34C	264	-1.894	0.000	0.008	0.731
17410534-TABLE1	53	-1.895	0.000	0.007	0.721
PUJANA_XPRSS_INT_NETWORK	133	-1.896	0.000	0.007	0.721
BERENJENO_TRANSFORMED_BY_RHOA_UP	430	-1.898	0.000	0.007	0.717
MODULE_408	18	-1.900	0.000	0.007	0.709
18637760-SUPPTABLE3	27	-1.901	0.000	0.007	0.698
CELL_CYCLE_PHASE	137	-1.902	0.000	0.007	0.695
LI_WILMS_TUMOR_VS_FETAL_KIDNEY_1_DN	133	-1.906	0.000	0.007	0.673
14737219-CSRGENES	409	-1.915	0.000	0.006	0.627
MUELLER_PLURINET	252	-1.916	0.000	0.006	0.620
15843827-TABLES14	28	-1.924	0.002	0.005	0.584
KIM_WT1_TARGETS_UP	168	-1.927	0.000	0.005	0.574
MODULE_403	42	-1.927	0.000	0.005	0.574
SHEPARD_CRUSH_AND_BURN_MUTANT_DN	134	-1.930	0.000	0.005	0.559
POOLA_INVASIVE_BREAST_CANCER_UP	209	-1.933	0.000	0.005	0.542
RHEIN_ALL_GLUCOCORTICOID_THERAPY_UP	53	-1.934	0.000	0.005	0.535
BOYAULT_LIVER_CANCER_SUBCLASS_G23_UP	40	-1.940	0.000	0.004	0.509
16707422-TABLES1B	34	-1.941	0.000	0.004	0.498
FRASOR_RESPONSE_TO_ESTRADIOL_DN	52	-1.942	0.000	0.004	0.493
SHEPARD_BMYB_MORPHOLINO_DN	149	-1.947	0.000	0.004	0.474
SUNG_METASTASIS_STROMA_DN	35	-1.948	0.000	0.004	0.472
11823860-SUPPTABLE2	131	-1.953	0.000	0.004	0.457
HOFFMANN_PRE_BI_TO_LARGE_PRE_BII_LYMPHOCYTE_DN	52	-1.953	0.000	0.004	0.455
RUGO_STRESS_RESPONSE_SUBSET_H	37	-1.961	0.000	0.003	0.416
REACTOME_DNA_STRAND_ELONGATION	25	-1.961	0.000	0.003	0.416
KYNG_DNA_DAMAGE_BY_GAMMA_AND_UV_RADIATION	37	-1.962	0.000	0.003	0.411
MCMURRAY_TP53_HRAS_COOPERATION_RESPONSE_UP	24	-1.975	0.000	0.003	0.337
MODULE_198	254	-1.986	0.000	0.002	0.300
MODULE_346	16	-1.986	0.000	0.002	0.298
MARKEY_RB1_CHRONIC_LOF_UP	102	-1.986	0.000	0.002	0.298
CHROMOSOME__PERICENTRIC_REGION	24	-1.992	0.000	0.002	0.279
PUJANA_BRCA2_PCC_NETWORK	327	-1.993	0.000	0.002	0.269
18347175-TABLES4.2	29	-1.997	0.000	0.002	0.246
17899371-GENETABLE4	295	-1.998	0.000	0.002	0.239
SASAKI_ADULT_T_CELL_LEUKEMIA	117	-2.002	0.000	0.002	0.230
CHROMOSOMAL_PART	74	-2.003	0.000	0.002	0.221
REACTOME_TELOMERE_MAINTENANCE	46	-2.005	0.000	0.002	0.215
18288381-TABLE4	49	-2.008	0.000	0.002	0.204
SONG_TARGETS_OF_IE86_CMV_PROTEIN	43	-2.008	0.000	0.002	0.203
BERENJENO_TRANSFORMED_BY_RHOA_FOREVER_DN	28	-2.019	0.000	0.001	0.172
RODRIGUES_THYROID_CARCINOMA_DN	50	-2.019	0.000	0.001	0.170
MODULE_196	16	-2.022	0.000	0.001	0.159
17873908-TABLE3	15	-2.022	0.000	0.001	0.159
CELL_CYCLE_PROCESS	152	-2.030	0.000	0.001	0.142
REN_BOUND_BY_E2F	47	-2.032	0.000	0.001	0.135
TTCCGTT,MIR-191	24	-2.032	0.000	0.001	0.135
REACTOME_CELL_CYCLE_MITOTIC	231	-2.033	0.000	0.001	0.134
17510386-TABLE4	30	-2.040	0.000	0.001	0.118
15994935-TABLE3B	35	-2.051	0.002	0.001	0.099
MODULE_253	15	-2.052	0.000	0.001	0.097
15318932-TABLE2	36	-2.052	0.000	0.001	0.096
RHODES_UNDIFFERENTIATED_CANCER	57	-2.053	0.000	0.001	0.096
VECCHI_GASTRIC_CANCER_EARLY_UP	286	-2.055	0.000	0.001	0.093
FOURNIER_ACINAR_DEVELOPMENT_LATE_DN	17	-2.059	0.000	0.001	0.085
LY_AGING_PREMATURE_DN	23	-2.062	0.000	0.001	0.082
MITSIADES_RESPONSE_TO_APLIDIN_DN	182	-2.062	0.000	0.001	0.081
FRASOR_RESPONSE_TO_SERM_OR_FULVESTRANT_DN	42	-2.063	0.000	0.001	0.081
18698033-TABLES1-AURKA	170	-2.063	0.000	0.001	0.081
RUIZ_TNC_TARGETS_DN	108	-2.064	0.000	0.001	0.081
MOLENAAR_TARGETS_OF_CCND1_AND_CDK4_DN	32	-2.065	0.000	0.001	0.081
20156340-GENIUS-ERPOSHER2NEG_SIGNATURE	170	-2.068	0.000	0.001	0.075
MODULE_252	198	-2.072	0.000	0.001	0.066
FERREIRA_EWINGS_SARCOMA_UNSTABLE_VS_STABLE_UP	92	-2.075	0.000	0.001	0.062
FINETTI_BREAST_CANCER_BASAL_VS_LUMINAL	15	-2.078	0.000	0.001	0.059
17205517-TOP100POORPROGNOSISGENES	61	-2.088	0.000	0.000	0.049
17072343-TABLE1	33	-2.094	0.000	0.000	0.043
VANTVEER_BREAST_CANCER_METASTASIS_DN	83	-2.100	0.000	0.000	0.035
SENGUPTA_NASOPHARYNGEAL_CARCINOMA_UP	198	-2.104	0.000	0.000	0.033

GARGALOVIC_RESPONSE_TO_OXIDIZED_PHOSPHOLIPIDS_TURQUOISE_DN	33	-2.105	0.000	0.000	0.032
FINETTI_BREAST_CANCER_KINOME_RED	15	-2.106	0.002	0.000	0.029
LINDGREN_BLADDER_CANCER_CLUSTER_3_UP	228	-2.107	0.000	0.000	0.029
MODULE_125	41	-2.108	0.000	0.000	0.029
MODULE_57	51	-2.111	0.000	0.000	0.028
SHEDDEN_LUNG_CANCER_POOR_SURVIVAL_A6	330	-2.117	0.000	0.000	0.026
M_PHASE	91	-2.119	0.000	0.000	0.025
GRAHAM_NORMAL_QUIESCENT_VS_NORMAL_DIVIDING_DN	67	-2.121	0.000	0.000	0.025
GOLDRATH_ANTIGEN_RESPONSE	303	-2.128	0.000	0.000	0.023
MICROTUBULE_CYTOSKELETON	112	-2.131	0.000	0.000	0.022
REACTOME_MITOTIC_PROMETAPHASE	60	-2.132	0.000	0.000	0.022
16849555-22GENES	16	-2.132	0.000	0.000	0.022
17495134-TABLES1	67	-2.133	0.000	0.000	0.022
MODULE_158	40	-2.136	0.000	0.000	0.022
MODULE_197	141	-2.141	0.000	0.000	0.020
PYEON_HPV_POSITIVE_TUMORS_UP	56	-2.168	0.000	0.000	0.014
SHEPARD_BMYB_TARGETS	58	-2.169	0.000	0.000	0.014
18416598-SUPPTABLE3	37	-2.171	0.000	0.000	0.013
MITOTIC_CELL_CYCLE	121	-2.176	0.000	0.000	0.013
SPINDLE	33	-2.178	0.000	0.000	0.013
MORF_CCNF	51	-2.185	0.000	0.000	0.008
17510386-TABLE3	63	-2.187	0.000	0.000	0.008
FUJII_YBX1_TARGETS_DN	114	-2.190	0.000	0.000	0.008
18338247-SUPPTABLE4A	140	-2.192	0.000	0.000	0.008
17899371-25GENELIST	21	-2.192	0.000	0.000	0.008
WILCOX_PRESPONSE_TO_ROGESTERONE_UP	99	-2.192	0.000	0.000	0.008
18398820-TABLE5	35	-2.193	0.000	0.000	0.008
PUJANA_BREAST_CANCER_LIT_INT_NETWORK	90	-2.196	0.000	0.000	0.008
GNF2_RFC3	33	-2.210	0.000	0.000	0.005
CHEMNITZ_RESPONSE_TO_PROSTAGLANDIN_E2_UP	93	-2.220	0.000	0.000	0.005
MODULE_54	196	-2.222	0.000	0.000	0.004
WU_APOPTOSIS_BY_CDKN1A_VIA_TP53	29	-2.233	0.000	0.000	0.004
17076897-TABLES3	37	-2.237	0.000	0.000	0.004
KOBAYASHI_EGFR_SIGNALING_6HR_DN	16	-2.238	0.000	0.000	0.004
MARKEY_RB1_ACUTE_LOF_DN	191	-2.245	0.000	0.000	0.004
MORI_IMMATURE_B_LYMPHOCYTE_DN	44	-2.251	0.000	0.000	0.003
MISSIAGLIA_REGULATED_BY_METHYLATION_DN	88	-2.252	0.000	0.000	0.002
HORIUCHI_WTAP_TARGETS_DN	229	-2.261	0.000	0.000	0.001
GNF2_CKS1B	34	-2.264	0.000	0.000	0.001
16919171-TABLE3	38	-2.267	0.000	0.000	0.001
YU_MYC_TARGETS_UP	38	-2.267	0.000	0.000	0.001
KUNINGER_IGF1_VS_PDGFB_TARGETS_DN	17	-2.274	0.000	0.000	0.001
18992152-TABLE3	28	-2.277	0.000	0.000	0.001
18231641-W2GENELIST	48	-2.277	0.000	0.000	0.001
FURUKAWA_DUSP6_TARGETS_PCI35_DN	46	-2.279	0.000	0.000	0.001
BASAKI_YBX1_TARGETS_UP	205	-2.280	0.000	0.000	0.001
LY_AGING_MIDDLE_DN	15	-2.281	0.000	0.000	0.001
BLUM_RESPONSE_TO_SALIRASIB_DN	298	-2.282	0.000	0.000	0.001
BENPORATH_CYCLING_GENES	445	-2.289	0.000	0.000	0.001
M_PHASE_OF_MITOTIC_CELL_CYCLE	67	-2.298	0.000	0.000	0.001
LY_AGING_OLD_DN	43	-2.310	0.000	0.000	0.001
EGUCHI_CELL_CYCLE_RB1_TARGETS	19	-2.312	0.000	0.000	0.001
GNF2_BUB1B	41	-2.345	0.000	0.000	0.000
MORI_LARGE_PRE_BII_LYMPHOCYTE_UP	48	-2.348	0.000	0.000	0.000
MITOSIS	65	-2.353	0.000	0.000	0.000
WINNEPENNINCKX_MELANOMA_METASTASIS_UP	104	-2.358	0.000	0.000	0.000
GNF2_FEN1	42	-2.376	0.000	0.000	0.000
BENPORATH_PROLIFERATION	109	-2.393	0.000	0.000	0.000
15171711-TABLE6	37	-2.409	0.000	0.000	0.000
CHIANG_LIVER_CANCER_SUBCLASS_PROLIFERATION_UP	121	-2.434	0.000	0.000	0.000
CROONQUIST_NRAS_VS_STROMAL_STIMULATION_DN	68	-2.436	0.000	0.000	0.000
GRAHAM_CML_DIVIDING_VS_NORMAL_QUIESCENT_UP	139	-2.436	0.000	0.000	0.000
ODONNELL_TARGETS_OF_MYC_AND_TFRC_DN	28	-2.440	0.000	0.000	0.000
18992152-TABLE4	30	-2.452	0.000	0.000	0.000
CHANG_CYCLING_GENES	30	-2.456	0.000	0.000	0.000
SCIAN_CELL_CYCLE_TARGETS_OF_TP53_AND_TP73_DN	22	-2.470	0.000	0.000	0.000
LE_EGR2_TARGETS_UP	92	-2.472	0.000	0.000	0.000
17150101-TABLES1A	49	-2.481	0.000	0.000	0.000
GNF2_SMC4L1	66	-2.481	0.000	0.000	0.000
17150101-TABLES1L	218	-2.490	0.000	0.000	0.000
TANG_SENESCENCE_TP53_TARGETS_DN	36	-2.492	0.000	0.000	0.000

WHITEFORD_PEDIATRIC_CANCER_MARKERS	87	-2.521	0.000	0.000	0.000
KAUFFMANN_MELANOMA_RELAPSE_UP	54	-2.525	0.000	0.000	0.000
GNF2_TTK	30	-2.525	0.000	0.000	0.000
GNF2_RRM1	70	-2.526	0.000	0.000	0.000
16921376-GENELIST	55	-2.529	0.000	0.000	0.000
HOFFMANN_LARGE_TO_SMALL_PRE_BII_LYMPHOCYTE_UP	78	-2.535	0.000	0.000	0.000
GNF2_H2AFX	27	-2.541	0.000	0.000	0.000
GNF2_MKI67	23	-2.543	0.000	0.000	0.000
GNF2_MCM4	45	-2.547	0.000	0.000	0.000
GNF2_RFC4	50	-2.562	0.000	0.000	0.000
18662380-S3-AURKA	254	-2.567	0.000	0.000	0.000
FARMER_BREAST_CANCER_CLUSTER_2	25	-2.574	0.000	0.000	0.000
GNF2_SMC2L1	27	-2.588	0.000	0.000	0.000
ZHAN_MULTIPLE_MYELOMA_PR_UP	26	-2.601	0.000	0.000	0.000
17150101-TABLES1H	199	-2.610	0.000	0.000	0.000
LEE_EARLY_T_LYMPHOCYTE_UP	56	-2.620	0.000	0.000	0.000
17150101-TABLES1C	79	-2.625	0.000	0.000	0.000
CROONQUIST_NRAS_SIGNALING_DN	53	-2.629	0.000	0.000	0.000
GNF2_RRM2	34	-2.634	0.000	0.000	0.000
KOBAYASHI_EGFR_SIGNALING_24HR_DN	194	-2.653	0.000	0.000	0.000
18498629-GENELIST	136	-2.674	0.000	0.000	0.000
18271932-GENELIST	23	-2.682	0.000	0.000	0.000
16478745-SUPPTABLE1	145	-2.688	0.000	0.000	0.000
18427120-DATAS4	307	-2.692	0.000	0.000	0.000
GNF2_BUB1	23	-2.695	0.000	0.000	0.000
GNF2_CENPE	35	-2.705	0.000	0.000	0.000
AMUNDSON_GAMMA_RADIATION_RESPONSE	31	-2.715	0.000	0.000	0.000
17804718-HUMANGENELIST	53	-2.717	0.000	0.000	0.000
NAKAYAMA_SOFT_TISSUE_TUMORS_PCA2_UP	71	-2.727	0.000	0.000	0.000
GNF2_ESPL1	32	-2.734	0.000	0.000	0.000
16478745-SUPPTABLE1-SHORT	80	-2.746	0.000	0.000	0.000
CROONQUIST_IL6_DEPRIVATION_DN	69	-2.757	0.000	0.000	0.000
GNF2_PCNA	55	-2.770	0.000	0.000	0.000
SOTIRIOU_BREAST_CANCER_GRADE_1_VS_3_UP	113	-2.807	0.000	0.000	0.000
18427120-DATAS5	122	-2.810	0.000	0.000	0.000
KANG_DOXORUBICIN_RESISTANCE_UP	43	-2.825	0.000	0.000	0.000
GNF2_CKS2	39	-2.829	0.000	0.000	0.000
ODONNELL_TFRC_TARGETS_DN	83	-2.849	0.000	0.000	0.000
GNF2_CDC2	49	-2.872	0.000	0.000	0.000
GNF2_HMMR	39	-2.881	0.000	0.000	0.000
ROSTY_CERVICAL_CANCER_PROLIFERATION_CLUSTER	109	-2.938	0.000	0.000	0.000
GNF2_CCNA2	54	-2.947	0.000	0.000	0.000
GNF2_CENPF	49	-2.960	0.000	0.000	0.000
GNF2_CCNB2	49	-2.967	0.000	0.000	0.000
GNF2_CDC20	48	-3.033	0.000	0.000	0.000

Table S 9 – Expression of muscle fiber type genes in the Six1 and/or Six4 knock-downs. Genes involved in Fast and Slow Type muscle fibers are listed in the following table. Agilent Probe ID, GeneSymbol, Gene ID, Genebank accession number and the fold changes of each individual knock-down (siSix1, siSix4, siSix1&4 and siMyog) against the control siRNA are provided. Fold change values are in log 2 base. ND – Not Detected.

	ProbeID	Gene Symbol	Gene ID	Genbank	siSix1 vs. Control	siSix4 vs. Control	siSix1&4 vs. Control	siMyog vs. Control
Fast Type	A_51_P244856	Atp2a1	11937	NM_007504	-2.064434	-1.9174475	-2.233686	-1.494857
	A_52_P686130	Atp2a1	11937	NM_007504	-2.825479	-2.5515095	-3.574885	-1.801323
	A_51_P416858	Myl1	17901	NM_021285	-2.439735	-0.744762	-2.432315	-0.877047
	A_52_P273120	Myl1	17901	NM_001113387	-2.805207	-0.94129	-2.780981	-0.625612
	A_52_P383572	Mylpf	17907	NM_016754	-2.776151	-1.61983	-3.198656	-1.921084
	A_52_P622418	Tnnt3	21957	NM_011620	-2.496427	-1.712659	-3.347478	-2.394351
	A_51_P345699	Tnnt3	21957	NM_011620	-2.366695	-1.713335	-3.294333	-2.158767
	A_51_P353232	Tnnc2	21925	NM_009394	-2.730774	-1.299901	-3.178287	-1.496824
	A_51_P332742	Srl	106393	NM_175347	-2.4321385	-1.753972	-3.049895	-2.608059
	A_51_P167668	Myh3	17883	NM_001099635	-2.4449535	-1.5499425	-2.9557595	-3.2844955
	A_52_P247953	Myh3	17883	NM_001099635	-2.627133	-1.582406	-3.115868	-3.613156
	A_51_P328539	Eno3	13808	NM_007933	-0.4840235	-0.265624	-0.389728	-0.4711575
	A_51_P359424	Smtnl1	68678	NM_024230	ND	ND	ND	ND
	A_51_P500156	Pvalb	19293	NM_013645	-1.234949	-1.1823147	-1.3341507	-0.6434444
	A_51_P199104	Casq1	12372	NM_009813	ND	ND	ND	ND
	A_52_P602147	Myh4	17884	NM_010855	-2.3604716	-1.335844	-2.6485313	-3.6337035
	A_51_P338072	Myh4	17884	NM_010855	-2.0007177	-1.120715	-2.3071657	-2.552921
	A_51_P364788	Myh1	17879	NM_030679	-1.097558	-0.311186	-1.2653315	-1.2842553
	A_52_P257638	Myh1	17879	AK009974	-0.860621	0.0499266	-0.9201314	-0.8699824
	A_52_P656699	Actn3	11474	NM_013456	-1.67295	-1.085425	-1.909024	-1.397088
	A_52_P305767	Myoz1	59011	NM_021508	-0.6049493	-0.629922	-0.735308	-0.554799
	A_51_P379750	Myoz1	59011	NM_021508	-0.711257	-0.50732	-0.8397976	-0.4770623
	A_51_P355122	Tnni2	21953	NM_009405	-1.985113	-1.144532	-2.388581	-1.474189
	A_51_P395652	Myh2	17882	NM_001039545	-2.02132	-0.315837	-2.3214674	-3.1788
	A_51_P225134	Myh2	17882	NM_001039545	-0.4902865	-0.3009675	-0.5164819	-0.6109625
	A_52_P552036	Myh2	17882	NM_001039545	-0.4685898	-0.0773468	-0.5300441	-0.6976523
	A_51_P419319	Aqp4	11829	NM_009700	ND	ND	ND	ND
	A_51_P144090	Slc16a3	80879	NM_030696	1.015424	0.390648	0.878618	0.657137
	A_51_P160198	Nos1	18125	NM_008712	ND	ND	ND	ND
	A_51_P264495	Pgam2	56012	NM_018870	-2.444136	-1.188136	-2.736102	-1.482386
Slow Type	A_51_P476725	Myl6b	216459	AK017625	0.1883934	0.0480844	0.0168554	0.4252638
	A_51_P413130	Myl4	17896	NM_010858	-2.301066	-1.453581	-2.757282	-3.497861
	A_52_P544523	Myl4	17896	NM_010858	-2.273345	-1.532838	-2.746343	-3.510601
	A_51_P362627	Tnnt1	21955	NM_011618	-1.318493	-0.475367	-1.620889	-0.619345
	A_52_P657360	Tnni1	21952	NM_021467	-2.533049	-1.685021	-3.06634	-2.476817
	A_51_P435704	Tnni1	21952	NM_021467	-2.554234	-1.676488	-3.048237	-2.43687
	A_51_P338746	Prdm1	12142	NM_007548	ND	ND	ND	ND
	A_51_P491648	Vgll2	215031	NM_153786	-2.355175	-0.46919	-2.081245	-0.832746
	A_52_P569001	Atp2a2	11938	NM_001110140	0.332401	-0.155704	0.009503	-0.5345095
	A_51_P392459	Myl2	17906	NM_010861	ND	ND	ND	ND